Wnt secretion, trafficking and pathway activation: towards a structure-based understanding and modulation



Luca Vecchia

Medical Sciences Doctoral Training Centre

Kellogg College

University of Oxford

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

Doctor of Philosophy in Structural Biology

at the University of Oxford

Hilary Term 2021

Abstract

The Wnt pathway is a key morphogen route during embryonic development, regulating tissue organization, patterning and organ growth. In adult life, when maintained in a homeostatic range it controls stem cell self-renewal and regenerative processes, while its dysregulation promotes diseases like cancer and neurodegeneration.

A peculiar feature of the Wnt protein is its lipid modification by a palmitoleic acid, crucial for receptor engagement and functionality at the target cells. This acylation, however, reduces Wnt solubility, dictating the need for a dedicated machinery upstream of receptor engagement, to shield the lipid from the aqueous environment and promote the formation of concentration gradients.

The work presented here focused first on the elucidation of the structural and molecular mechanisms allowing Wnt to overcome this “solubility problem”. In particular, solving the structure of the *Drosophila melanogaster* heparan-sulfate proteoglycan Dlp in complex with a lipidated Wnt peptide, I showed that this class of membrane-anchored proteins, glypicans, has a novel role in the extracellular shielding of the Wnt lipid. The functional relevance of the interaction in modulating the extracellular Wnt spreading was confirmed by mutational and *in vivo* experiments in flies. The work has been published in *Nature*.

Additionally, upon thorough troubleshooting, I optimized a protocol for the recombinant production of the transmembrane protein Wntless - the main intracellular Wnt carrier - embedded in polymer-based nanoparticles, to use in future structural studies.

Moreover, given its importance in pathway activation, I also focused on the Wnt lipidation for therapeutic purposes in Alzheimer's disease. To the goal, a structure-based drug development campaign was set up to find inhibitors of Notum, an extracellular modulator that specifically removes the Wnt palmitoleate from the morphogen, therefore downregulating the signalling. The work led to the identification of a patented compound with optimal pharmacological properties, currently undergoing *in vivo* preclinical testing.

Acknowledgements

I am especially grateful to Yvonne not only for the constant and careful supervision over the course of my D.Phil. and for the ability to fine-tune the work on different projects to finally reach a wonderful result, but also for her commitment and fruitful discussions on the future perspectives for my career. Many thanks also to my day-to-day supervisor, Yuguang, for his supervision on my projects, for involving me in the Notum work and for the practical tips and help during the experimental work. In this sense, a great and particular thanks also to Tomas for the additional support and informal supervision.

Thanks to Gergely, Vitul, James, Reinis and Daniel for having shared and made enjoyable the daily life in the lab.

A huge acknowledgement to the crucial work of the support team: Karl, for his invaluable contribution to several projects, as detailed below; Margaret, for lab assistance and friendly chats; Weixian for tissue culture help; and Tom, for assistance with crystallization.

Finally, a special thanks to Mattia, for having shared together the journey of this D. Phil. and having represented a constant source of peer support, and Luigi, for our discussions about science and life and for his help and training in crystallography.

Declaration of work

All the work presented in this thesis is my own, except where otherwise explicitly stated in the respective chapters. The contributions are also summarized here:

Chapter 3 is based, including figures, on the article: McGough, I. J.; Vecchia, L.; Bishop, B.; Malinauskas, T.; Beckett, K.; Joshi, D.; O'Reilly, N.; Siebold, C.; Jones, E. Y.; Vincent, J. P., **Glypicans shield the Wnt lipid moiety to enable signalling at a distance**. Nature 2020 Sep;585(7823):85-90, of which I am co-first author.

Some sentences are adapted from the article or quoted *verbatim*. The functional *in vivo* fly experiments, and the related figures, were performed and designed by Dr Ian McGough (Vincent lab, Francis Crick Institute London). Dr Karl Harlos from the Division of Structural Biology (University of Oxford) assisted with mounting and cryo-cooling of crystals. Prof Christian Siebold and Dr Benjamin Bishop from the Division of Structural Biology contributed to the expression and purification of of *Drosophila* Dally.

Some part of the text and some figures of **Chapter 5** are based on the following two articles and patent, of which I am, respectively, co-author and co-inventor: Atkinson BN, Steadman D, Zhao Y, Sipthorp J, Vecchia L, Ruza RR, Jeganathan F, Lines G, Frew S, Monaghan A, Kjær S, Bictash M, Jones EY, Fish PV. **Discovery of 2-phenoxyacetamides as inhibitors of the Wnt-depalmitoleating enzyme NOTUM from an X-ray fragment screen.**

Medchemcomm. 2019 Apr 29;10(8):1361-1369; Mahy W, Patel M, Steadman D, Woodward HL, Atkinson BN, Svensson F, Willis NJ, Flint A, Papatheodorou D, Zhao Y, Vecchia L, Ruza RR, Hillier J, Frew S, Monaghan A, Costa A, Bictash M, Walter MW, Jones EY, Fish PV. **Screening of a Custom-Designed Acid Fragment Library Identifies 1-Phenylpyrroles and 1-Phenylpyrrolidines as Inhibitors of Notum Carboxylesterase Activity.** J Med Chem. 2020 Sep 10;63(17):9464-9483; **"Compounds"** (patent application n. PCT/EP2019/073177, filing date 30.08.2019).

The medicinal chemistry experiments in the hit-to-lead optimization process were performed by the lab of Prof Paul Fish (University College London), while the *in vivo* data presented in **Figure 5.18** were obtained by Dr Sarah Jolly and colleagues (University College London). The initial fragment screening mentioned in **Figure 5.3** was performed by Dr Yuguang Zhao from the Division of Structural Biology. The hit-to-lead optimization of the different series was a collaborative effort of the lab involving myself, Dr Yuguang Zhao, Dr James Hillier, and Mr Reinis Reinholds Ruza. The list of the compounds on which I have worked is listed in Appendix **Table 7.1**. In particular, I have personally carried out the whole pipeline from crystal growth to compound soaking optimization, data collection, structure solution and analysis for all the compounds presented in **Chapter 5**, and the same holds for the compounds highlighted in yellow in **Table 7.1**. Furthermore, for the other compounds listed in the latter table, I have been differentially involved at various levels of the aforementioned steps. Dr Karl Harlos helped with mounting and cryo-cooling of crystals.

Table of Contents

TABLE OF FIGURES 12

TABLE OF TABLES 16

LIST OF ABBREVIATIONS 18

CHAPTER 1 - INTRODUCTION 25

THE WNT PATHWAY IN HEALTH AND DISEASE 25

GENERAL OVERVIEW OF THE SIGNALLING PATHWAY 29

THE WNT MORPHOGEN 33

THE WNT PATHWAY UPSTREAM OF RECEPTOR ENGAGEMENT 38

The source of Wnt: the process of Wnt secretion *38*

Shielding the lipid: the Wnt solubility problem *43*

Wnt extracellular modulators *52*

AIMS AND RESEARCH QUESTIONS 57

CHAPTER 2 - MATERIALS AND METHODS **60**

GENERAL METHODS COMMON TO ALL CHAPTERS 60

Polymerase chain reaction and site-directed mutagenesis *60*

Agarose gel electrophoresis and gel extraction *62*

DNA enzymatic digestion, ligation and transformation *62*

Purification of plasmid DNA *64*

SDS-PAGE *65*

Western Blot 66

METHODS SPECIFIC TO CHAPTER 3 67

Construct design 67

Generation of lentiviral-transduced HEK293S GnTI- TetR cell lines expressing Dlp core 69

Small-scale expression of Dlp mutants 70

Large scale expression and purification of Dlp core wild-type and Dally core 70

Large scale expression and purification of Dlp core mutants 72

Biolayer interferometry 72

Fluorescence-detection size-exclusion chromatography 73

Crystallization, data collection and structure determination 74

METHODS SPECIFIC TO CHAPTER 4 76

Construct design 76

Generation of stably-transfected Drosphila S2 cell lines 81

Generation of lentiviral-transduced HEK293S GnTI– TetR cell lines 82

Small-scale protein expression screening 83

Wntless full-length/transmembrane protein detergent screening 84

Large-scale protein expression and purification 86

SMALP extraction and purification of Wntless full-length/transmembrane protein 90

METHODS SPECIFIC TO CHAPTER 5 91

Construct design 91

Large scale expression and purification of Notum 91

Crystallization of Notum, soaking of inhibitors and structure determination 92

METHODS SPECIFIC TO CHAPTER 6 94

Construct design, expression, purification and preliminary crystallization of human Gpc6 94

Preliminary BLI analysis of Dlp^{core} interaction with Hedgehog palmitate-mimicking peptide *95*

CHAPTER 3 - UNDERSTANDING WNT EXTRACELLULAR SPREADING: THE ROLE OF GLYPICANS. 97

IDENTIFICATION OF THE WNT CARRIER MOLECULE IN *DROSOPHILA* 97

PROTEIN EXPRESSION AND PURIFICATION 102

BIOPHYSICAL STUDIES ON GLYPICAN CORE - WNT LIPID INTERACTIONS 108

CRYSTAL STRUCTURE OF *DROSOPHILA* GLYPICAN DLP IN COMPLEX WITH A PALMITOLEOYLATED WNT PEPTIDE 114

FROM STRUCTURE TO FUNCTION: DESIGN OF MUTATIONS TO IMPAIR WNT BINDING TO DLP^{CORE} 128

DISCUSSION 133

What we have learnt *133*

Glypicans in Wnt morphogen diffusion *134*

Open questions *137*

CHAPTER 4 - WNT INTRACELLULAR TRANSPORT AND SECRETION: BIOCHEMICAL STUDIES ON THE CARRIER WNTLESS 140

THE WNT-BINDING 1ST EXTRACELLULAR LOOP OF WNTLESS PRODUCED AS SOLUBLE RECOMBINANT PROTEIN: EXPRESSION AND PURIFICATION ATTEMPTS 142

FULL-LENGTH TRANSMEMBRANE WNTLESS: EXPRESSION AND PURIFICATION ATTEMPTS. 151

Isolation of human Wls transmembrane protein in SMALPs *156*

DISCUSSION 157

Addendum: discussion points from the structure of human Wls-human Wnt8a *161*

CHAPTER 5 - TARGETING THE WNT EXTRACELLULAR MODULATORS 168

BACKGROUND AND RATIONALE FOR TARGETING THE WNT PATHWAY IN ALZHEIMER'S DISEASE 168

The Wnt pathway in the adult brain *168*

Wnt pathway in the ageing brain and role in Alzheimer's disease *169*

Targeting Notum *172*

EXPRESSION, PURIFICATION AND CRYSTALLIZATION OF NOTUM 173

FROM FRAGMENT SCREENING TO LEAD CANDIDATE: OVERVIEW OF SERIES DEVELOPMENT AND REPRESENTATIVE CRYSTAL STRUCTURES 175

Phenoxyamides *178*

Heteroaryl compounds *186*

Carboxylic acid compounds *195*

DISCUSSION 202

CHAPTER 6 - GENERAL DISCUSSION, FUTURE PERSPECTIVES AND CONCLUDING REMARKS 208

HANDOVER OF WNT 209

GLYPICANS: A MORE GENERAL HUB FOR LIPID-MODIFIED MORPHOGENS? 213

THERAPEUTIC TARGETING OF THE WNT PATHWAY 219

APPENDIX 223

LIST OF PUBLICATIONS DERIVED FROM THE WORK PRESENTED IN THIS THESIS 227

COPYRIGHT INFORMATION 228

REFERENCES 230

Table of Figures

Figure 1.1. Canonical Wnt signalling activation cascade at the target cells..... 31

Figure 1.2. Structure of Wnt in complex with the extracellular domain of the canonical receptor Frizzled..... 35

Figure 1.3. Overview of the Wnt secretory route. 40

Figure 1.4. The Wnt intracellular carrier Wntless..... 42

Figure 1.5. Models of Wnt transport on subcellular structures..... 45

Figure 1.6. Model of Wnt transport on cytonemes. 48

Figure 1.7. The glypican family of heparan sulphate proteoglycans. 51

Figure 1.8. Structure of the Wnt extracellular deacylase human Notum and modes of Wnt binding. 56

Figure 3.1. *In vivo* studies in *Drosophila* ruled out the involvement of lipoprotein particles, exosomes, Swim and multimers in Wnt transport..... 101

Figure 3.2. Schematics and rationale for Dlp^{core} construct design. 104

Figure 3.3. Western Blot on small-scale expression of the Dlp^{core} constructs from lentiviral-transduced HEK293S GnTI(-) TetR cells. 105

Figure 3.4. Representative size-exclusion chromatography profiles for the two Dlp^{core} constructs and $Dally^{core}$..... 107

Figure 3.5. Biophysical analysis confirmed a direct interaction between Dlp^{core} and the Wnt lipid..... 112

Figure 3.6. *In vivo* data confirmed the differential trapping activity between glypicans members..... 113

Figure 3.7. The glypican CRD is not the Wnt binding domain. 115

Figure 3.8. Highest diffracting crystal obtained for Dlp^{core} in complex with Wnt7a peptide..... 117
Figure 3.9. Structure of Dlp^{core} in complex with palmitoleoylated Wnt peptide. 121
Figure 3.10. Secondary structure nomenclature for the Dlp^{core}-Wnt7a peptide complexed structure. 122
Figure 3.11. Several types of omit maps showed and supported the modelling of the electron density in the binding pocket as PMS from the palmitoleoylated Wnt peptide..... 125
Figure 3.12. Structure-guided mutations impaired Dlp lipid-binding ability. 131
Figure 3.13. Mutational analysis showed that the difference between Dlp and Dally is not fully attributable to the cavity environment. 132
Figure 3.14. Model of Dlp functional interaction with Frizzled receptors. 135
Figure 4.1. Multiple sequence alignment of the 1st extracellular loop of Wntless family members from different species. 143
Figure 4.2. Human Wls 1st extracellular loop constructs were retained in the cell. 144
Figure 4.3. A *Drosophila* Wls loop construct could be expressed in *Drosophil*a cells, but the protein aggregated. 147
Figure 4.4. Human Wls loop cysteine truncations promoted expression in mammalian cells, but the protein still aggregated. 149
Figure 4.5. Switch to inducible expression improved the purification behaviour of human Wls loop. 151

Figure 4.6. Human Wls transmembrane constructs well-behaved in small-scale test, but not upon expression scaling up. 154
Figure 4.7. Detergent screening by FSEC on human Wls aa. 1-495. 155
Figure 4.8. Extraction of human Wls aa. 1-541 in SMALPs improved protein purity and yield recovered. 157
Figure 4.9. Structure of human Wntless in complex with human Wnt8a..... 166
Figure 5.1. The blood-brain-barrier (BBB) and its involvement in neurodegeneration. 172
Figure 5.2. Representative purification and crystallization of Notum. 175
Figure 5.3. Overview of fragment screening and chemical series development for Notum inhibitors..... 177
Figure 5.4. Phenoxyamide series medicinal chemistry optimization..... 180
Figure 5.5. Overview of the structures of ARUK3000763 (marine) and ARUK3000400 (wheat)..... 181
Figure 5.6. Structure of Notum in complex with ARUK3000763. 183
Figure 5.7. Structure of Notum in complex with ARUK3000400. 184
Figure 5.8. Initial hit from fragment screening for the heteroaryl series. 186
Figure 5.9. Overview of the aminooxadiazoles ARUK3001213, 1233 and 1244. 188
Figure 5.10. Structure of Notum in complex with ARUK3001233. 190
Figure 5.11. Structure of Notum in complex with ARUK3001213. 191
Figure 5.12. Structure of Notum in complex with ARUK3001244. 192
Figure 5.13. Variable positioning of the β11-β12 loop containing the catalytic H389..... 193

Figure 5.14. Physicochemical and *in vivo* mouse pharmacokinetics properties of lead compound ARUK3001185. 194
Figure 5.15. Carboxylic acid hits and pyrrole-3 scaffold medicinal chemistry optimization. 196
Figure 5.16. Structure of Notum in complex with ARUK3001246. 200
Figure 5.17. Comparison between the pyrrole-3 and pyrrolidine-3 carboxylic acids. 201
Figure 5.18. Notum expression in mouse and human brain in physiological and pathological conditions. 204
Figure 6.1. The Dlp-Wnt peptide structure and the Hedgehog pathway. 215
Figure 6.2. Preliminary results for the study of glypican role in the Hedgehog pathway. 218

Table of Tables

Table 2.1. Typical PCR reaction volumes and concentrations..... 60

Table 2.2. Typical PCR reaction conditions. 61

Table 2.3. First set of human Wntless 1st extracellular loop constructs designed. 77

Table 2.4. Drosophila Wntless 1st EC loop constructs. 79

Table 2.5. Cysteine-deleted versions of the human Wls 1st EC loop (in yellow the expressing constructs)..... 79

Table 2.6. Screened constructs for human Wls transmembrane protein (highlighted the expressing constructs). TM = Transmembrane region; EC = extracellular. 80

Table 2.7. Detergents conditions screened. 84

Table 2.8. Summary of conditions tested for the purification of *Drosophila* Wntless 1st extracellular loop..... 87

Table 2.9. Summary of conditions tested for the purification of human Wntless 1st extracellular loop 88

Table 3.1. Data collection and refinement statistics..... 127

Table 5.1. Data collection and refinement statistics – phenoxyamides..... 185

Table 5.2. Data collection and refinement statistics – heteroaryl compounds 189

Table 5.3. Data collection and refinement statistics – carboxylic acids 199

Table 7.1. Compounds in the Notum drug development project on which I have worked in collaboration with Dr Yuguang Zhao, Dr James Hillier and Mr Reinis

Reinholds Ruza (the compounds highlighted are those for which I carried out the whole pipeline)..... 223

List of Abbreviations

Å: Angstrom

aa.: amino acids

AD: Alzheimer's disease

AP2: Clathrin adaptor protein 2

APC: Adenomatous polyposis coli

APP: Amyloid precursor protein

ATP: Adenosine triphosphate

AUC: Analytical ultracentrifugation

AUC (Figure 5.14): Area under curve

BBB: Blood-brain-barrier

BLI: Bio-layer interferometry

BSA: Bovine serum albumin

β-TRCP: β-transducin repeats-containing protein E3 ubiquitin ligase

CBP: cAMP-response element binding protein (CREB)-binding protein

CFS: cerebrospinal fluid

CHES: N-cyclohexyl-2-aminoethanesulfonic acid

CHS: Cholesteryl hemisuccinate

CK-1: Casein kinase 1

cLogP: Logarithm of octanol/water partition coefficient

Cmax: maximum plasma concentration

CMC: Critical micellar concentration

CMV: Cytomegalovirus

CNS: Central nervous system

CNS MPO: Central nervous system multiparameter optimization

CoA: Coenzyme A

CRD: Cysteine-rich domain

CSS: Complex significance score

CYMAL-6: 6-cyclohexyl-1-hexyl-β-D-maltoside

CYMAL-7: 7-cyclohexyl-1-hexyl-β-D-maltoside

DALLY: Division abnormally delayed protein

DDM: n-dodecyl-β-D-maltoside

DDTM: n-dodecyl 1-thio-β-D-maltoside

DKK1: Dickkopf Wnt signalling pathway inhibitor 1

DLP: Dally-like protein

DMNG: Decyl maltose neopentyl glycol

DMSO: Dimethyl sulfoxide

dNTPs: Deoxynucleoside triphosphates

DTM: n-decyl-β-D-thiomaltopyranoside

DTT: Dithiothreitol

EC_{50}: Half maximal effective concentration

ECS: Extracellular spaces

EDTA: Ethylenediaminetetraacetic acid

EM: Electron microscopy

ER: Endoplasmic reticulum

ESCRT: Endosomal sorting complexes required for transport

FAP: Familial adenomatous polyposis

FBS: Fetal bovine serum

FSEC: Fluorescence-detection size-exclusion chromatography

FZ: Frizzled receptor

GAG: Glycosaminoglycan

GFP: Green fluorescent protein

GnTI: N-acetyl-glucosaminyltransferase I

GPCR: G protein-coupled receptor

GPI: Glycosylphosphatidylinositol

GSK-3: Glycogen synthase kinase 3

HA: Hemagglutinin

HEK293: Human embryonic kidney 293 cells

HH: Hedgehog

HLM: Human liver microsomes

HPLC: High-performance liquid chromatography

HRP: Horseradish peroxidase enzyme

HRV: Human rhinovirus

HS: Heparan sulphate

IC_{50}: Half maximal inhibitory concentration

IMAC: Immobilized metal affinity chromatography

IRES: Internal ribosome entry site

ISF: interstitial fluid

K: Kelvin

Kd: Dissociation constant

kDa: Kilodalton

LB: Lysogeny broth

LCP: Lipidic cubic phase

LDS: Lithium dodecyl sulfate

LGR4/5/6: Leucine-rich repeat-containing G-protein coupled receptor 4/5/6

LiPE: Lipophilic efficiency

LLE: Ligand-lipophilicity efficiency

LMNG: Lauryl maltose neopentyl glycol

LogD7.4: octanol/water distribution coefficient for ionizable compounds at pH 7.4

LPP: Lipoprotein particle

LRP1: LDL receptor-related protein 1

LRP5/6: Low-density lipoprotein receptor-related protein 5/6

LSQ: Least- squares superposition

LTP: Long-term potentiation

MBOAT: Membrane bound O-acyl transferase

MCS: Multiple cloning site

MDR1-MDCK: Madin Darby canine kidney cells overexpressing the MDR1 gene

MLM: Mouse liver microsomes

MME: Monomethylether

MMT: Malic Acid, MES and Tris

MMTV: Mouse mammary tumour virus

mPPB: Mouse plasma protein binding

MVB: Multivescicular bodies

MWCO: Molecular weight cut-off

NADPH: Nicotinamide adenine dinucleotide phosphate

NDSB-201: Non-detergent sulfobetaine 201

NPC1: Niemann-Pick disease Intracellular Cholesterol Transporter 1

NPC2: Niemann-Pick disease Intracellular Cholesterol Transporter 2

ObG: n-octyl-β-D-glucopyranoside

OGNG: Octyl glucose neopentyl glycol

OPTS: Octanoyloxypyrene-1,3,6-trisulfonate

PAM: Palmitoleic acid

PBS: Phosphate-buffered saline

PCP: Planar cell polarity

PCR: Polymerase chain reaction

PEG: Polyethylene glycol

PEI: Polyethylenimine

PISA: Proteins, Interfaces, Structures and Assemblies server

PMS: Palmitoleoylated serine

po: orally admistered

POPG: 1-palmitoyl-2-oleoyl-3-sn-phosphatidyl glycerol

PS1: Presenilin-1

RAGE: receptor for advanced glycosylation end products

RBC: red blood cell

REST: RE1 silencing transcription factor

RMSD: Root mean square deviation

RNF43: Ring finger protein 43

ROR: Receptor tyrosine kinase like orphan receptor

ROS: reactive oxygen species

rpm: revolutions per minute

RSPO: R-spondin

RYK: Receptor like tyrosine kinase

SAR: Structure-activity-relationship

SCD: Stearoyl-CoA desaturase

SDS: Sodium dodecyl sulfate

SDS-PAGE: Sodium dodecyl sulphate–polyacrylamide gel electrophoresis

SEC: Size-exclusion chromatography

sFRPs: secreted Frizzled-related proteins

SMA: Styrene-maleic acid polymer

SMALP: Styrene-maleic acid lipid particle

SMILES: Simplified molecular input line entry system

SWIM: Secreted Wg-interacting molecule

t1/2: half-life

TAE: Tris-acetate-ethylenediaminetetraacetic acid

TCF/LEF: T-cell factor/lymphoid enhancer factor transcription factors

TetR: Tet repressor proteins

TLS: Translation-libration-screw

Tmax: time at which the maximum plasma concentration is achieved

UDM: n-undecyl β-D-maltoside

UDTM: n-undecyl β-D-1-thiomaltopyranoside

UV: Ultraviolet

V: Volt

VSP4: Vacuolar sorting protein 4

v/v: volume/volume

WB: Western blot

WG: Wingless

WIF-1: Wnt inhibitory factor 1

WLS: Wntless

w/v: weight/volume

ZNRF3: Zinc and ring finger 3 E3 ubiquitin ligase

Chapter 1 - Introduction

The Wnt pathway in health and disease

Already present in sponges, and evolutionarily conserved in all the animal kingdom[1,2], the Wnt pathway is a fundamental cell-cell communication signalling route regulating stem cell renewal, cell proliferation and differentiation. In particular, the Wnt protein themselves, which activate the signalling cascade, can be defined as morphogenetic growth factors that stimulate cell proliferation and act as directional cues to control tissue patterning and organization[3].

The pathway exerts its effect from the early stages of the embryonic life, both in vertebrates and in invertebrates. The name itself – Wnt stands for Wingless-type MMTV integration site – has its origin from a developmental defect shown to cause lethal segmentation defects in *Drosophila* embryos, involving ablation of the *Wingless* (Wg) gene (the rest of the name relates instead to the oncogenic effect caused by the integration of the mouse mammary tumour virus (MMTV) in the locus of the *int1* gene, the mouse homologue of Wingless)[4,5]. Wnt is also responsible for the formation and dorsoventral/antero-posterior specification of the body axis of several animals, and ectopic expression of its components in *Xenopus* embryos has been shown to determine duplication of the dorsal axis and lead for example to two-headed frogs[6].

In the adult life, the pathway keeps a key role especially in the maintenance of the stem cell niches in several organs. Although the best known and studied examples are the intestinal, bone, and hair follicle niches, lineage tracing

experiments have reported the presence of stem cells positive for the membrane receptor Lgr5 (Wnt target gene and pathway component) also in the stomach, pancreas, liver, mammary gland, kidney and ovary[3]. Moreover, activation of the Wnt pathway can maintain the pluripotency and stem-cell phenotype of embryonic stem cells in culture, and allow establishment and growth of complete organoids from single, dissociated Lgr5+ stem cells[7]. Thanks to this activity, Wnt can both functionally oversee the renewal of fast-turnover tissues, like the gastrointestinal tract, and guide wound healing and tissue regeneration after damage in homeostatic conditions.

Consequently, due to the key role of this family of proteins in several functions in human physiology, mutations and alterations in the core components of the pathway and their regulators have been associated with a plethora of hyperproliferative and degenerative abnormalities. While the latter, with a specific focus on neurodegeneration, will be treated in **Chapter 5** in the context of the drug development project to boost Wnt signalling in Alzheimer's disease, the many roles and crucial relevance of the pathway in the oncological setting will be discussed here.

The most prominent example of Wnt pathway implication in cancerogenesis is represented by colorectal cancers. The first report dates back to 1991, when the genetic cause of a hereditary disease called Familial Adenomatous Polyposis (FAP) – characterized by the development of a large number of intestinal polyps in the early adulthood, some eventually acquiring malignant traits – was found in inactivating mutations in the *Adenomatous Polyposis Coli* (APC) gene, a negative regulator of Wnt signalling[8,9]. Loss of APC was then discovered even in patients

with non-familiar colorectal cancer[10]. Over time, further associations were uncovered, resulting in genetic alterations of Wnt regulators to be drivers of over 92% of sporadic colorectal cancers according to the 2012 report from The Cancer Genome Atlas[11]. Despite the vast majority of mutations mapping to APC (>70%), other pathway components have been implicated, including the key cytoplasmic signal transducer β-catenin and the intracellular negative regulator Axin (5-6% of cases) – see below for an explanation of the molecular role of these proteins. Fusions and mutational events in the R-spondin/Lgr5/Rnf43 axis, a membrane complex involved in Wnt signalling regulation, account for an additional 18% of cases[12].

Wnt importance in tumorigenesis is however not restricted to colorectal cancer. Activating mutations in the pathway are present, among others, in hepatocellular, ovarian, breast, stomach, pancreas and kidney carcinomas[13], and Wnt aberrations affect all the phases of the natural history of tumours. The widespread role of the pathway in the latter sense can be particularly highlighted if its dysregulations are mapped onto the landmark conceptual framework of the "hallmarks of cancer"[14] - sustained cell proliferation, *de novo* angiogenesis, genome instability, evasion from growth suppression, resistance to cell death, replicative immortality, deregulated cellular energetics, activation of invasion and metastasis, evasion from immune system control and tumour-promoting inflammation. These are defined as pathological capabilities acquired by tumours, in a context and tissue-specific manner, to sustain malignant growth and escape therapy, and are useful to encompass tumour diversity and find a logical organizing principle in the disease. A direct or indirect role for the Wnt

pathway have been reported for all of these 10 hallmarks; particularly significant, from a therapeutic perspective, is the Wnt pathway's ability to favour the acquisition of an invasive phenotype through multiple mechanisms, including epithelial-to-mesenchymal transition and interaction with the tumour microenvironment to promote cytokine secretion, inflammation, and matrix degradation by production of extracellular metalloproteinases[15].

Finally, the pathway has also been implicated in the plasticity of the so-called "drug-tolerant persister cells", which form the minimal residual disease surviving pharmacological treatment and are responsible for therapy resistance and tumour relapse. In skin basal cell carcinomas treated with the Hedgehog pathway inhibitor vismodegib, for example, a subpopulation of cells transdifferentiates and switch to a Hedgehog-independent phenotype, where the cellular growth is sustained by Wnt signalling. Upon treatment discontinuation, they then revert to their original identity and to a Hedgehog-controlled proliferation, thereby promoting cancer recurrence[16].

Although, as discussed so far, hyperactivation of the pathway is usually associated with malignancies, it is however worth keeping in mind that the picture is more nuanced than a simple linear relationship between elevated signalling and cancer progression, reflecting the deep biological complexity, redundancy, and regulatory cross-talks existing among the pathway effectors. Thus, increased Wnt levels also correlate with improved prognosis and prolonged survival in some medulloblastoma, breast, ovarian and prostate cancers[13], and this dictates the need of a careful mechanistic understanding and biological analysis of the local disease context when devising strategies to target the pathway.

Nevertheless, based on this concise overview, it is understandable how this signalling route has attracted much attention as a promising option for therapy. In this sense, while a key focus of both the fundamental and translational research has been initially represented by the “downstream” signalling cascade, (from receptor engagement at the target cells to activation of the transcriptional programme) more proximal events are also emerging as key regulatory nodes and potential drug targets – and represent the focus of the research discussed in this thesis.

In the following sections, therefore, a brief overview of the downstream pathway, still necessary to fully appreciate the molecular basis of the pathway and compare the envisaged therapeutic strategies, will be first presented. The attention and the core of the analysis will however then mainly switch first to the structural and molecular features of the Wnt ligands themselves, that makes Wnts such a peculiar family of proteins, and then to the events regulating the morphogen secretion at the producer cells, its extracellular spreading to reach the target cells, and the extracellular modulation of its activity.

General overview of the signalling pathway

From a molecular point of view, the “canonical” downstream activation of the signalling starts when the Wnt ligand engages, on the target cells, members of the Frizzled (Fz) class of 7-pass membrane receptors and promote the formation of a complex with the single-pass membrane co-receptors Lrp5/6[3].

This in turn causes a conformational rearrangement triggering a series of phosphorylation events, with the most prominent determined by the glycogen synthase kinase 3 (GSK-3) at the serine of a conserved PPPSP motif in the cytoplasmic tail of Lrp5/6[17]. The latter favours recruitment at the plasma membrane of the scaffold protein Axin. This protein is part of a complex called "destruction complex", also composed of the already cited APC, of GSK-3 and of the further kinase casein kinase 1 (CK-1).

In the absence of Wnt **(Figure 1.1 a**), the destruction complex binds the intracellular protein β-catenin and phosphorylates it through the action of GSK-3 and CK-1; this phosphorylation allows recognition by β-Trcp, an E3 ubiquitin ligase which subsequently causes β-catenin ubiquitination and degradation. Upon receptor engagement from Wnt and activation of signalling (**Figure 1.1 b**), the recruitment of Axin at the plasma membrane determines the disruption of the aforementioned degrading activity of the destruction complex; β-catenin can thus shuttle to the nucleus, where it engages DNA-bound TCF/LEF transcription factors and activates target genes – mostly in a tissue- and/or development stage-specific manner, but with the gene *Axin 2* as a global transcriptional target, therefore used in functional studies as the main indicator of pathway activation.

The precise molecular events underneath the described cascade, in particular how the receptor conformational rearrangement blocks the destruction complex activity, are still only partially elucidated. Proposed models range from disassembly of the destruction complex caused by Axin recruitment by phosphorylated Lrp5/6[18], to relocation of the full, intact destruction complex to the membrane with dissociation of β-Trcp, blocking β-catenin ubiquitinoylation (as in

Figure 1.1 b)[19], to dynamic inhibition of β-catenin phosphorylation[20]. An important role in linking the activated receptor to the destruction complex can also be played by the intracellular scaffold protein Dishevelled: the latter presents a DIX domain which undergoes head-to-tail polymerization upon signalling activation, and two other domains (PDZ and DEP) both interacting with Frizzled; it can therefore act as a binding platform for components of the destruction complex and bridge Frizzled to Lrp5/6 upon receptor engagement[21].

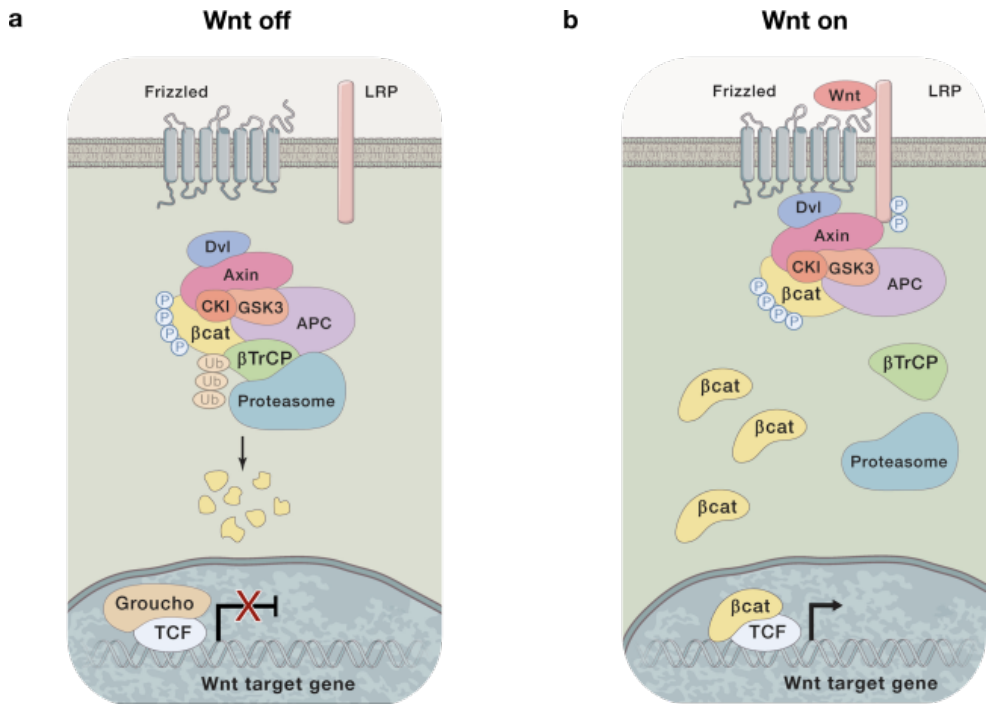


Figure 1.1. Canonical Wnt signalling activation cascade at the target cells. **a)** In the absence of Wnt, β-catenin is phosphorylated and targeted to proteosomal degradation by the action of a complex of proteins known as the destruction complex, having in Axin 1 its core scaffold protein. **b)** When Wnt reaches its target cells, it binds to Frizzled receptors and activates a cascade of events leading to Lrp coreceptor phosphorylation, disruption of the activity of the destruction complex, and intracellular accumulation of unphosphorylated β-catenin. β-catenin can then shuttle to the nucleus and activate downstream transcription. In the displayed model, recruitment of the destruction complex to the site of Lrp phosphorylation promotes the dissociation of the β-Trcp E3

ubiquitin ligase in the context of an otherwise active destruction complex. β-catenin is therefore still phosphorylated but not degraded, and this leads to saturation of the destruction complex, hence accumulation of "active" β-catenin. Alternative molecular mechanisms for β-catenin stabilization have also been suggested, as discussed in the text. The figure is derived from ref. [3].

In addition to this "canonical" way of signal activation, Wnt can also find a plethora of non-Frizzled receptors on the target cells. While the canonical β-catenin-centred pathway oversees cell proliferation, stemness maintenance and cell-fate specification, the non-canonical pathways mainly control events like planar cell polarity, cytoskeleton rearrangement and cell migration. However, this classification and a strict distinction between the two Wnt pathway branches is largely a simplification: many downstream components are shared (e.g. Dishevelled), some Frizzled receptors have also been implicated in the non-canonical pathway, and non-canonical signalling can directly act onto, and modulate, the canonical arm[22]. Whereas a comprehensive review of non-canonical receptors is out of the scope of this **Introduction**, the Ror and Ryk single-pass transmembrane receptors are of particular relevance, and their altered expression has been associated with several haematological and solid malignancies[22].

Finally, it is also worth briefly discussing the role in Wnt signalling of R-spondin/Lgr5/Rnf43, since, as noted above, genetic alterations of this module account, for example, for a significant portion of colorectal cancers. R-spondins are secreted proteins that bind seven-pass membrane receptors Lgr4/5/6 with high affinity, and in turn bridge them to the E3 ubiquitin ligases Rnf43/Znrf3[23,24], whose targets include Frizzled receptors. Formation of the Rspo/Lgr/Rnf43-Znrf3 trimeric complex promotes its endocytosis and degradation, therefore blocking

Frizzled ubiquitylation and enhancing Wnt signalling[25]. Hence, this circuit acts as amplifier and potentiator of Wnt signalling, in particular in the context of adult stem cell niches[3].

The Wnt morphogen

One of the major obstacles in analysing, from a molecular and biophysical perspective, the events of signal initiation derives from the multiple difficulties associated with the expression of the Wnt ligands themselves at levels suitable for structural studies.

Whereas the discovery of the Wnt genes in *Drosophila* and mouse dates back to the 80's, it wasn't in fact until 2003 that the first member of the family (Wnt3a) was purified, with a multi-step, detergent-based process - harbouring relatively low yields[26]; the first crystal structure was then only obtained in 2012[27], followed by a second in 2019[28†]. These two structures, however, are closely related and only reflect a single molecular entity, the complex between the Wnt protein (*Xenopus* Wnt8 for the 2012 structure, human Wnt3 for the 2019 one) and the extracellular Cysteine Rich Domain (CRD) of the canonical receptor Frizzled (mouse Frizzled 8 for both structures). Moreover, in both cases, Wnt and Frizzled were co-expressed and formed a stable complex, devoid of signalling activity,

† A third structure, of the *Drosophila* Wnt ligand WntD, was actually solved in 2013 (ref. 29 Chu, M. L. *et al.* structural Studies of Wnts and identification of an LRP6 binding site. *Structure* **21**, 1235-1242, doi:10.1016/j.str.2013.05.006 (2013)., but this protein appears to be only a distantly related member of the family, which does not signal through the standard Wnt pathway effectors and lacks all the standard post-translational modifications of Wnts, including the biologically central acylation. Moreover, very recently, a structure of human Wnt8a in complex with the carrier protein human Wntless (see below) was also solved, but the latter will be discussed at the end of **Chapter 4** for its recent solution and the relevance to the work presented in that chapter.

from the early steps of the protein production procedure. This dictates caution in the extrapolation of generalizable information and highlights the still unmet need for the structure of an uncomplexed Wnt protein. Nevertheless, the two structures are extremely relevant as prototypes to interpret from a mechanistic perspective many functional findings, and also provide an explanation for the intrinsic difficulties in efficient Wnt protein recombinant expression and purification.

First of all, Wnt in fact presents a complicated folding. As shown for the XWnt8-mFz8 structure, 22 cysteines, conserved in all the 19 Wnt family members of the vertebrate genomes, stabilize a three-dimensional architecture composed of two domains connected by an unstructured, solvent-exposed linker. This grants the protein with the overall shape of a left hand, with a “thumb” and an “index” protruding from the palm, which clamp the Frizzled CRD from two opposite edges, the “site 1” and “site 2” interfaces (**Figure 1.2 a** and **b**).

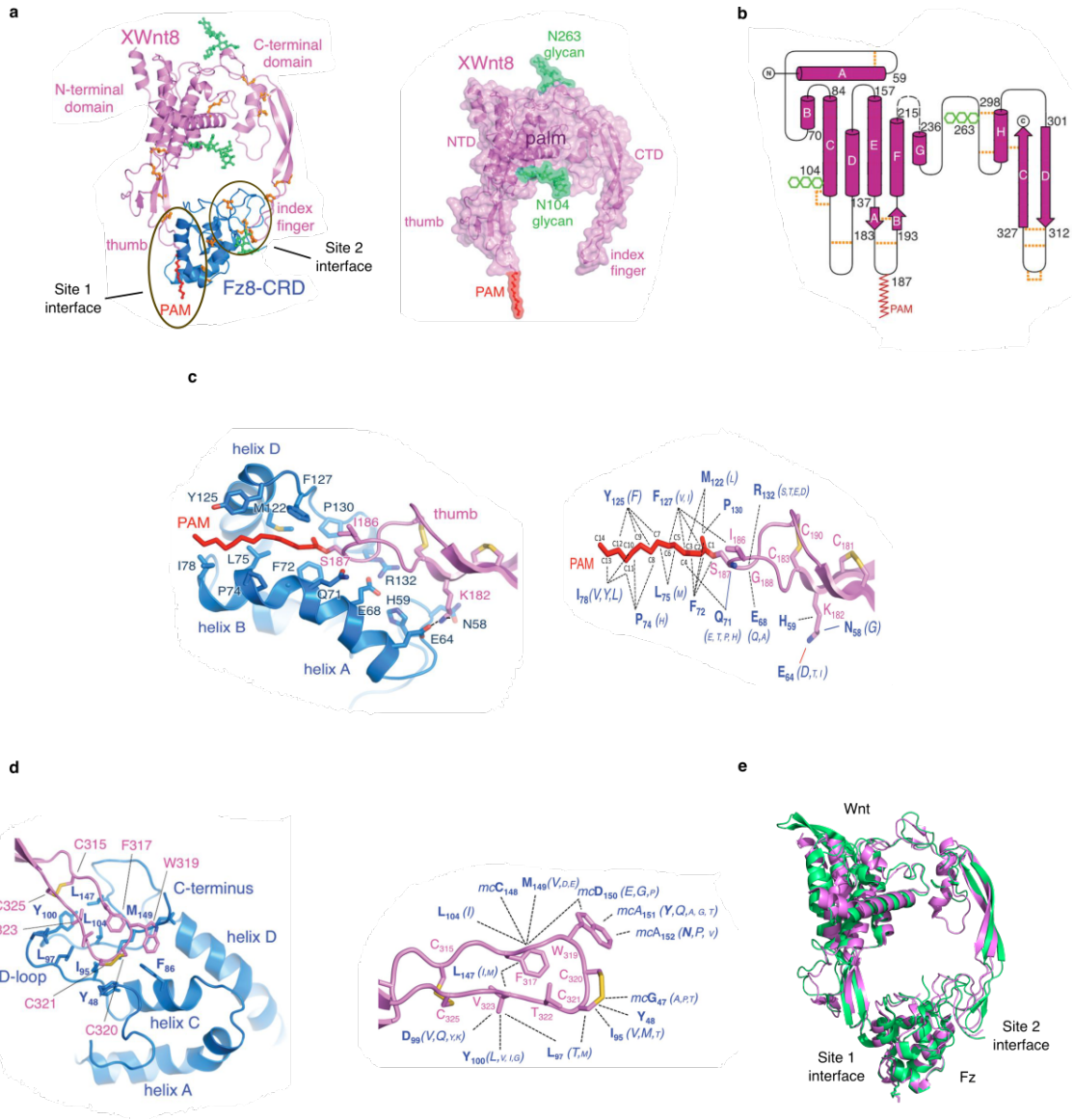


Figure 1.2. Structure of Wnt in complex with the extracellular domain of the canonical receptor Frizzled.
a) Ribbon and surface representation of the structure of Xenopus Wnt 8 in complex with the extracellular cysteine-rich domain of mouse Frizzled 8 (PDB ID: 4F0A). **b)** Diagram of XWnt8 secondary structure. Yellow lines represent disulphide bonds between evolutionarily conserved cysteines. **c)** Close-up view of the site 1 binding interface, with the Frizzled lipid-binding groove. Frizzled residues making contacts with the Wnt lipid or the Wnt thumb are highlighted, and the sequence variation in other Frizzled members at each position is shown in parenthesis. The different font size indicates the relative degree of conservation of that amino acid at that position among Frizzleds. Dashed lines represent hydrogen bonds. **d)** Close-up view of the site 2 binding interface. **e)** Superposition between the XWnt8-mFz8 structure and chains A and B of the hWnt3-mFz8 structure (PDB ID: 6AHY). The similarity between Wnt3 N-terminal

domain and XWnt8 is 0.92 Å root mean square deviation (for 135 Cα atoms), and 0.97 Å (for 37 Cα atoms) for the C-terminal domain[28]. Panel a-d are derived from ref. [27] (Janda, C. Y., Waghray, D., Levin, A. M., Thomas, C. & Garcia, K. C. Structural basis of Wnt recognition by Frizzled. *Science* **337**, 59-64, doi:10.1126/science.1222879 (2012)). Reprinted with permission from AAAS.

Specifically, the Wnt N-terminal domain – the thumb – presents a saposin-like fold, with a cluster of α-helices from which a long β-stranded hairpin loop emerges. At the tip of the loop, a highly conserved serine residue is present, post-translationally modified with an O-linked, hence irreversible, acylation, generally in the form of a monounsaturated, 16-carbon long palmitoleic fatty acid (PAM). This lipid is responsible for the majority of the Wnt-Fz contacts at site 1, through van der Waals interactions of the acyl tail with several non-polar residues of Fz8. The identity and/or hydrophobic character of the latter are highly conserved among the 10 vertebrate Frizzled receptors - although with some variations that might account for a certain difference in affinity and thus specificity. These residues form a deep groove, where the palmitoleate sits in an elongated form, shielded from the extracellular environment. A small number of polar contacts are also formed by Fz with protein residues of the Wnt thumb, although their relative contribution to the interaction seems secondary (**Figure 1.2 c**)**.**

The Wnt C-terminal domain, whose tertiary structure resembles the so-called cystine-knot cytokine-like fold, is instead mainly formed by two long antiparallel β-strands pinned together by three disulphide bridges located in the loop region between the strands. As for site 1, even site 2 Wnt-Fz contacts are mainly hydrophobic (though they are based on protein-protein, and not protein-lipid interactions) and the involved residues are generally conserved in the Wnt and Frizzled families (**Figure 1.2 d**).

In line with this conservation pattern, the general architecture of the hWnt3-mFz8 complex is very similar, and nearly identical at the level of the binding interfaces (**Figure 1.2 e**). The main difference between the two structures lies in the oligomerization state observed in the crystal. While the hWnt3-mFz8 complex has been crystallized in a symmetric 2:2 stoichiometry, with one palmitoleic acid in each Fz CRD binding pocket, no such dimer was observed in the XWnt8-mFz8 structure. Intriguingly, structures of other Frizzled CRDs, crystallized in complex with isolated PAMs rather than full-length Wnt, support instead a 1:2 Wnt/Frizzled architecture[30]. This might suggest that different Wnt/Frizzled pairs might have different oligomerization properties, and provide a certain level of specificity in the context of an otherwise broadly degenerate interaction (as evident from the high sequence conservation of the interface residues). A detailed analysis of the finding is however out of the scope of this **Introduction**, but the interested reader can find further information in the excellent discussion of ref. [28].

What is instead central to the goal of the work presented here is that the covalent lipid modification of Wnt is essential for the whole biological activity of Wnt[31]: mutagenesis of the conserved serine residue in mouse Wnt3a yields non-functional and poorly secreted protein[32], while serine-to-alanine mutation of *Drosophila* Wingless disrupts its interaction with the dedicated intracellular Wnt transporter Wntless[33] (see below).

On the other hand, the acylation provides Wnt with a strong hydrophobic character, which might also account for the difficulties in its purification and structural characterization in a recombinant, isolated form.

Most importantly, before the morphogen is able to exert its function at the target cells, its lipid modification mandates the need for a dedicated machinery able to shield the palmitoleic acid and accompany Wnt in all its travels from its synthesis in the endoplasmic reticulum, through the secretory route and in the extracellular space.

The available knowledge about the upstream Wnt signalling - the focus of my thesis - and about the partners of the morphogen in this journey is dissected in the next paragraphs.

The Wnt pathway upstream of receptor engagement

The source of Wnt: the process of Wnt secretion

In addition to the palmitoleoylation, Wnt proteins undergo further extensive post-translational modifications in their secretory route. While, as noted, addition of a lipid moiety is crucial for receptor engagement and Wnt biological activity, glycosylation appears to be mainly relevant to enhance secretion. Moreover, some Wnts have also been reported to present additional modifications (e.g. tyrosine sulfation of Wnt 5a and Wnt11, or incorporation of a glycosylphosphatidylinositol anchor to Wnt1 and Wnt3a), but these seem protein-specific and not a general feature of the family[34].

The Wnt acylation occurs in the endoplasmic reticulum by the action of a transmembrane protein of the MBOAT (membrane-associated O-acyl transferase) family, Porcupine[34]. Namely, Porcupine recognizes a disulphide-

constrained 21-amino acid sequence around the conserved serine at the tip of the Wnt thumb[35], and catalyses the transfer of fatty acid groups from 13 to 18-carbon fatty acyl-Coenzyme A (CoA) substrates to Wnt, with a preferential usage of 16-carbon, cis-Δ9 unsaturated palmitoleoyl-CoA[36], as already mentioned. The preferential incorporation of unsaturated lipids is further supported by the finding that pharmacological inhibition of the ER (Endoplasmic reticulum)-resident integral membrane protein stearoyl-CoA desaturase (SCD), responsible for the desaturation of the saturated precursor palmitoyl-CoA, blocks Wnt lipidation[35]. Since the structure of Porcupine has not yet been solved, however, the molecular reasons for this preference are still elusive, as are the mechanisms regulating the enzyme catalytic activity.

Once acylated, Wnt is then directed through the secretory route (**Figure 1.3**).

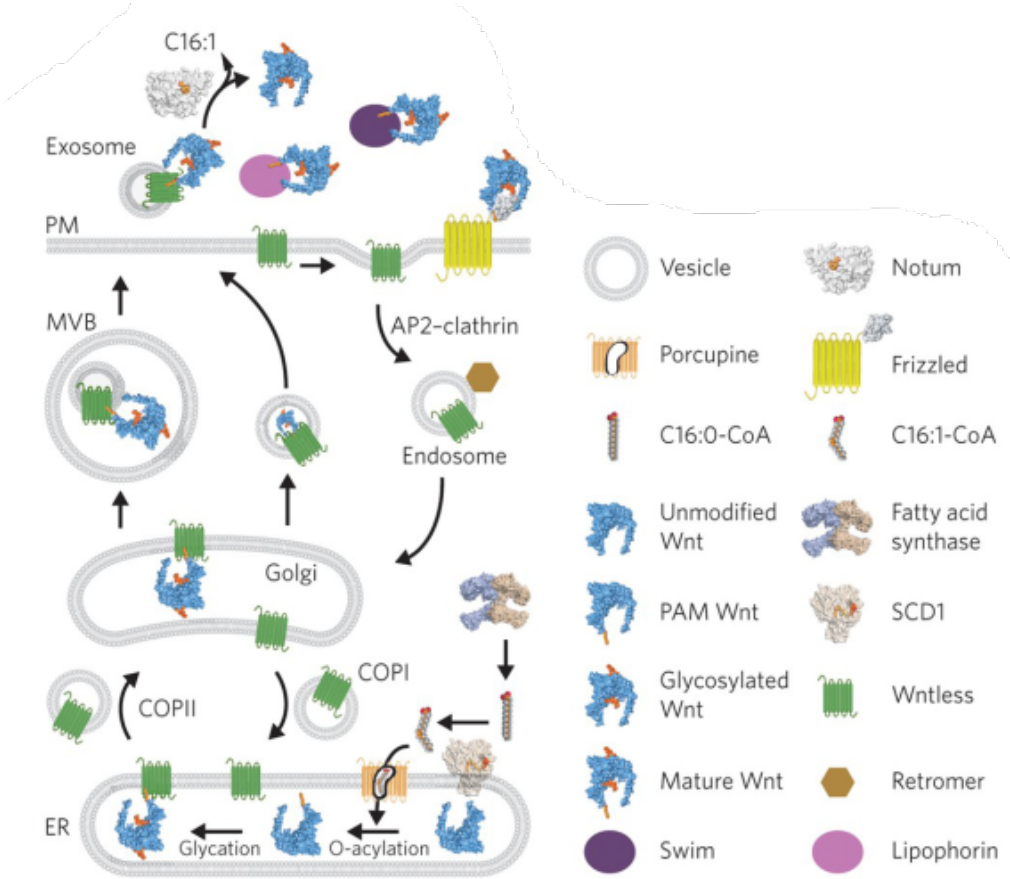


Figure 1.3. Overview of the Wnt secretory route.
The figure is derived from ref. [31].

There, a second multipass transmembrane protein is central to the whole Wnt intracellular trafficking process: Wntless (Wls).

The protein has been reported to bind Wnt in the ER after its lipidation by Porcupine, and to carry it up to the plasma membrane[37,38]. Some reports also suggest that in specific settings (*Drosophila* neuromuscular junction) Wntless can remain complexed with Wnt even in the extracellular space and the complex be internalized by the post-synaptic neurons[39]. Finally, upon trafficking to the plasma membrane and Wnt release, Wls is recycled back to the endoplasmic reticulum in a retromer-dependent manner, mainly driven by motifs present in the third

intracellular loop and in the protein C-terminal tail[40,41] (**Figure 1.4 a**); there, it can in turn bind to newly synthesized Wnt, and cycle it again to the plasma membrane.

Even more importantly, Wntless appears to be a molecule solely dedicated to the Wnt pathway[38]. Relatedly, Wnt directly and tightly regulates the Wntless protein levels, by stabilizing the latter and preventing its ubiquitinoylation, which targets Wntless to degradation in the absence of this stabilizing effect[42].

However, data are lacking with respect to Wntless topology and structure‡. It was initially even unclear how many transmembrane regions Wls presented, and whether it contained a signal sequence. Since N-terminal Flag-tagged Wls can't be detected in non-permeabilized cellular conditions[43], it's likely that it presents an even number, namely 8, transmembrane portions without signal sequence processing (**Figure 1.4 a**), making unlikely the postulated hypothesis of Wls function as an orphan GPCR receptor.

‡ This was the case until very recently. Then, a structure of human Wntless in complex with human Wnt8a has been solved, and this has shed light on several topics discussed in this paragraph, including the number of transmembrane regions and the boundaries and fold of the 1st extracellular loop. Since the lack of structural information about Wntless was one of the bases for part of the research presented in this thesis, the information gathered from the recent Wls-Wnt structure will be presented later, at the end of **Chapter 4,** in order to ensure the logic flow of the text. There, some experimental findings obtained in this thesis and their relevance for future studies will also be retrospectively analysed on the basis of the Wls-Wnt structure.

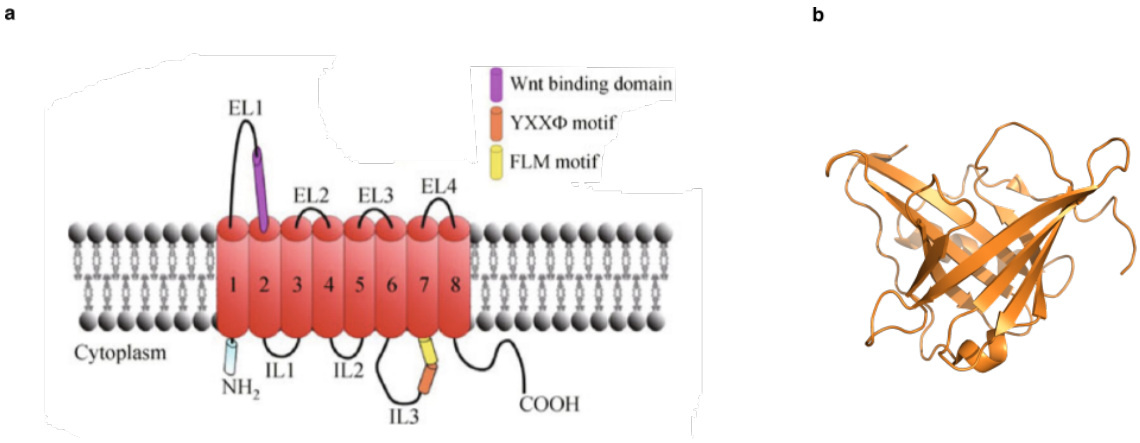


Figure 1.4. The Wnt intracellular carrier Wntless.
a) Wntless predicted topology. The YXXΦ motif (where X can be any residue and Φ is a residue with a bulky hydrophobic side chain) is an endocytic signal for the clathrin adaptor protein 2 (AP2, see also **Figure 1.3**); the FLM sequence seems instead to drive the interaction with the retromer subunit Vsp35 and facilitate retromer transport. **b)** Representative lipocalin fold, from apolipoprotein D (PDB ID: 2hzq). Panel a is derived from ref. [44].

Sequential loop deletions showed that the Wnt binding domain seems to reside in the first extracellular/intraluminal loop[45], and secondary structure predictions suggest that the latter can present a β-barrel-based lipocalin-like fold[46] (**Figure 1.4 b**); however, details about the exact domain boundaries are lacking. The interaction between Wls and Wnt should also strongly depend on the Wnt lipid modification, as mutations in Porcupine or in the Wnt acylated residues cause Wnt accumulation in the ER[33,46].

How Wls separates from Wnt also lacks clarity: a pH-based mechanism has been proposed[46], but the report has not been further confirmed by other studies.

Irrespective of the release mechanism, as Wnt reaches the extracellular space, to move from the producing cells to the receiver cells it has once again to face the issue of poor solubility caused by its lipid modification. In other words, it has

to find ways to overcome what is known in the community as the “Wnt solubility problem”.

Shielding the lipid: the Wnt solubility problem

Wnt trafficking: proposed models

The question about how Wnt is able to travel in an aqueous environment despite its hydrophobicity intersects with a further, debated topic: what is the extent of the Wnt signalling range?

Drosophila Wingless has been classically proposed to act as a long-range morphogen in the formation of the wing imaginal discs[47]; in vertebrates, a long-range Wnt gradient has been, for example, suggested to organize the anterior–posterior axis and neural patterning in *Xenopus*[48].

However, juxtacrine action could suffice in the mouse intestinal stem cell crypts[49].

Even in *Drosophila*, the prototypical model for long-range signalling, membrane tethering of Wg was actually reported not to perturb early fly embryogenesis and to harbour morphologically normal flies, though it still altered post-juvenile organ growth, causing several phenotypes such as delayed development, sterility and reduced food intake[50,51] – supporting the idea that Wnt distant spreading is still essential, at least for some functions or when the body has reached too large dimensions to rely only on juxtacrine signalling.

It is therefore likely that the two mechanisms of Wnt spreading (long-range vs. autocrine/short-range activity) coexist, and, at least in polarized epithelial cells, it

has been speculated[52] that the intracellular Wnt distribution can regulate the signalling range, with apical and basolateral secretion producing short- and long-range gradients, respectively, and the specific glycosylation of the different Wnt family members regulating which secretory route each morphogen takes[53]. In line with this, transcytosis has been shown to molecularly mediate Wnt trafficking from the apical to the basolateral plasma membrane in *Drosophila* wings[54].

In terms of morphogen transport, whereas autocrine/juxtacrine signalling probably obviates the need of Wnt lipid shielding as the fatty acid might remain embedded in the membrane, long-range signalling dictates once again the need of carriers to help Wnt movement in the process.

Several models of Wnt transport have been proposed in the literature, falling into three main classes[34,55]:

1. Wnt can travel on subcellular structures (lipoproteins or exosomes/microvescicles) (**Figure 1.5**).

Lipoproteins are globular vesicles composed of a phospholipid monolayer, surrounding a central core of triacylglycerols and other lipids, and stabilized by proteins called apolipoproteins. In *Drosophila*, Wg has been shown to colocalize with the fly apolipoprotein lipophorin, and RNA interference knockdown of the latter reduced the extent of long-range signalling[56]. Moreover, Wnt3a is associated with lipoproteins in cultured mouse fibroblast media[57]. However, the functional importance of this association has not been evaluated and the broader *in vivo* role of lipoproteins in Wnt transport remains to be fully demonstrated as, for example, lipophorins are expressed in the fly fat body and not in the wing disc,

and a molecular mechanism has therefore to be found to explain how physiologically Wg can be loaded onto lipoprotein particles (LPP) in the latter tissue[34].

Exosomes are another type of membranous particles loaded with cargo proteins and released in the extracellular space after the fusion of intracellular multivesicular bodies (MVB) with the plasma membrane. They have been associated with Wnt transport both *in vitro*[58,59] and *in vivo*, where they have been reported to transport Wg and Wls across the neuromuscular junction of *Drosophila*[39]; furthermore, Wg colocalized with exosomal markers in fly wing discs[58]. On the other hand, knockdown of the key exosome release regulator Rab11 did not affect the extracellular levels of Wg in the latter tissue[59], raising even for exosomes the question of the functional relevance of this form of transport for the *in vivo* morphogen gradient formation.

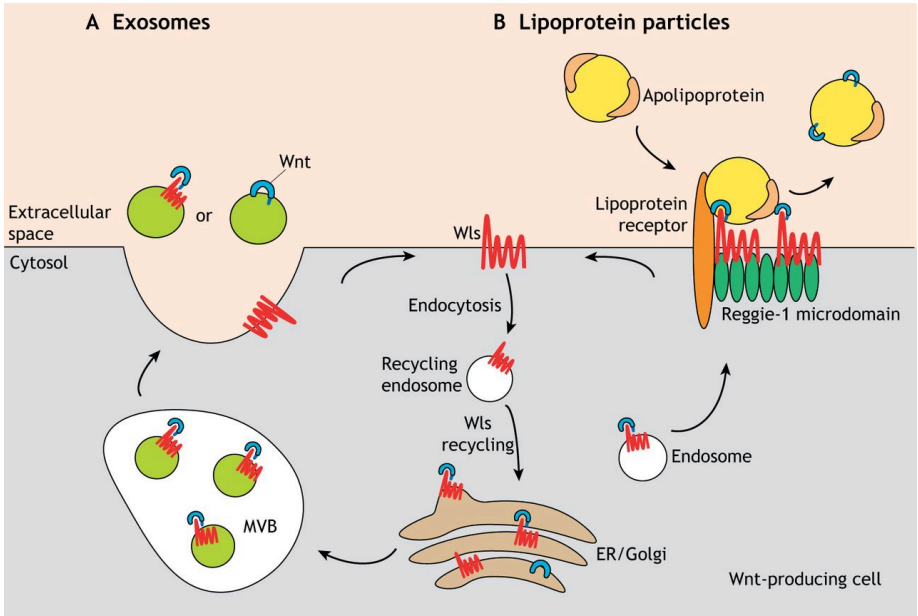


Figure 1.5. Models of Wnt transport on subcellular structures.
The figure is derived from ref. [55].

2. Wnt can form complexes with dedicated protein carriers to shield its lipid from the aqueous environment.

Two main candidates have in particular gathered attention in this sense: Swim (Secreted Wg-interacting Molecule) and afamin.

Swim – which should present a lipocalin-like fold similar to that of Wls, according to structure prediction analysis - has been identified by RNA interference studies in *Drosophila* to affect the long-distance signalling range and the activation of distal Wnt target genes[60].

Afamin is a glycoprotein of the albumin superfamily of serum proteins, which has been recently and unexpectedly found to complex with 12 different Wnt ligands in Wnt-producing HEK293 and Expi293F cells. Additionally, it was shown that the addition of exogenous afamin to serum-free culture media restored Wnt solubility in a dose-dependent manner, and that the morphogen carried by the complex is biologically active when tested in reporter-based luciferase assays[60]. A year after this report, the structure of afamin was solved, and it was shown to present a hydrophobic binding groove potentially able to accept an acyl modification like that of Wnt; indeed, an electron density compatible with a palmitoleic acid (probably derived from the cell culture media) could be visualized in one of the molecule forming the asymmetric unit[61].

However, once again, these models lack a clear confirmation of a wide-ranging role *in vivo*. Vertebrate homologues of Swim, in fact, have yet to be identified, and the evidence supporting its role in *Drosophila* was obtained, as noted, by RNA interference and not genetic knockout methods[34]. In the case of afamin, a

clear functional association between its tissue expression pattern and that of Wnt is still missing, and the protein family is absent in invertebrates, suggesting that this way of extracellular Wnt transport is not an ancient and evolutionarily conserved mechanism[55], and questioning therefore the extent of its functional relevance.

3. Wnt may actually never leave the cell.

The already cited study suggesting the dispensability of long-range signalling in mouse intestinal crypts[49] also hinted that the movement of the morphogen away from the source happens in a cell-dependent manner through subsequent cell divisions.

In a long-range signalling context, Wnt might also be presented to target receptors by long, actin-based cellular protrusions up to 700 um-long called cytonemes, a subset of specialized filopodia (**Figure 1.6**). Wnt movement from the producing cells associated with these extensions has been for example reported in zebrafish[62] and in human cells[63], and, in zebrafish in particular, Wnt8 can reach by this route distant target sites during neural plate formation. Formation of these structures has also been shown to be directly controlled by Wnt, through Ror2-mediated activation of the non-canonical, β-catenin independent PCP pathway. This leads to actin polymerization and *de novo* biogenesis of cytonemes delivering Wnt to the receiving cells, where the canonical signalling is activated[64] – also suggesting an interesting functional relationship between canonical and non-canonical pathways.

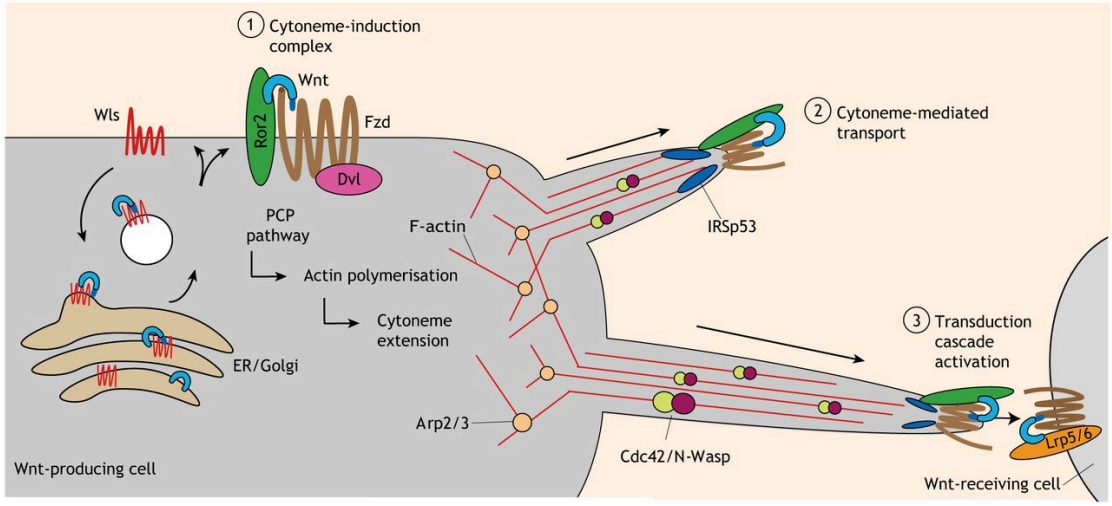


Figure 1.6. Model of Wnt transport on cytonemes.
The figure is derived from ref. [55].

In addition to the main models presented above, and intersecting with some of them, a further class of proteins has been proposed to be involved in the regulation of Wnt movement in the extracellular space, although subsequently slightly neglected in comparison to the other models: the membrane-associated heparan sulphate proteoglycans glypicans. As the data presented in this thesis point instead towards a crucial role of the latter in determining the extent of Wnt distribution, a detailed analysis will be dedicated to this group of molecules in the next paragraph.

Glypicans and their involvement in Wnt transport

Glypicans are glycoproteins composed of a central core protein which connects to the membrane through a stalk region terminating in a GPI anchor. From the stalk segment, containing Ser-Gly glycosaminoglycan (GAG) attachment sites,

long heparan sulphate chains (HS), composed of 50-250 repeated disaccharide units of glucosamine and uronic acid, extend (**Figure 1.7 a**).

Glypicans have been indicated to interact with a plethora of membrane-bound and extracellular proteins, including the morphogens Hedgehog and Wnt, and many of these interactions involve the long, negatively charged, sugar protrusions[65].

Indeed, the first associations between heparan sulphate proteoglycans and impaired Wnt signalling were identified in *Drosophila* mutants for members the evolutionary conserved enzymatic machinery responsible for attachment and modification of GAG chains to the core protein: namely, *Sugarless*, homologue of the vertebrate UDP-glucose dehydrogenase[66-68], and *Sulfatless*, homologue of the vertebrate heparan sulphate N-deacetylase/N-sulphotransferase[69]. In line with these initial results, accumulating evidence then further linked glypicans to the regulation of Wnt distribution. Two glypicans, Dlp (or Dally-like protein) and Dally, are encoded by *Drosophila*, while humans present six glypicans (**Figure 1.7 b**). Genetic null flies for *Dlp* and for *Dally* showed their essential role in the formation of extracellular Wg gradient[70]; the two proteins also both were shown to bind Wnt in cell culture[71]. Furthermore, Dally was shown to cooperate with *Drosophila* Frizzled 2 to transduce Wnt signalling, possibly acting as a co-receptor[69]. In vertebrates, the zebrafish glypican knypek potentiates Wnt11 signalling[72], and glypican 4 interacts with Wnt11 in axis formation in *Xenopus* embryos[73].

Despite their undisputed importance in Wnt modulation, two main considerations pointed against a primary role of glypicans as key molecules able to tackle the

issue of Wnt hydrophobicity and solubility. Since the presented studies indicated an interaction driven by the sugar chains, hence electrostatic-based, the question is how much the binding is specific, and selective for some glypicans/Wnt pairs, a general caveat to take into account for heparan-sulphate binding proteins[74]. Second, these interactions are charge-based and polar, therefore not suggesting any possible link with an ability to shield the Wnt lipid.

Interestingly, actually, some levels of differential activity were reported for Dlp vs. Dally. Dlp was in fact shown to have a biphasic activity in Wnt signalling, acting as a negative regulator of Wnt signalling close to the source, inhibiting the activation of the high-threshold target gene *senseless*, and a positive regulator at a distance, enhancing the expression of the low-threshold target gene *distal-less*[71]. This effect was also suggested to depend on the relative expression at the cell surface of Wg, Dlp, and Frizzled receptors, explaining the biphasic activity by a competition between Dlp and Fz for binding. In the same study, Wg was also shown to directly bind to the Dlp protein core and not only to the GAG chains, hinting to a more specific interaction[75].

However, the later solved structure of Dlp[76] (and the very similar human glypican 1 structure[77,78]) didn't suggest - as can be observed in **Figure 1.7 c** - any obvious binding interface, even less for a lipid moiety, making further structural studies necessary to elucidate the issue.

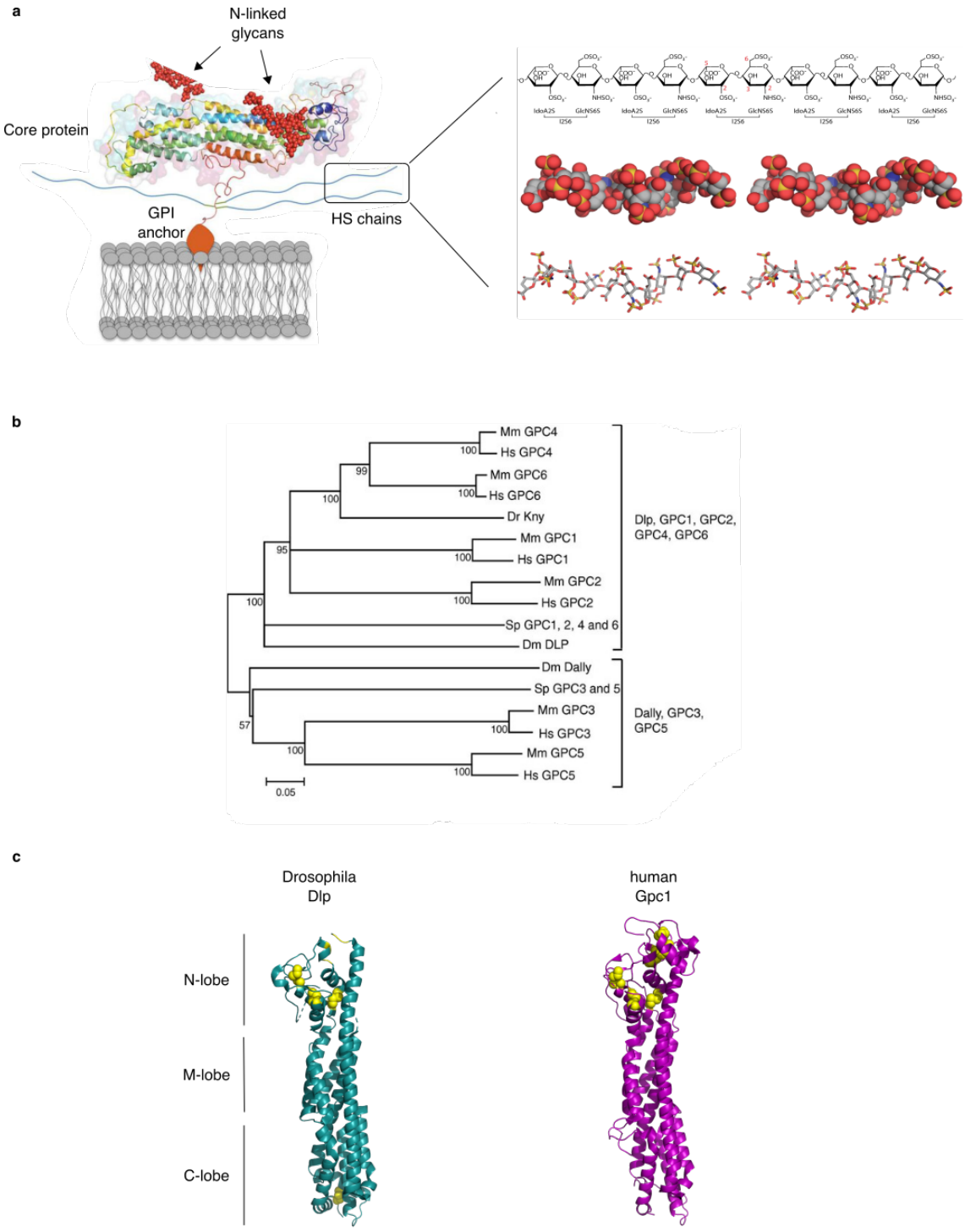


Figure 1.7. The glypican family of heparan sulphate proteoglycans.
a) Structural features of glypicans and predicted membrane topology. The magnified box shows a representative heparan-sulphate decasaccharide repeat in chemical, space-filling and stick representation, based on a heparin structure (PDB ID: 1E0O). **b)** Phylogenetic tree of the evolutionarily relations and subfamily classification of glypicans in vertebrates and invertebrates. The abbreviations refer to human (Hs), mouse (Mm), zebrafish (Dr), purple sea urchin (Sp), and *Drosophila melanogaster* (Dm), respectively. **c)** Crystal structures of

Dlp (PDB ID: 3ODN) and human Gpc1 (PDB ID: 4YWT), the two glypicans whose structure have been solved. Disulphides are shown as yellow spheres. Glypicans present a cylindrical all-helical fold that can be divided into three parts: the N-terminal N-lobe, or Cys-rich lobe, which contains most of the evolutionary conserved cysteines of the glypicans; the central M-lobe; and the C-lobe, also called protease-site lobe as the region undergoes furin protease processing in many glypicans[77]. Further information about the structural features of the glypicans will be discussed in **Chapter 3**. Panel a is adapted from refs. [74,78], whereas panel b is derived from ref. [79].

Wnt extracellular modulators

Several extracellular activators and inhibitors of the Wnt pathway exist[80,81]: they can either act by modulating the relative abundance of Wnt receptors and co-receptors, or directly interact with Wnts to modulate their extracellular spread and activity.

In the group of proteins that directly act on Wnt, both transmembrane and secreted regulators can be retrieved.

Among the transmembrane proteins, the Tiki1 and Tiki2 metalloproteases are of particular importance. Both have been functionally demonstrated to antagonize Wnt signalling, and Tiki2 in particular has been identified by mass spectrometry to cleave Wnt at an 8 amino acid-long sequence at the N-terminus: removal of the latter causes oligomerization and subsequent inactivation of the morphogen[82].

In the secreted class, secreted Frizzled-related proteins (sFRPs) and Wif-1 in particular have been suggested to directly interact with Wnt to reduce signalling, preventing its binding to receptors: in the first case, this can be mediated by the CRD-like domain present in sFRPs[83], while Wif-1 has been crystallized in

complex with endogenous phospholipids and shown to present a lipid-binding pocket[84], raising the possibility that it might also accommodate the Wnt lipid.

However, one candidate in particular has been unambiguously associated and structurally confirmed to be a dedicated Wnt effector acting directly on Wnt, and in particular on the Wnt lipid: the extracellular hydrolase Notum.

Its structure was solved by our group in 2015[85], confirming the identity of the protein as a Wnt-dedicated feedback inhibitor, and uncovering the molecular basis for this specificity. For its crucial role in the pathway, Notum also represents an attractive therapeutic target, and this thesis presents important results obtained in the effort of developing Notum inhibitors to boost Wnt signalling in the context of Alzheimer's disease.

For its direct action on the Wnt lipid, and the centrality to the work presented here, a specific section will therefore now be dedicated to a detailed description of the Notum structure and mechanism of action. The rationale for the translational application of the drug development campaign to the neurodegenerative area and specifically to Alzheimer's will be instead discussed in the related results chapter (**Chapter 5**).

Notum

Initially, Notum was thought to be a phospholipase, exerting its effect on the Wnt pathway by enhancing glypican shedding in the extracellular space, cleaving their GPI anchor[86]. However, due to the already mentioned pleiotropic involvement of glypicans in several morphogens' signalling[65], such an activity would have been

expected to have some influence on additional pathways rather than Wnt only. Instead, the enzyme appeared to be functionally dedicated to Wnt, expressed upon activation of Wnt signalling, and specifically acting on the Wnt cascade[87-89]. The apparent contradiction was clarified by the solution of the Notum crystal structure[85], which demonstrated that the protein was not a phospholipase. In fact, the protein presented the canonical α/β hydrolase superfamily fold - with the classical central β-sheet core composed of 8, mostly parallel, β-strands, surrounded by 6 α-helices, extended on both sides by additional strands and helices (**Figure 1.8 a** and **b**). Bioinformatic structure similarity analysis suggested in particular that the enzyme could be provided with carboxylesterase activity: in other words, with the ability to cleave carboxy-oxoester bonds, like the one that connects the palmitoleic acid with the conserved serine at the tip of the Wnt thumb. Indeed, enzymatic assays using chromogenic carboxylesterase substrates experimentally validated the prediction, and cellular luciferase assays and mass spectrometry analysis confirmed that Notum directly reduced Wnt signalling by deacylating the morphogen.

The structure of apo wild-type Notum, and of Notum^{S232A} (mutant for one of the amino acids of the Ser232-Asp340-His389 catalytic triad and therefore enzymatically dead) in complex with a hWnt7a-mimicking peptide showed that the catalytic triad is adjacent to a wide (~380 Å^3), but shallow hydrophobic pocket, where electron density consistent with the palmitoleate of the Wnt peptide could be observed in the complexed structure (**Figure 1.8 c**). The carboxylic acid carbon is positioned at the ideal 3.3 Å distance from the C_β of the catalytic serine, supporting also from a structural point of view the identification of Notum as a

carboxylesterase; moreover, the optimal reciprocal positioning of the G127-W128, S232-A233 and G126-G127 backbone amide bonds indicates that the classical oxyanion hole transition state should be particularly stable.

As noticed, a key feature of the Notum binding pocket is its shallowness. This, in turn, mandates that the cavity cannot accommodate extended fatty acids longer than 8-10 carbons. The latter can therefore only fit if they are *cis*-unsaturated at C9-C10, like the Wnt palmitoleate, but not, for example, the saturated C16 palmitic acid of Hedgehog. This feature provides the basis for the specificity of Notum for Wnt as observed in functional studies: the cis-unsaturation determines in fact a kink in the lipid chain geometry (**Figure 1.8 d**) that allows the long acyl chain to snuggle in the cavity, and provide the basis for the specificity of Notum for Wnt as observed in functional studies.

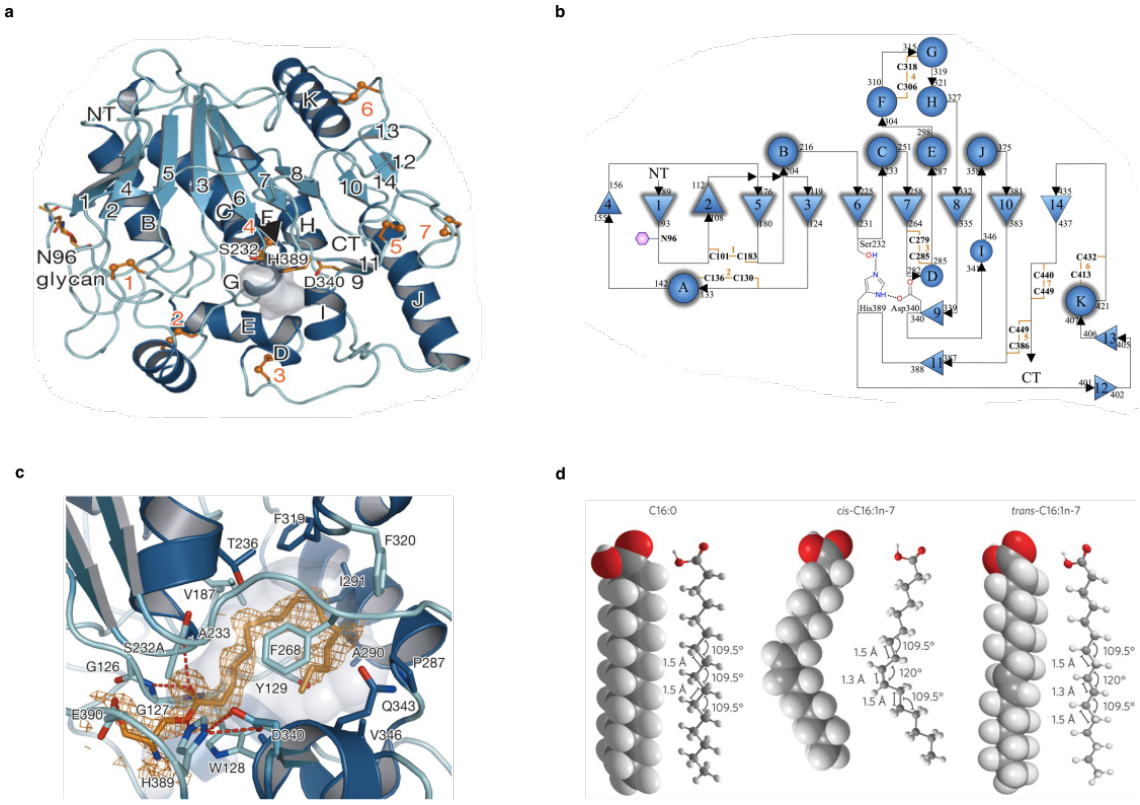


Figure 1.8. Structure of the Wnt extracellular deacylase human Notum and modes of Wnt binding.
a) Cartoon representation of the structure of apo-Notum (PDB ID: 4uz1). β-strands are indicated by numbers, α-helices by letters. Disulphides are shown as sticks in orange. The binding pocket is presented in surface visualization in grey. NT and CT correspond to N-terminus and C-terminus, respectively**. b)** Topology diagram of Notum. Circles indicate helices, triangles indicate strands. Glycosylation is shown as a pink hexagon. Shading indicates the secondary structure elements belonging to the canonical α/β hydrolase fold. **c)** Close-up view of the Notum binding pocket from the structure in complex with a hWnt7a-based palmitoleoylated peptide (PDB ID: 4uzq). Red dashes indicate hydrogen bonding. The omit electron density map is contoured at ±2σ. **d)** Ball and stick representation of saturated 16-carbon palmitic acid (C16:0), left, unsaturated *cis*-Δ9-palmitoleic acid (C16:1), middle, and *trans*-Δ9-palmitoleic acid (C16:1), right, showing the geometry difference and the kink introduced by the *cis*-Δ9 unsaturation. Panel a-c are derived from ref. [85], panel d from ref. [31].

Aims and research questions

As evident from what is discussed in the previous sections, before being able to reach and engage its target membrane receptors - either the canonical Frizzleds or the plethora of non-canonical receptors - and hence exert the biological activity central to so many conditions during development, health and disease, the Wnt morphogen needs to overcome several hurdles. The most prominent is represented by its reduced solubility in the aqueous environment, caused by the post-translational modification with the evolutionary conserved, and functionally crucial, palmitoleic acid lipid moiety.

In comparison with the better characterised downstream cascade, the events and partners helping Wnt in the task still awaits a deeper understanding. In the case of the extracellular transport of Wnt, in particular, several, and sometimes contrasting, candidates have been suggested to be able to shield the lipid, but, as noted, either they lack an undisputed proof for their *in vivo* involvement, or, in case of evidence for a wide-ranging functional role, a clear molecular explanation for their lipid-sequestering ability is still missing.

The work presented in this thesis aimed therefore at tackling the issue by focusing from a structural perspective on the modes of Wnt lipid engagement, with two main aims: on the one hand, to understand what are the molecular mechanisms helping Wnt to overcome the solubility problem; on the other hand, from a translational point of view, to investigate if we can act on the Wnt lipidation to open further therapeutic avenues to target the pathway in disease states.

With respect to the first, more biologically-oriented aim, we devoted our attention in particular to Wnt extracellular transport, the area where less clarity exists. To this purpose, we partnered with the Vincent lab at the Francis Crick Institute in London to first functionally re-screen, in the powerful experimental setting of the *Drosophila* wing, the main models proposed in the literature and identify, among them, the candidate with the most striking evidence for a role in Wnt lipid shielding. This represented the preliminary basis to then validate, explain and provide a molecular support to the observation by biophysical, biochemical and structural methods.

In addition to this main part of the work, still in the "basic research" area of the project, we also tried to understand how the Wnt is able in the first place to get to the extracellular space, since its acylation is equally likely to pose a problem, in terms of solubility, in the intracellular milieu and in the secretory route. In other words, we wanted to define the structural and molecular basis of Wnt intracellular transport as well. While in this sense the identity of the carrier is clearer, and is represented by the transmembrane protein Wntless, as discussed above the biochemical and structural information about the protein is however extremely limited: we therefore set to clarify the matter both by focusing on the full-length protein and on the minimal unit reported to be the domain sufficient to promote Wnt binding, i.e. the 1st extracellular loop.

The results related to these two sections are reported in **Chapter 3** and **Chapter 4**, respectively.

From the translational perspective, instead, the overarching question was if we could focus on the Wnt lipid - due to its central role in determining signalling

activation - also to devise new and more effective strategies to modulate the pathway, targeting extracellular modulators of the Wnt lipidation.

For the availability of high-resolution information to inform a structure-guided drug design campaign and for its function specific and dedicated to the Wnt pathway, the best candidate in this sense was represented by Notum. This part of the project aimed at developing inhibitors to prevent its enzymatic action, and therefore boost Wnt pathway activation, in situations where a downregulated signalling has been causally associated to degenerative processes, like in the case of Alzheimer's disease. **Chapter 5** therefore summarizes the results obtained in this area, while the advantages of targeting Notum with respect to other potential candidates noted in this **Introduction** chapter (e.g. Porcupine, downstream cascade) are evaluated in the conclusive general **Discussion** chapter.

The final **Discussion** chapter also presents a summary of the most important future directions opened by the results presented in **Chapter 3** and **Chapter 4** on the extracellular and intracellular Wnt transport.

Chapter 2 - Materials and Methods

General methods common to all chapters

Polymerase chain reaction and site-directed mutagenesis

The target genes of interest (see chapter-specific paragraphs of Materials and Methods for details of construct boundaries) were amplified by polymerase chain reaction (PCR)[90].

In this thesis, Pyrobest™ DNA polymerase (Takara) was used, a thermostable DNA polymerase from *Pyrococcus* species, exploited for its high fidelity in DNA replication, thanks to its 3'-5' exonuclease proofreading activity.

A typical reaction was performed as detailed in **Table 2.1** and **Table 2.2**, adjusted when needed on the predicted annealing temperature for each primer pairs (related to primer dimension and sequence composition, in turn determining their estimated melting temperature):

Table 2.1. Typical PCR reaction volumes and concentrations.

Reagent	Volume (concentration)
DNA	100 ng
dNTP mixture (2.5 mM each)	4 µl (0.2 mM each)
Pyrobest™ Buffer II (10x)	5 µl (1x)

Reagent	Volume (concentration)
Sense primer 5 µM	10 µl (0.5 µM)
Antisense primer 5 µM	10 µl (0.5 µM)
Pyrobest™ DNA polymerase (5 units/µl)	0.5 µl (2.5 units)
Water	up to 50 µl
Final volume	50 µl

Table 2.2. Typical PCR reaction conditions.

Step	Temperature and duration
1. Initial denaturation	98 °C for 2 min
2. 30 cycles of:	
Denaturation	98 °C for 20 sec
Annealing	55 °C for 30 sec
Elongation	72 °C for 1 min§
3. Final elongation	72 °C for 5 min

For the generation of mutants, site-directed mutagenesis was performed applying a two-step overlap extension PCR protocol[91], first generating two DNA fragments

§ For elongation of a 1 kb DNA fragment.

with overlapping ends at the site of the mutation(s), which were then “fused” with a second PCR reaction.

Agarose gel electrophoresis and gel extraction

To visualize the success of PCR reactions, as well of enzymatic DNA digestions during cloning (see below), the DNA fragments were generally run on 1% agarose gels. In typical conditions, 0.5 g of agarose powder was dissolved in 50 ml of TAE (Tris-Acetate-EDTA) buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) by microwave heating. Before cooling, 5 µl of 10000x SYBR™ Safe DNA Gel Stain (Thermo Fisher Scientific) were added. Upon gel solidification and addition of 5x loading dye containing bromophenol blue, xylene cyanol and orange G (Qiagen) to the samples, the DNA mixtures were run at 100 V for 30 min, and the results analysed by UV transillumination at a wavelength of 302 nm.
Agarose gel bands corresponding to the fragments of interest were excised from the gel, weighted and purified using the QIAquick Gel Extraction Kit (Qiagen), following the manufacturer’s instructions.

DNA enzymatic digestion, ligation and transformation

Enzymatic restriction of PCR-amplified fragments and of suitable vectors (see chapter-specific methods) was performed to generate hanging DNA strands with complementary sequence to allow ligation of the PCR product into the vector of choice. The choice of the enzymes used in each cloning was based on the

features of the vector multiple cloning site (MCS) and on the goals of construct design. Namely, the two vector suites used for cloning purposed in this thesis, pHL-SEC[92] and pHR-CMV-TetO_2[93], present the same MCS, and the enzyme pair AgeI/KpnI was employed when a simple change of insert without modifying the vector features (e.g. secretion signal sequence, purification tag) was needed; combinations of the latter enzymes with EcoRI and XhoI were instead employed when more extensive modifications were required. The enzymes were purchased from New England Biolabs in a high-fidelity format where available (to reduce non-specific sequence cleavage) and a typical reaction was set in a final volume of 50 µl, using 1 µl of each enzyme and performed overnight at a temperature dependent on the specifics of each enzyme as per the manufacturer's instruction. The cleaved DNA products were run and purified from agarose gels as described above.

To ligate the cleaved cohesive vector and insert DNA ends, the LigaFast™ Rapid DNA Ligation System from Promega was used. Typically, a reaction was performed using 100 ng of vector and a 3x molar excess of insert, in a final volume of 10 µl including 5 µl of 2X Rapid Ligation Buffer (60 mM Tris-HCl (pH 7.8), 20 mM MgCl2, 20 mM DTT, 2 mM ATP, 10% (w/v) PEG) and 0.5 µl of T4 DNA ligase (1.5 units), incubated for 5 minutes at 25 °C.

The whole ligation mix was then used to transform 50 µl of chemically-competent *Escherichia coli* DH5α cells (Thermo Fisher Scientific). The mixture of cells and ligation mix was incubated on ice for 20 minutes, heat-shocked at 42 °C for 45 seconds to facilitate incorporation of the target DNA, and incubated again on ice for 5 minutes. The cells were then recovered adding 150 µl of SOC medium

(Thermo Fisher Scientific, 2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4, 20 mM glucose) and incubated for 1 hour at 37 °C, shaking at 200 rpm. Finally, the mixture was plated onto antibiotic-selective Lysogeny broth (LB)-agar plates (using ampicillin at a concentration of 100 µg/ml) and incubated overnight at 37 °C.

Purification of plasmid DNA

Transformed colonies were inoculated into 10 ml of LB broth (10 g/L peptone, 5 g/L yeast extract, 10 g/L NaCl), with the appropriate antibiotic(s) and incubated overnight at 37 °C, shaking at 200 rpm. The following day, the cultures were centrifuged for 10 min at 8000xg, the supernatant discarded and the pellet resuspended in 250 µl of Buffer P1 (50 mM Tris·Cl (pH 8.0), 10 mM EDTA, 100 µg/ml RNase A) from the QIAprep Spin Miniprep Kit (Qiagen). The samples were then further processed following the manufacturer's instructions, and the purified plasmid DNA generally eluted in 80 µl of EB buffer (10 mM Tris·Cl, pH 8.5). The DNA was sequenced by Sanger sequencing[94] using the commercial service offered by Eurofins Genomics.

For the purification of plasmid DNA to use in large-scale transient transfection of mammalian cells, one colony from a transformation with the construct of interest was inoculated into 10 ml of LB and incubated for 8 hours at 37 °C, shaking at 200 rpm. The culture was then diluted into 3 litres of LB + antibiotic(s) and grown overnight at 37 °C, shaking at 200 rpm. The day after, the cells were centrifuged for 25 min at 8000xg, resuspended in 125 ml of buffer P1 from the HiSpeed

Plasmid Giga EF Kit (Qiagen), and further processed according to the manufacturer's instructions. The DNA was finally eluted generally in 3 ml of EB buffer.

SDS-PAGE

Upon expression and purification of the different constructs produced in the present thesis, protein purity was evaluated by Sodium Dodecyl Sulphate - Polyacrylamide Gel Electrophoresis (SDS-PAGE) in reducing conditions, and non-reducing conditions where appropriate. In reducing conditions, a suitable volume of 4x NuPAGE LDS Sample Buffer (Thermo Fisher Scientific) was added to the samples, and the latter were also supplemented with β-mercaptoethanol at 2.5% (v/v) final concentration, and boiled at 99 ºC for 5 min. In non-reducing conditions, the loading buffer was added and the samples immediately run on the gel.

NuPAGE Bis-Tris Gels (Thermo Fisher Scientific) at a 12% (w/v) acrylamide concentration were used, run for 35 min at 165 V using NuPAGE MES SDS Running Buffer (Thermo Fisher Scientific, 50 mM MES, 50 mM Tris Base, 0.1% (w/v) SDS, 1 mM EDTA, pH 7.3) or NuPAGE MOPS SDS Running Buffer (Thermo Fisher Scientific, 50 mM MOPS, 50 mM Tris Base, 0.1% (w/v) SDS, 1 mM EDTA, pH 7.7), according to the specific molecular weight to resolve. The different ionic composition and pKa of each buffer, in fact, influence the speed of migration and make the MES-based buffer more suitable to separate small-to-medium sized proteins, the MOPS one more useful for medium-to-large proteins.

The BenchMark™ Unstained Protein Ladder (Thermo Fisher Scientific) was used as molecular weight ladder. The gels were stained using the Expedeon Instant Blue™ Coomassie™-based protein stain.

Western Blot

For Western blotting, first SDS-PAGE was performed as detailed above. After SDS-PAGE, proteins from the gel were transferred onto a nitrocellulose membrane (0.45 µm pore size, Amersham Hybond™ ECL™ membrane from GE Healthcare) in a wet transfer electrophoretic apparatus using the XCell SureLock™ Mini-Cell system (Thermo Fisher Scientific). The transfer was performed in 25 mM Bicine, 25 mM Bis-Tris, 1 mM EDTA, pH 7.2.

Upon transfer, for TwinStrep-tagged proteins, the membrane was first blocked with 3% (w/v) BSA (Bovine Serum Albumin, Sigma-Aldrich) at room temperature for 1 hour in PBS, 0.1% (v/v) Tween-20, and subsequently incubated overnight at 4 °C with a StrepTactin-Horseradish Peroxidase (HRP) Conjugate probe (Biorad), used at a 1:25000 dilution in 25 ml of PBS, 0.1% Tween-20, 3% BSA. The day after, the membrane was washed 4 times with PBS, 0.1% Tween-20 for 5 minutes; after that, the chemiluminescent signal was detected using the Clarity™ Western ECL Substrate kit (Biorad) and the ChemiDc™ MP Imaging System (Biorad). The molecular weight ladder used in this case was the Precision Plus Protein™ WesternC™ (Biorad), containing strep-tagged standards.

For His-tagged proteins, the blocking step was performed with 5% (w/v) fat-free milk powder in PBS, 0.1% Tween-20, for 1 hour at room temperature.

Subsequently, the membrane was incubated overnight at 4 °C with anti-PentaHistidine primary antibody (Qiagen, 1:2500 dilution, in 10 ml of PBS, 0.1% Tween-20, 5% milk). The following day, the membrane was first washed 4x5 min with PBS, 0.1% Tween-20, and then incubated for 1 hour at room temperature with a goat anti-mouse IgG HRP-conjugate antibody (GE Healthcare) used at a 1:2500 dilution, in 10 ml of PBS, 0.1% Tween-20, 5% milk. Upon 4 more washes with PBS, 0.1% Tween-20, the blot was developed using the same apparatus described above. The molecular weight ladders used here were both the BenchMark™ Prestained Protein Ladder (Thermo Fisher Scientific) - to assess the success of the transfer procedure - and the BenchMark™ Unstained Protein Ladder (Thermo Fisher Scientific) - containing His-tagged standards to visualize the ladder in Western Blot. The two ladders were pre-mixed in the same SDS-PAGE well (10 ul of Prestained + 1 ul of Unstained).

Methods specific to Chapter 3**

Construct design

Drosophila melanogaster Dally-like protein (Dlp, UNIPROT database reference Q9VUG1) constructs encompassed aa. D74-Q617 of the protein sequence, thus including the protein core and stopping just before the heparan sulphate

** Parts of the text in this section are quoted verbatim from the Methods section which I wrote for the following article, of which I am co-first author: ref. 95 McGough, I. J. *et al.* Glypicans shield the Wnt lipid moiety to enable signalling at a distance. *Nature*, doi:10.1038/s41586-020-2498-z (2020).

attachment sites. Namely, as also detailed in the related chapter, two wild-type constructs with these N-terminal and C-terminal boundaries were designed: one, $Dlp^{core\ furinmut}$, mutated for the furin recognition sites located at aa. 400 and 437 (by mutations K398Q, K399A, R402Q, R438Q, R441A); and one where the amino acid stretch between the two sites was removed, to allow a more homogeneous furin processing, $Dlp^{core\ \Delta 400\text{-}437}$. The two constructs were cloned into pHR-CMV-TetO2_3C-Twin-Strep_IRES-EmGFP (Addgene plasmid # 113884), allowing lentiviral-based generation of stable mammalian cell lines expressing the protein with a Human Rhinovirus (HRV) 3C Protease-cleavable Twin-Strep tag at the C-terminus, and an internal IRES offering both a direct readout of the success of the infection procedure, and the possibility of fluorescent cell sorting-based enrichment of the infected cells. The same plasmid was also used for transient transfection of mammalian cells in further batches of protein production.

The Dlp mutants tested in FSEC were based on the $Dlp^{core\ furinmut}$ scaffold and additionally contained mutations S168M and M175C/N367C respectively. They were cloned in the pHL-SEC-AVI-His, encoding for C-terminal Avitag (substrate sequence for biotinylation) and 6xHis-tag, and used to transiently transfect mammalian cells.

Drosophila melanogaster Dally core protein used in biophysical studies (UNIPROT database reference Q24114) spanned residues S65–I541 of the protein sequence, thus excluding the heparan sulphate attachment sites as in the case of Dlp, and was cloned in pHL-SEC-His, carrying a C-terminal 6xHis-tag. The plasmid was then employed for transient transfection of mammalian cells.

Generation of lentiviral-transduced HEK293S GnTI- TetR cell lines expressing Dlp core

The procedure was carried out mainly following the protocol outlined in ref. [93]. Briefly, 4 ug of DNA mixture containing a 1:1:1 ratio of pHR-CMV-TetO2_3C-Twin-Strep_IRES-EmGFP (Addgene plasmid #113884, containing $Dlp^{core\ furinmut}$ and $Dlp^{core\ \Delta 400\text{-}437}$) and of lentiviral packaging and envelop plasmids (psPAX2, Addgene plasmid #12260, and pMD2.G, Addgene plasmid #12259, respectively) were used to transfect virus-producing HEK293T Lenti-X cell (Takara/Clontech), with PEI (polyethylenimine) used as a transfection agent (at a 1:2.5 DNA:PEI ratio), in a 6-well plate with 2 ml of DMEM/F12 containing 2% fetal bovine serum (FBS) (Thermo Fisher Scientific). After 2 days of incubation at 37 °C (5% CO_2), the virus-containing supernatant was harvested and used to infect target HEK293S GnTI- TetR cells[96] (allowing inducible protein expression with doxycycline) for 3 days before changing media to DMEM/F12, 10% FBS. The cells were then progressively expanded to a 75 cm^2 and 175 cm^2 flask-format, before being harvested, washed twice with PBS, and seeded at a density of 0.5 × 10^6 cells/mL in shaking flasks in FreeStyle 293/1% FBS (Thermo Fisher Scientific), to adapt them to suspension culture.

At the same time of the expansion, a further 6-well containing the transduced cells in DMEM/F12, 10% FBS was prepared, to check in small-scale by Western Blot the expression of the target proteins. When the cells reached 80% confluency, the media was changed to 2% FBS and the expression induced with 1 µg/ml of doxycycline for 96 h at 37 °C.

Small-scale expression of Dlp mutants

In the case of Dlp mutants, to check for their successful expression, the small-scale tests in 6-well format were performed by transient transfection. 4 µg of construct DNA, combined with PEI at a 1:2 ratio, were used to transfect ~80% confluent HEK293S GnTI- (ATCC, cat. no. CRL-3022) in 2 ml of DMEM/2% FBS; the unconcentrated supernatant was collected and tested by Western Blot after cell growth at 37 °C for 96 hours in a humidified incubator with 5% CO_2.

Large scale expression and purification of Dlp core wild-type and Dally core

As already mentioned, both Dlp core proteins constructs ($Dlp^{core\ furinmut}$ and $Dlp^{core\ \Delta 400\text{-}437}$) were expressed, secreted in the culture medium, either by transient transfection of HEK293S GnTI- cells or by induction of expression of lentiviral-transduced HEK293S GnTI- TetR cells, in which case 1 µg/ml of doxycycline for 96 h at 37 °C was used for induction when the cells reached a cell density of $3x10^6$ cells/ml in 3 liters/batch of suspension culture. For transient transfection, 12 roller bottles of adherent HEK293S GnTI- cells/batch were transfected with 6 mg of DNA premixed with 12 mg PEI according to published protocols[92], and the conditioned medium was collected seven days post-transfection.

Irrespective of the expression procedure, the collected conditioned medium was dialysed against 50 mM Tris-HCl, pH 8, 200 mM NaCl using a QuixStand benchtop system (GE Healthcare) connected to a 60 cm Xampler Cartridge (GE Healthcare) with a 10-kDa nominal MWCO. Initial purification was performed by

Strep-tag affinity chromatography (StreptactinXT High Capacity column, IBA Lifesciences). In brief, the protein was loaded overnight with continuous circulation on the column, before being eluted in 100 mM Tris-HCl, pH 8, 150 mM NaCl, 1 mM EDTA, 50 mM biotin. The protein was further purified by size-exclusion chromatography (SEC) (Superdex 200 Increase 10_300 GL column, GE Healthcare) in a final buffer containing 10 mM Tris-HCl, pH 8, 35 mM NaCl. Between the affinity chromatography and SEC steps, the Twin-Strep tag was removed by incubation for 48 h at 4 °C with HRV 3C Protease, added at a 1:50 (w/w) protease:protein ratio.

Dally^{core} was expressed by transient transfection of HEK293T cells (ATCC, cat. no. CRL-3216) in the presence of the class I α-mannosidase inhibitor, kifunensine[92,97]. Conditioned medium was collected five days after transfection. Medium was concentrated and dialysed, using the same apparatus described above, into 25 mM phosphate pH 8, 250 mM NaCl. The sample was incubated with TALON beads for 1 h at 16 °C. Beads were washed with 25 mM phosphate pH 8, 250 mM NaCl, followed by washes in wash buffer containing 5 mM imidazole. The protein was eluted in wash buffer containing 250 mM imidazole. Samples were further purified by SEC using a HiLoad 16_600 Superdex 200 pg column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl.

Man(5)GlcNAc(2) and Man(5–9)GlcNAc(2) residual glycosylation produced by expression in GnTI- HEK293S or by HEK293T cells with kifunensine, respectively, was preserved for biophysical and FSEC analysis of protein-lipid interaction, while the samples for crystallization were further deglycosylated by *in situ* deglycosylation by endo-β-N-acetylglucosaminidase F1 treatment[97].

Large scale expression and purification of Dlp core mutants

Dlp core M175C/N367C and S168M, expressed in large scale for FSEC analysis, were expressed by transient transfection of adherent GnTI- HEK293S following the same procedure detailed above, dialysed against 20 mM Tris, pH 8, 500 mM NaCl, 20 mM imidazole using the same apparatus described for Dlp core, and initially purified by Nickel-based immobilized-metal affinity chromatography (HisTrap HP 5 ml, GE Healthcare), eluting the protein in step of 40 mM, 60 mM, 80 mM and 250 mM imidazole, with the rest of the buffer composition equal to the loading buffer. The purest fractions were collected and the proteins were further purified by SEC using the same experimental setting described for the wild-type protein.

Biolayer interferometry

Data were collected with Octet RED96 and Octet RED384 instruments (FortéBio). All steps were performed in 10 mM Tris, 35 mM NaCl, pH 8. Initially, Dip and Read Streptavidin Biosensors (FortéBio) were activated by dipping them in buffer for 10 min before starting the analysis. Sensors were then loaded with 0.05 mg/ml of biotinylated palmitoleoylated and non-palmitoleoylated human Wnt7a peptide. After baseline reference collection in wells containing only buffer, biosensors were dipped in analyte dilution series, before allowing dissociation in the same buffer wells where the baseline reference was collected. Peptide-loaded sensors tested against buffer were used for blank subtraction. Unloaded

sensors dipped into the analyte dilution series were used as reference to account for non-specific binding. Upon reference subtraction, data were analysed by the FortéBio (v.11.1) and SigmaPlot (v.14.0) data analysis software. The model used for curve fitting in kinetic and steady-state analysis was a 1:1 Langmuir binding isotherm model. Data were collected in three independent experiments.

Fluorescence-detection size-exclusion chromatography

Fluorescently labelled C16:0 palmitic acid bodipy FL C16 (4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene-3-hexadecanoic acid, Sigma-Aldrich) was incubated with the proteins of interest for 16 h at 4 °C in 10 mM Tris-HCl, pH 8, 35 mM NaCl, 7.5% (v/v) DMSO. The proteins were used at a concentration of 2 μM, and the fluorescent lipid was added at a 1:20 (w/w) protein:lipid ratio.
After incubation, the samples were run on a Superose 6 Increase 3.2/300 column (GE Healthcare) in a FSEC Prominence HPLC system (Shimadzu) equipped with a RF-10AXL fluorescent detector. The running buffer was 10 mM Tris-HCl, pH 8, 35 mM NaCl and the runs were performed at a flow rate of 0.08 ml/min. The fluorescent dye was excited at 505 nM and emission registered at 512 nM; absorbance at 280 nM was also registered to compare the elution profiles with those of the fluorescence emission. Bodipy FL C16 in absence of proteins was run as control. Data were collected in three independent experiments.

Crystallization, data collection and structure determination

Purified proteins were concentrated by ultrafiltration to 3.5 mg/ml and incubated for 16 h at 4 °C with palmitoleoylated human Wnt7a peptide (resuspended in 10 mM Tris-HCl, pH 8, 35 mM NaCl and added at a final concentration of 5 mM). After incubation, 96-well crystallization plates were set up at 20 °C using a Cartesian Technologies pipetting robot dispensing nanolitre-scale drops consisting of 100 nl of protein solution and 100 nl of reservoir solution. The crystallization conditions screened derived from a set of ~500 solutions from commercially available kits (Hampton Research, CA, USA and Emerald Bio-Structures, deCode Genetics, WA, USA) reformatted in-house in our Division (see ref. [98]). Crystals appeared after 2 weeks in a variety of conditions with 20% (w/v) PEG4000 as precipitant and different pH values.

Since these initial crystals diffracted at a maximum resolution of 3 Å with a high degree of anisotropy in the c* dimension, optimization of the crystal growth was performed, designing a two-dimensional pH/precipitant screening matrix (5–9 pH interval using 0.1 M MMT buffer[99], 15–25% w/v PEG4000 concentration interval). The best diffracting crystal grew in 16.25% (w/v) PEG4000, 0.1 M MMT pH 9.

Prior to crystal freezing, crystals were additionally soaked with palmitoleoylated peptide resuspended in DMSO, at a final concentration of 1 mM peptide in 10% (v/v) DMSO. The crystals were soaked for 1 h before flash-cooling by dipping into liquid nitrogen, using mother liquor supplemented with 25% (v/v) glycerol as cryoprotectant.

The diffraction data for the best diffracting crystal, whose structure is presented, were collected at 100 K at Diamond Light Source (Didcot, UK) beamline I04 equipped with an Eiger2 X 16M detector using a wavelength of 0.9795 Å.

Data were processed using Xia2[100], scaled using DIALS[101] and merged using AIMLESS[102,103].

The structure of Dlp^{core} in complex with human Wnt7a palmitoylated peptide was thus solved to a resolution of 2.21 Å by molecular replacement with PHASER[104] using the previously published apo-form of Dlp (PDB ID: 3ODN)[76] as a search model. Iterative cycles of refinement and manual rebuilding were performed using phenix.refine[105] (with coordinate, Translation-Libration-Screw and individual B-factor refinement) and COOT[106].

SMILES for the chemical structure of the peptide was designed using CHEMDRAW and peptide model and restraints for refinement were generated by the GRADE webserver (http://grade.globalphasing.org). Refinement models were validated with MOLPROBITY[107]. No Ramachandran outliers are present in the final model, with 97.83% of the residues in the favoured region of the Ramachandran plot.

Inspection of crystal contacts was performed using the Protein Interfaces, Surfaces and Assemblies (PISA) EMBL-EBI[108] webserver. Structural superpositions were performed in COOT, using the LSQ algorithm.

Structural figures were created with PYMOL. Lipid-binding cavity analysis and visualization were performed using CASTp[109] and its related PYMOL plugin.

Methods specific to Chapter 4

Construct design

Wntless 1st extracellular/intraluminal loop

The database references for human Wntless and *Drosophila melanogaster* Wntless are UNIPROT ID Q5T9L3 and Q95ST2, respectively. To help design of constructs, remote protein homology detection (using the HHPred[110] server), primary sequence conservation analysis (Clustal Omega[111] server), signal sequence prediction (SignalP[112] and Phobius[113] servers), secondary structure predictions (PSIPRED[114], QUICK2D[110] servers), disorder prediction (DisMeta server[115]) and *de novo* three-dimensional structure modelling (Robetta[116]) were preliminarily performed.

As described in **Chapter 4**, initially, constructs with different boundaries were designed around four conserved cysteines identified as the central structural core, and cloned in a pHL-Sec plasmid encoding the protein with a C-terminal 6x-Histag (**Table 2.3**).

No candidate expressed, therefore the expression system was changed to *Drosophila* S2 cells (Thermo Fisher Scientific). The construct expressing the *Drosophila* Wls loop, inserted in the pCO-Hygro plasmid (Thermo Fisher Scientific), with C-terminal 6x-His and HA (human influenza hemagglutinin) tags, was a kind gift from Dr Jean-Paul Vincent (Francis Crick Institute, London). The cells were also co-transfected with a pMT plasmid (Thermo Fisher Scientific)

encoding untagged *Xenopus laevis* Wnt8 under copper sulphate-inducible expression, kind gift of Dr Yuguang Zhao (Division of Structural Biology, Oxford). The construct features are summarized in **Table 2.4**.

Following observations from the *Drosophila* expression pointing to cysteines as potential cause of aggregation, truncated versions of the loop were generated in a modified pHL-Sec vector encoding a C-terminal mono-Venus fusion partner (for an easier visual assessment of protein behaviour in downstream processing) and a C-terminal 12x-His tag. These constructs are listed in **Table 2.5** and were transiently transfected in mammalian cells.

Finally, the best behaving among these constructs was also employed for the generation of stable mammalian cell lines, after having been cloned in pHR-CMV-TetO2_3C-mVenus-Twin-Strep (Addgene plasmid #113891), which allows doxycycline-inducible expression of the protein with a C-terminal mono-Venus fusion partner and a Twin-Strep tag (**Table 2.5**).

Table 2.3. First set of human Wntless 1st extracellular loop constructs designed.

Protein	Start residue	End residue	Vector (tag)	Expression system, transient transfection
Human Wntless	A 36	N 227	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells
	A 36	G 228	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells

	A 36	G 229	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells
	A 37	N 227	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells
	A 37	G 228	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells
	A 37	G 229	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells
	G 38	N 227	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells
	G 38	G 228	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells
	G 38	G 229	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells
	S 44	N 227	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells
	S 44	G 228	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells
	S 44	G 229	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells
	K 49	D 181	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells
	K 49	G 229	pHL-SEC (6x-His)	HEK293T, Wnt-expressing mouse L-cells

Table 2.4. Drosophila Wntless 1st EC loop constructs.

Protein	Start residue	End residue	Vector (tag)	Expression system, stable transfection
***Drosophila* Wntless**	A 36	N 236	pCO-Hygro (6xHis + HA)	*Drosophila* Schneider S2, constitutive
***Xenopus* Wnt8**	A 23	A 338	pMT-puro	*Drosophila* Schneider S2, inducible

Table 2.5. Cysteine-deleted versions of the human Wls 1st EC loop (in yellow the expressing constructs).

Protein	Start residue	End residue	Vector (tag)	Expression system
Human Wntless	A 37	E 147	pHL-SEC- mVenus (12x-His)	HEK293S GnTI–, transient transfection
	D 52	G 229	pHL-SEC- mVenus (12x-His)	HEK293S GnTI–, transient transfection
	T 60	G 229	pHL-SEC- mVenus (12x-His)	HEK293S GnTI–, transient transfection
	E 84	G 229	pHL-SEC- mVenus (12x-His)	HEK293S GnTI–, transient transfection
	A 87	G 229	pHL-SEC- mVenus (12x-His)	HEK293S GnTI–, transient transfection
	E102	G 229	pHL-SEC- mVenus (12x-His)	HEK293S GnTI–, transient transfection

	E84	G229	pHR-CMV-TetO2_3C-mVenus (Twin-Strep)	HEK293S GnTI– TetR, stable cell line by lentivirus infection, inducible

Wntless full-length/transmembrane protein

The following constructs were designed and cloned into pHR-CMV-TetO2_3C-mVenus-Twin-Strep (**Table 2.6**).

Table 2.6. Screened constructs for human Wls transmembrane protein (highlighted the expressing constructs). TM = Transmembrane region; EC = extracellular.

Protein	Start residue	End residue	Vector (tag)	Expression system	Construct features
Human Wntless	M 1	G 229	pHR-CMV-TetO2_3C-mVenus (Twin-Strep)	HEK293S GnTI– TetR, stable cell line by lentivirus infection, inducible	1st TM + 1st EC loop
	M 1	A 380	pHR-CMV-TetO2_3C-mVenus (Twin-Strep)	HEK293S GnTI– TetR, stable cell line by lentivirus infection, inducible	All EC loops

	M 1	S 495	pHR-CMV-TetO2_3C-mVenus (Twin-Strep)	HEK293S GnTI– TetR, stable cell line by lentivirus infection, inducible	w/o C-term tail
	M 1	E 541	pHR-CMV-TetO2_3C-mVenus (Twin-Strep)	HEK293S GnTI– TetR, stable cell line by lentivirus infection, inducible	Full-length
	A 36	A 380	pHR-CMV-TetO2_3C-mVenus (Twin-Strep)	HEK293S GnTI– TetR, stable cell line by lentivirus infection, inducible	All EC loops w/o 1st TM
	A 36	S 495	pHR-CMV-TetO2_3C-mVenus (Twin-Strep)	HEK293S GnTI– TetR, stable cell line by lentivirus infection, inducible	w/o C-term tail and 1st TM
	A 36	E 541	pHR-CMV-TetO2_3C-mVenus (Twin-Strep)	HEK293S GnTI– TetR, stable cell line by lentivirus infection, inducible	w/o 1st TM

Generation of stably-transfected Drosphila S2 cell lines

Stable cell transfection with *Drosphila* Wls loop-containing pCO-Hygro and *Xenopus* Wnt8-containing pMT-Puro was performed according to the manufacturer's instructions

(https://www.thermofisher.com/order/catalog/product/K412001) in a 6-well format in Schneider's *Drosophila* Medium (Thermo Fisher Scientific) supplemented with 2% FBS. The media was changed to 10% FBS after 8 hours. Selection of positively-transfected clones was achieved by adding 1 mg/ml hygromycin and 15 ug/ml puromycin for 2 weeks, then lowered at 500 ug/ml and 10 ug/ml respectively for maintenance. The cells were progressively expanded to a 75 cm^2 and 175 cm^2 flask-format, before being adapted to grow in suspension culture. Cells were cultured at an incubation temperature of 28 °C, without the need for CO_2 supplementation.

Generation of lentiviral-transduced HEK293S GnTI– TetR cell lines

4 ug DNA containing a 1:1:1 ratio of pHR-CMV-TetO2_3C-mVenus, encoding for the respective protein of interests, and viral packaging and envelop plasmids (psPAX2 and pMD2.G, see also glypican methods) were used to transfect virus-producing HEK293T Lenti-X cell with PEI (1:2.5 ratio) in a 6-well plate with 2 ml of DMEM/F12 containing 2% FBS. The virus-containing supernatant was collected after three days, filtered and used to infect target HEK293S GnTI- TetR cells for 3 days before changing media to DMEM/F12/10%FBS. Cell expansion and adaptation to suspension growth was performed as described for the glypican project.

Small-scale protein expression screening

Wntless 1st extracellular loop

For transiently transfected constructs, 4 µg of construct DNA, combined with Lipofectamine2000 at a 1:2 ratio were used to transfect ~80% confluent HEK293T cells or mouse L-cells in a six-well plate with 2 ml of DMEM/2% FBS. Results were tested by Western Blot, harvesting unconcentrated supernatant after cell growth at 37 °C and 30 °C for 72 and 96 hours, respectively, in presence and in absence of 4 mM valproic acid.

For *Drosophila* expression, cells were tested by WB at the end of the 2 weeks of selection.

For lentiviral expression, transduced cells were seeded to ~50% confluency in a six-well plate with 2 mL DMEM/F-12 containing 10% FBS. The day after, FBS content was changed to 2% and induction performed with 0.1, 0.5, 1, 2, 5 ug/ml doxycycline for 48h and 72h at 37 °C and 30 °C. Results were assessed by WB.

Wntless full-length/transmembrane protein

After lentiviral infection, induction of expression was performed in 6-well plates testing the same range of conditions described just above. Cells were then harvested and washed twice with PBS. The pellets were resuspended in 180 ul buffer (50 mM Tris, 150 mM NaCl, pH 8.0) and 20 ul 10% (w/v) DDM (n-Dodecyl β-D-maltoside):2% (w/v) CHS (Cholesteryl hemisuccinate), hence at a 5:1 ratio (final detergent concentration 1%:0.2% (w/v)). Membrane solubilization was

obtained by incubation for 3 hours at 4 °C. Samples were then centrifuged at 30,000xg for 1 hour and 10 ul from the supernatant were loaded onto a Superose 6 Increase 3.2/300 column (GE Healthcare) in an FSEC Prominence HPLC system (Shimadzu) equipped with a RF-10AXL fluorescent detector. The runs were performed at a flow rate of 0.08 ml/min and the FSEC buffer was: 50 mM Tris, 150 mM NaCl, pH 8.0, 0.03% DDM:CHS 5:1. The excitation wavelength was 515 nm and the emission was registered at 528 nm.

Wntless full-length/transmembrane protein detergent screening

To ensure a more reproducible comparison between detergents, a single pellet (2.5 g) was resuspended in 12 ml buffer (50 mM Tris, 150 mM NaCl, pH 8.0) and then separated in 100 ul aliquots. Different detergents from **Table 2.7** below were added to different aliquots to a 1% (w/v) final concentration and processed and analysed by FSEC as described for the expression screening. To test the protein behaviour even at reduced detergent concentrations (similar to the conditions the protein would encounter in downstream purification), 10 ul from each aliquot were also diluted in 90 ul resuspension buffer and assessed.

Table 2.7. Detergents conditions screened.

Detergent class	Detergent
CYMAL	Cymal 6
	Cymal 7

	Cymal 6/CHS 10:1
	Cymal 7/CHS 10:1
NG CLASS (Neopentyl glycol)	LMNG
	LMNG/CHS 20:1
	OGNG
	OGNG/CHS 20:1
	OGNG/CHS 40:1
	DMNG
	DMNG/CHS 10:1
	DMNG/CHS 20:1
MALTOSIDES	DDM-Low alpha
	DDM-Low alpha/CHS 10:1
	DDM-Low alpha/CHS 20:1
	DDM
	DDM/CHS 5:1
	DDM/CHS 20:1 (Solution grade)
	UDM/CHS 20:1
	UDM-Low alpha
	UDM-Low alpha/CHS 10:1
	UDM-Low alpha/CHS 20:1
	UDM-High alpha
	UDM-High alpha/CHS 10:1

THIOMALTHOSIDES	DDTM
	DDTM/CHS 20:1
	UDTM
	UDTM/CHS 20:1
	DTM
	DTM/CHS 20:1
GLUCOSYDES	ObG
	ObG/CHS 20:1

Large-scale protein expression and purification

Optimization attempts in Drosophila S2 cells (Wntless 1st extracellular loop)

Cells were harvested when reaching a density of 15x10^6/ml, and the supernatant dialyzed against IMAC loading buffer employing a QuixStand benchtop system (GE Healthcare) connected to a 60 cm Xampler Cartridge (GE Healthcare) with a 5-kDa nominal MWCO. For XWnt8 induction, copper sulphate at a final concentration of 500 μM was used, added at a cell density of 15x10^6/ml and left for 24 hours before harvesting. Several purification trials were performed using different combinations of chromatographic steps and buffer/additives (**Table 2.8**). Additives were either added before sample loading on IMAC to improve the initial protein behaviour, or to the aggregated gel filtration fractions trying to rescue aggregation.

Table 2.8. Summary of conditions tested for the purification of *Drosophila* Wntless 1st extracellular loop.

	Conditions tested
Purification strategy	IMAC + anion exchange
	IMAC + gel filtration
	IMAC + anion exchange + gel filtration
Buffer type and pH	Sodium-phosphate buffer, pH 7.2
	Hepes buffer, pH 7.4
	Tris buffer, pH 8
NaCl concentration in IMAC loading	300 to 750 mM in different trials
NaCl concentration in gel filtration	150 to 500 mM in different trials
Additives (alone or in combination)	L-arginine (100 mM and 200 mM)
	Glycerol (10%, v/v)
	NDSB-201 (200 mM)
	DTT (0.5 mM and 1 mM)

Optimization attempts in transiently transfected HEH293S GnTi- (Wntless 1st extracellular loop)

Upon large scale transfection using 12 roller bottles/batch (2.5 l/batch), cells were left incubating for 5 days at 30 °C, then the supernatant harvested and dialyzed against IMAC loading buffer as described for the attempts in *Drosophila*. As

above, several purification conditions were tested for different batches, listed in **Table 2.9**. 3c cleavage was performed at a 1:10 (w/w) protease:target protein ratio.

Table 2.9. Summary of conditions tested for the purification of human Wntless 1st extracellular loop

	Conditions tested
Purification strategy	IMAC + 3c cleavage + reverse IMAC + gel filtration
	IMAC + anion exchange + 3c cleavage + reverse IMAC + gel filtration
	IMAC + 3c cleavage + reverse IMAC
Buffer type and pH	Hepes buffer, pH 7.4
	Tris buffer, pH 8
	Citrate buffer, pH 4.5 and 5.5
	CHES buffer, pH 9.5
NaCl concentration in gel filtration	150 to 1M in different trials
Additives (alone or in combination)	Glycerol (10%, v/v)
	NDSB-201 (300 mM)
	0.001% w/v LMNG

Purification protocol for the optimized human Wntless 1st extracellular loop (lentiviral-transduced HEK293S GnTI- TetR, inducible expression)

Cells were grown either as adherent cells in 12 to 15 roller bottles/batch (2.5 to 3 l, induced after 3 days from seeding in roller bottles) or as suspension cells (2.5 l/batch, induced when cell density was $3x10^6$/ml), induced with 1 ug/ml doxycycline for 48 hours at 37 °C. Harvested supernatant was dialyzed against 50 mM Tris, pH 8, 200 mM NaCl and loaded o.n. with recirculation onto a 5 ml StreptactinXT High Capacity column (IBA Lifesciences). On-column cleavage was performed adding 3c protease at a 1:10 (w/w) ratio and incubating o.n. at 16 °C. Cleaved loop was washed out from the column and size-exclusion chromatography was performed in 50 mM Tris, pH 8, 200 mM NaCl using a Superdex 200 Increase 10_300 GL column (GE Healthcare).

Wntless full-length/transmembrane protein

HEK293S GnTI– TetR cells lentivirally-transduced with the construct of interest were grown as suspension cells and induced as described for the loop. Harvested pellet was resuspended in 50 mM Tris, pH 8, 200 mM NaCl, 1% of the appropriate detergent, 1x cOmplete Protease Inhibitor (Sigma Aldrich), incubated at 4 °C for 3 h and centrifuged at 120,000xg for 45 min at 4 °C and loaded o.n. with recirculation on a Streptactin column. The column was washed with a buffer containing 50 mM Tris, pH 8, 200 mM NaCl, and a concentration of the appropriate detergent at 3 times the critical micellar concentration (CMC). On-

column cleavage was performed as described above for the optimized human Wntless 1st extracellular loop and the cleaved protein recovered. The cleaved protein was then run on a gel filtration Superdex 200 Increase 10_300 GL column in final buffer 50 mM Tris, pH 8, 200 mM NaCl, appropriate detergent at 3 times the CMC.

SMALP extraction and purification of Wntless full-length/transmembrane protein

The protocol for membrane protein extraction in SMALPs is based on ref. [117-119]. Cell pellets were resuspended in 10 mM Hepes, pH 7.5, 1 mM EDTA, 1x cOmplete Protease Inhibitor (Sigma Aldrich) at a 1:3 (v/v) pellet:buffer ratio, and the mixture was incubated at 4 °C for 2 h with rotation. The solution was then centrifuged at 120,000xg for 45 min at 4 °C, the supernatant decanted, and the pellets were resuspended in 40 ml of 10 mM Hepes, pH 7.5, 1 M NaCl, 1x cOmplete Protease Inhibitor (Sigma Aldrich). The resuspended membranes were then homogenized, centrifuged as before, resuspended in 40 ml of 10 mM Hepes, pH 7.5, 1 mM EDTA, 1x cOmplete Protease Inhibitor (Sigma Aldrich) after discarding the supernatant, homogenized and centrifuged again.

Isolated membranes were thus resuspended in 50 mM Tris, pH 8.0, 500 mM NaCl, 10% (v/v) glycerol, with the help of a pre-chilled homogeniser, to a final membrane concentration of 80 mg/ml. A 5% (w/v) SMA solution in 50 mM Tris, pH 8.0, 500 mM NaCl, 10% (v/v) glycerol was added to the membrane at a 1:1 (v/v) ratio and the sample was incubated with rotation o.n. at 4 °C.

The solution was then diluted 10 times in 50 mM Tris, pH 8, 200 mM NaCl, and loaded o.n. with recirculation onto a 5 ml StreptactinXT High Capacity column (IBA Lifesciences). Elution was performed by running 35 ml of 1x buffer BXT (100 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA, IBA Lifesciences) on the column at a flow-rate of 0.5 ml/min.

Methods specific to Chapter 5

Construct design

For the studies presented in **Chapter 5**, the construct harbouring the previously solved apo- and palmitoleate-bound structures[85] of human Notum (UNIPROT database reference Q6P988) was used: it encompasses aa. S81-T451 of the protein sequence, with the C330S mutation, and is cloned into pHL-Sec with a C-terminal 6xHis-tag.

Large scale expression and purification of Notum

The protein was produced, secreted in the culture media, by transient transfection of adherent HEK293T cells, using standard protocols, already described above. The class I α-mannosidase inhibitor, kifunensine was also added during the transfection step. The cells were left in incubation for 5 days at 37 °C in a roller bottle incubator before harvesting.

The supernatant was collected by centrifugation at 8000xg for 45 min. After that, it was dialyzed against PBS supplemented with 20 mM imidazole, using a QuixStand benchtop system (GE Healthcare) connected to a 60 cm Xampler Cartridge (GE Healthcare) with a 10-kDa nominal MWCO. The sample was then loaded onto a His-Trap HP 5 ml column, for IMAC-based purification. The protein was eluted with a single step of PBS + 250 mM imidazole.

Further purification to homogeneity was obtained by running size-exclusion chromatography (HiLoad Superdex 200 16_600 pg, GE Healthcare), with the column equilibrated in 20 mM Hepes, 150 mM NaCl, pH 7.4. Before SEC, the sample for crystallization was deglycosylated by endo-β-N-acetylglucosaminidase F1 treatment. The purity of the eluted fractions was assessed by SDS-PAGE.

Crystallization of Notum, soaking of inhibitors and structure determination

As the previously solved structures of Notum were obtained in 9 different crystal forms, the associated crystallization conditions were screened and optimized to identify the experimental set-up most consistently harbouring high resolution-diffracting crystals (ideally ~ or > 1.5 Å). This was found in a variation of the original crystal form III (1.5 M ammonium sulphate, 0.1 M sodium acetate, pH 4.6). In particular, high diffracting crystals grew in 1.4 ammonium sulphate, 0.1 citric acid, pH 4 at a concentration of 4.5 mg/ml. In the successive iterations of drug development and screening, crystal formation was further optimized by

designing two-dimensional pH/precipitant screening matrices with a 3.7–4.3 pH interval and 1.3-1.5 M ammonium sulphate concentration interval.

Upon growth, crystals were soaked with inhibitors. Compounds were resuspended in 100% (v/v) DMSO and then diluted to a final concentration of 15% (v/v) DMSO in the crystallization condition supplemented with 40% ethylene glycol (v/v).

In the initial iterations of drug development, soaking was performed in triplicate for 5 min, 10 min or overnight, and diffraction datasets were collected for each time point after flash-freezing the crystals in liquid nitrogen. As the crystals didn't show to be damaged by overnight incubation with inhibitors - either morphologically when visually monitored during soaking, and in their diffraction quality - and this time point provided the best-quality difference maps upon structure solution, in the successive rounds of compound screening overnight soaking was the procedure of choice.

Diffraction data were collected at 100 K at Diamond Light Source either at beamline I03 or I04-1 for the different compounds, and automatically processed generally through the Xia2 pipeline. Molecular replacement using the apo-form of Notum (PBD ID: 4UZ1) was performed using PHASER, followed by iterative cycles of refinement using phenix.refine and manual rebuilding with COOT. Ligand restraints were generated from SMILES using the GRADE webserver (http://grade.globalphasing.org). Refinement models were validated with MOLPROBITY. The chemical contacts between the compounds and Notum residues were analysed by structure inspection in PYMOL and contact analysis using LIGPLOT+[120]. Structural figures were created with PYMOL.

Methods specific to Chapter 6

Construct design, expression, purification and preliminary crystallization of human Gpc6

Human Gpc6 aa. 24-488 (UNIPROT database reference Q9Y625) wild-type and mutants were cloned into the pHR-CMV-TetO2_3C-AVI-HIS vector (Addgene plasmid #113887) encoding for a 3c-cleavable C-terminal Avi-6xHis-tag.

The mutants L112C/D313C, N115C/D313C, S119C/D316C and M123C/D316C were designed based on three-dimensional modelling using the servers SWISS-MODEL[121] and Phyre2[122]. The aim of the mutations was to identify the best pairs of residues at an appropriate distance to form an artificial disulphide bond to lock the putative Gpc6 lipid-binding pocket (identified by similarity with the Dlp structure) in the closed, binding-ineffective conformation for downstream biophysical and biochemical studies.

The proteins were expressed, secreted in the culture media, by transient transfection of adherent HEK293T cells in presence of kifunensine, as described above, and collected after 5 days of incubation at 37 °C by centrifugation at 8000xg for 45 min.

For each protein, the supernatant was then dialyzed against 20 mM Tris, pH 8.0, 500 mM NaCl, 20 mM imidazole, using a QuixStand benchtop system (GE Healthcare) connected to a 60 cm Xampler Cartridge (GE Healthcare) with a 10-kDa nominal MWCO, and loaded onto a His-Trap HP 5 ml column, for IMAC-based purification. The protein was eluted with a single step of 20 mM Tris, pH 8.0, 500 mM NaCl, 500 mM imidazole.

The eluted sample was concentrated up to a volume of 1 ml using 10 KDa cut-off Amicon concentrators (Millipore) and then injected into a HiTrap Desalting 5 ml column (GE Healthcare) in buffer 20 mM Tris, 500 mM NaCl, pH 8, 20 mM Imidazole. Fractions corresponding to the protein peak were pooled and 3c-protease added at a protease:protein ratio of 1:50 w/w, incubated for 48h at 4 °C. The sample was then loaded on a HisTrap HP 5 ml (GE Healthcare) and the flow-through, containing the cleaved Gpc6, collected in 1 ml fractions. The fractions corresponding to the protein peak were pooled and concentrated to 0.5 ml using 10 KDa cut-off Amicon concentrators.

The concentrated fractions from the reverse IMAC were injected into a Superdex 200 Increase 10/300 GL column (GE Healthcare), with the column equilibrated in 10 mM Tris pH 8, 35 mM NaCl.

For the crystallization of wild-type Gpc6 apo form, the protein was concentrated at 3 mg/ml and nanolitre-scale drops consisting of 100 nl of protein solution and 100 nl of reservoir solution were set up using a Cartesian Technologies pipetting robot. Crystals in the shape of thin plates appeared in 14% (w/v) PEG 5000 MME, 0.08 M magnesium acetate, 0.1 M Sodium Citrate, pH 6, in presence of 8% (v/v) of the polymer additive pentaerythritol ethoxylate (3/4 EO/OH).

Preliminary BLI analysis of Dlp^{core} interaction with Hedgehog palmitate-mimicking peptide

A biotinylated peptide encompassing aa. 24 - 34 of human Sonic Hedgehog (UNIPROT database reference Q15465) and incorporating N-terminal palmitate (kind gift of Dr Christian Siebold, Division of Structural Biology, University of

Oxford) was used to load Dip and Read Streptavidin Biosensors (FortéBio), and was tested against a single concentration (1.25 μM) of the wild-type Dlp^{core} construct for which the Dlp structure in complex with the palmitoleoylated human Wnt7a peptide was obtained (hence, the Dlp^{core} mutated for the furin cleavage site, see above). All the steps and the analysis were performed as described above for the Dlp-Wnt peptide interaction. Data were collected on an Octet RED96 instrument.

Chapter 3 - Understanding Wnt extracellular spreading: the role of glypicans.[††]

Identification of the Wnt carrier molecule in *Drosophila*

To provide a molecular explanation about how Wnt is able to spread at long distances despite its palmitoleate modification hampering solubility, it was of paramount importance to first identify, among the plethora of possible interactors proposed in the literature to escort it in its travel, the one with the most convincing experimental evidence to be involved in direct interaction with the morphogen and its lipid *in vivo*.

To the aim, as already mentioned, a collaboration was established with the Vincent lab at the Francis Crick in London, long-standing collaborator of our lab for the functional counterpart of our studies. The Vincent lab mainly uses the already widely cited *Drosophila melanogaster* wing imaginal discs to perform its research into Wnt signalling. In more detail, the imaginal discs are sac-like epithelial structures that form in the larval stage of the insect development and constitute the precursors for the major adult organs; the wing disc is in particular well suited for developmental studies since its antero-posterior and dorso-ventral patterning are controlled by morphogens acting and spreading in distinct but

[††] The work presented in this chapter has been published in the following article, of which I am co-first author: ref. 95 McGough, I. J. *et al.* Glypicans shield the Wnt lipid moiety to enable signalling at a distance. *Nature*, doi:10.1038/s41586-020-2498-z (2020). The functional *Drosophila* studies have been performed by Dr Ian McGough (Vincent lab, Francis Crick Institute, London), while the Wnt peptide synthesis was carried out by Karen Beckett, Dhira Joshi and Nicola O'Reilly (Francis Crick Institute, London). Some figures are derived from the article, also some parts of the text are quoted *verbatim*.

intersecting compartments, and this helps to study the functional relationships of the latter[123]. Moreover, it has been demonstrated that, in the wing disc, Wnt can activate target genes at a distance of more than 10 cell diameters[47], making it an ideal model to study the mechanisms involved in the morphogen extracellular trafficking and long-range signalling. Finally, *Drosophila* also represent a model organism whose genetic engineering is relatively straightforward, fast and effective.

Indeed, fly genetic modifications were central to this part of the project; in addition, a Wnt morphotrap[124] approach was also devised. The latter was based on the expression of a membrane-tethered anti-GFP nanobody in the wing disc: upon tagging the Wnt homologue of *Drosophila*, Wingless, with GFP, the morphogen is readily trapped, and any bound interactor can be evaluated by, for instance, colocalization experiments.

To evaluate potential Wnt partners, the attention initially focused on lipoprotein particles (LPP) and exosomes. Whereas overexpression of the morphotrap captured endogenously expressed GFP-Wg at the cell surface as expected, no colocalization was observed either with lipophorin (a key component of LPP, see **Introduction**) or the ESCRT component Hrs (Endosomal Sorting Complexes Required for Transport), an exosome marker[125] (**Figure 3.1 a**). Genetic manipulations also confirmed the lack of experimental support for these two models (data not shown here, can be found in Extended Data Figure 2 of our paper[95]). With respect to LPP, deletion of *Lpr1* and *Lpr2* (LPP receptors) didn't in fact alter the extracellular distribution of Wingless; neither did the overexpression of the LPP receptor isoform *Lrp2E*, despite depleting, as foreseen, the

extracellular pool of LPPs in control experiments. For the exosome hypothesis, in absence of other validated tools to specifically inhibit their formation, a dominant negative form of the Vacuolar Sorting Protein 4 (VSP4, component of the ESCRT machinery)[126], VSP^{DN}, was overexpressed, to inhibit the formation of the multivesicular bodies, that are the last intracellular precursors of exosomes. Once again, Wg extracellular distribution was not modified, and the same held for deletions of the ESCRT protein Hrs. These results further confirmed that both LPP and exosomes do not play a direct role in binding Wnt and shielding its lipid *in vivo*, at least in the *Drosophila* wing model used here.

Next, a role for the lipocalin-like extracellular transporter Swim was also ruled out. In fact, genetic knockout for the entire region coding for Swim didn't affect the spreading of extracellular Wg, didn't alter the activation of the Wnt target gene *Distal-less* (**Figure 3.1 b**), and the flies homozygous for the deletion were healthy and fertile with no apparent phenotypic differences compared to the wild-type controls.

Finally, the possibility that the hydrophobic moiety of Wnt is sequestered by the formation of morphogen multimers or micelles was also considered, but deemed unlikely based on the experimental results: when two differentially tagged Wingless, HA-Wg and GFP-Wg, were co-expressed in the imaginal discs, GFP-Wg was captured by the morphotrap, but no colocalization was observed with HA-Wg (**Figure 3.1 c**).

In summary, no convincing evidence to support any of the main models proposed to explain Wnt transport in the extracellular space was found so far.

We therefore looked for potential alternatives carriers of Wnt. As noted in the **Introduction**, a further family of extracellular proteins, heparan sulphate proteoglycans or glypicans, had also been associated to Wnt signalling modulation; a direct role in the Wnt solubility problem was however neglected since the available literature pointed towards an interaction mainly based on relatively weak and non-specific polar or charge-based binding with the long, negatively charged heparan-sulphate chains – similar to those reported for glypicans with many other signalling proteins.

Triggered however by the lack of *in vivo* evidence for all the other modes of Wnt transport, we focused our investigation on this class of proteins.

Surprisingly, in this case, our collaborators' data not only confirmed the association (**Figure 3.1 d**), but additionally hinted towards a direct interaction with the Wnt lipid: indeed, mutation of the palmitoleoylated Ser239 of Wingless into an alanine – to prevent Ser esterification and morphogen acylation – was sufficient to disrupt Wg capturing by Dlp (Dally-like protein, one of the two glypicans present in *Drosophila*, as discussed in the **Introduction**). The fact that the observed effect was due to impaired Wg expression was also excluded, since WgS239A was readily produced and trapped by the morphotrap in control experiments (**Figure 3.1 e**).

These unexpected but encouraging data therefore prompted us to focus on the glypicans as the best candidates in our efforts to explain from a molecular and structural perspective the mechanisms responsible for the regulation of Wnt extracellular movement.

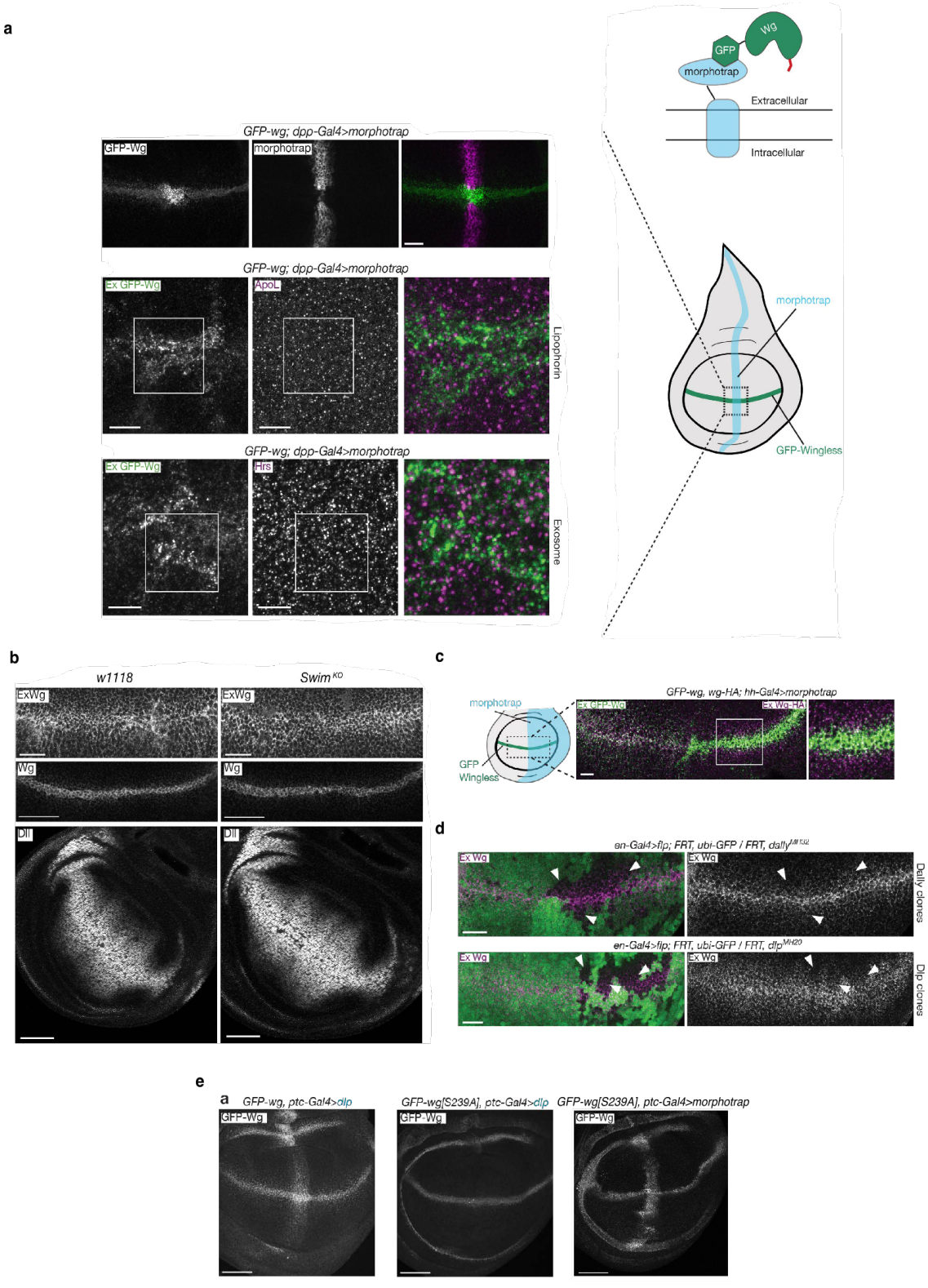


Figure 3.1. *In vivo* studies in *Drosophila* ruled out the involvement of lipoprotein particles, exosomes, Swim and multimers in Wnt transport.
a) Model explaining the morphotrap approach, and morphotrap-based co-localization experiments showing no interaction between morphotrap enriched

GFP-Wg (green) and endogenous lipoprotein component ApoL (violet, middle subpanel) or endogenous exosome marker Hrs (violet, bottom subpanel). The top subpanel shows the enrichment of GFP-Wingless by the transversally expressed morphotrap in the zone where the two expression domains overlap. **b)** *Swim* knock-out does not alter the expression and distribution of extracellular Wg and the expression of the low-threshold target gene *distal-less* (Dll). **c)** The morphotrap expressed in the disc posterior compartment traps GFP-Wg (green) but not HA-Wg (violet), ruling out that extracellular Wg forms multimers. **d)** Extracellular Wingless distribution is reduced in clones lacking the *Drosophila* glypicans Dlp and Dally, to a greater extent in absence of Dlp (clones marked by the absence of GFP (white arrowheads)). **e)** Wg S239A mutation, removing Wnt lipidation, impairs extracellular Wg trapping by Dlp expressed in a transversal stripe (middle subpanel), whereas wild-type Wg is readily captured (left subpanel); in control experiments (right subpanel), when the morphotrap is expressed in the stripe, it captures WgS239A, ruling out reduced expression or secretion of Wg due to the mutation. The panels were derived and adapted from ref. [95] and the data obtained and the figures prepared by Dr Ian McGough (Francis Crick Institute, London).

Protein expression and purification

In line with the findings just discussed, we first concentrated on the expression and purification of the core protein of Dlp (i.e. without the GAG attachment sites), to evaluate if it had in fact lipid-binding capacity.

The construct boundaries (aa. 74-617) were based on the structure of the apo form of Dlp, previously obtained by Kim et al. (2011)[76], with some modifications. Namely, since analysis of the Dlp sequence shows that the protein contains both a cryptic and a consensus furin cleavage site in a solvent-exposed loop (at aa. positions 400 and 437 respectively), and Kim's paper reported this leading to heterogeneous protein populations due to uncomplete processing, two different constructs were designed: a construct ($Dlp^{core\ \Delta 400\text{-}437}$) with a deletion between the two furin cleavage sites, preserving the sites but removing any intervening

amino acids to allow a more complete processing[‡‡] - mirroring what was successfully done in Kim's paper; and a construct spanning the full aa. 74-617 region without any deletion but mutation of both furin sites ($Dlp^{core\ furinmut}$, containing mutations K398Q, K399A, R402Q, R438Q, R441A).

In summary, $Dlp^{core\ \Delta 400\text{-}437}$ was designed to undergo complete furin cleavage, while $Dlp^{core\ furinmut}$ was designed to be cleavage-resistant (**Figure 3.2**). The rationale for the design of multiple constructs was, on the one hand, to enhance the production of more homogenous batches of protein, and, on the other hand, to investigate if furin cleavage was somehow involved in lipid binding, as this is not only a feature of Dlp, but a more general characteristic also present in other members of the family (e.g. Dally, human Gpc 3, 4 and 6, bioinformatically predicted to contain furin cleavage sites by the ProP 1.0 web server[127]).

As further detailed in Material and Methods, the constructs were designed with a cleavable Twin-strep tag at the C-terminus and initially used to stably transduce HEK293S GnTI(-) TetR cells, through lentiviral delivery. Stable transfection was used both to facilitated downstream production of repeated large-scale batches of protein, and because, based on the experience of our lab, higher yields are usually obtained when the vector suite we employed, pHR-STRUBI[93], is used for stable rather than transient expression. However, a further advantage of the plasmid is its suitability for transient expression as well, and indeed in subsequent batches of protein production transient transfection was also employed.

[‡‡] Note that, as confirmed by the structure, the two cleaved subunits remain connected by inter-subunit disulphide bonds.

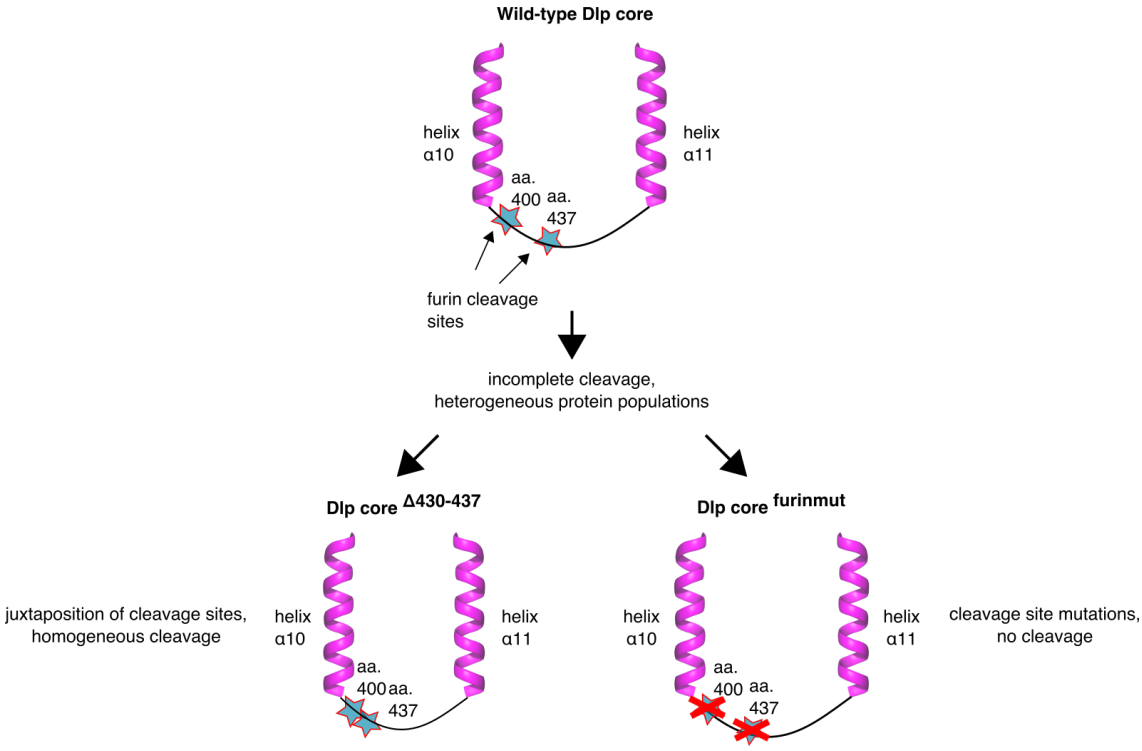


Figure 3.2. Schematics and rationale for Dlp^{core} construct design.
Helices numbered according to the apo structure of Dlp[76].

Upon expansion of the infected cells, a Western blot on the conditioned media was performed to check the success of the procedure. Bands at the expected molecular weight were visible for the two constructs when assessed with an HRP-conjugated streptavidin-based probe reactive against the Twin-strep tag. Moreover, the blot confirmed that the Dlp$^{core\ \Delta 400\text{-}437}$ underwent cleavage upon furin processing, while Dlp$^{core\ furinmut}$ was resistant to the latter as envisaged in construct design (**Figure 3.3**).

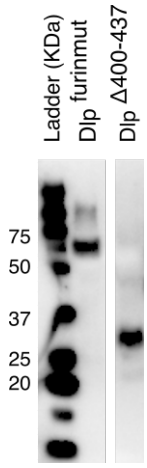


Figure 3.3. Western Blot on small-scale expression of the Dlp^{core} constructs from lentiviral-transduced HEK293S GnTI(-) TetR cells.
The proteins were probed with a StrepTactin-HRP conjugate (see Materials and Methods); the molecular weight ladder (Precision Plus Protein™ WesternC™, Biorad) contains strep-tagged protein standards. Since the gel was run in reducing conditions, only the tagged, C-terminal subunit of Dlp$^{core\ \Delta 400\text{-}437}$ is visible in the blot, confirming that the construct undergoes complete furin processing, while only a band at the expected molecular weight for the full-length protein (~ 66 KDa) is present for Dlp$^{core\ furinmut}$, suggesting that the designed mutations abolish furin cleavage.

The proteins were therefore scaled up and purified by a two-step purification protocol. Initially the conditioned media was dialyzed into 50 mM Tris-HCl, pH 8, 200 mM NaCl and underwent affinity-purification using Strep-tag^{TM} affinity chromatography. The tag was then cleaved and the protein further purified using size-exclusion chromatography in 10 mM Tris, 35 mM NaCl, pH 8. The low-salt buffer was chosen to facilitate downstream crystallization with a palmitoleoylated Wnt-mimicking peptide used as Wnt surrogate (see below), since the latter tended to be more soluble in low-salt conditions. As their monodisperse peaks in size-exclusion chromatography show, the proteins tolerated the buffer well (**Figure 3.4 a** and **b**).

We also purified the core protein of the other *Drosophila* glypican, Dally, for which we could count on the expertise of the Siebold lab in our Division, who previously worked on the protein. In this case, the selected construct spanned aa. 65-541, was expressed with a 6x-His-tag transiently in HEK293T cells, and purified to homogeneity by a two-step purification, composed of an affinity chromatography step and a size-exclusion step (see **Materials and Methods**) (**Figure 3.4 c**).

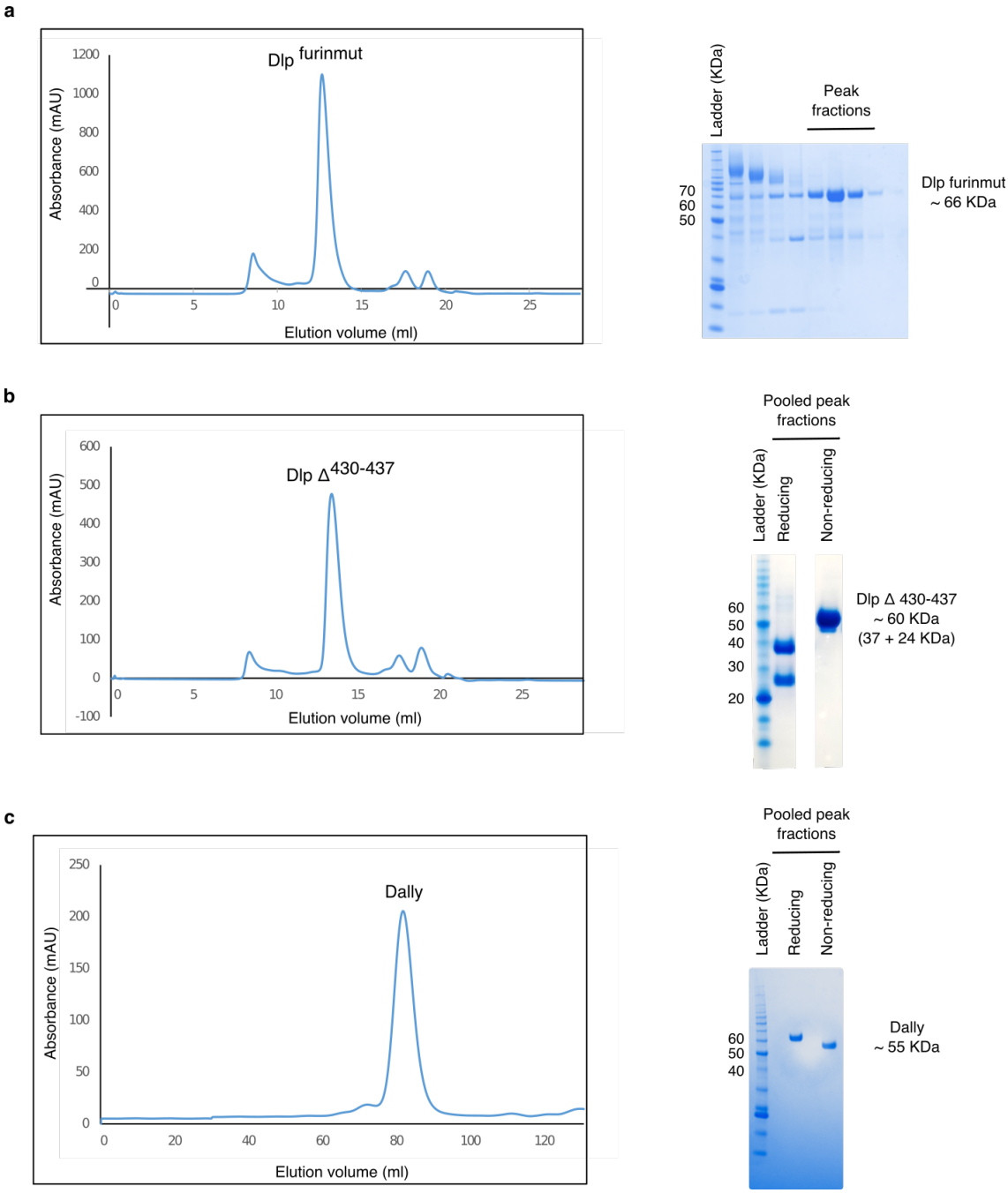


Figure 3.4. Representative size-exclusion chromatography profiles for the two Dlp^{core} constructs and Dally^{core}.
a) Size-exclusion chromatography (Superdex 200 Increase 10_300 GL, GE Healthcare) elution profile and related SDS-PAGE for Dlp$^{core\ furinmut}$. **b)** Size-exclusion chromatography (Superdex 200 Increase 10_300 GL, GE Healthcare) elution profile and SDS-PAGE of the pooled peak fractions in reducing and non-reducing conditions for Dlp$^{core\ \Delta 400\text{-}437}$. The furin-cleaved N-terminal and C-terminal subunits, held together by disulphides, run as separate bands in reducing conditions, and as a single band in non-reducing conditions, as expected. **c)** Size-exclusion chromatography (HiLoad 16_600 Superdex 200 pg,

GE Healthcare) elution profile and SDS-PAGE of the pooled peak fractions for Dally^{core}.

Biophysical studies on glypican core - Wnt lipid interactions

As a first step, we set out to confirm - using purified proteins and biophysical methods for protein-protein analysis - the functional data observed by our collaborators pointing towards a direct interaction between the core proteins of glypicans and Wnt. This aimed at ruling out the possible involvement of any additional and unidentified factor in determining the *in vivo* outcome, and, in case of a biophysical confirmation of the data, at quantifying the strength of the interaction, obtaining an affinity constant.

As a surrogate Wnt, we used a synthetic peptide encompassing aa. 204-213 of the human Wnt7a sequence, which includes the conserved Wnt serine modified with the palmitoleic acid. We used a Wnt peptide since, as already mentioned in the **Introduction**, the full-length Wnt proteins are extremely challenging proteins to express in quantities suitable for biochemical and structural studies, because of their reduced solubility, their complicated folding, and propensity to aggregate and lose activity when purified[128]. The choice for the Wnt7a sequence was dictated by the higher relative solubility of this peptide compared to the *Drosophila* Wingless sequence. Structurally, the peptide mimics the loop, stabilized by a conserved disulphide bond, that is located at the tip of the β-hairpin of the Wnt "thumb" (**Figure 3.5 a**): because of this bond, it is therefore a circular peptide. To test if the lipid of Wnt was crucial to the interaction, the peptides were synthesized in a palmitoleoylated and non-palmitoleoylated form, different only

for the presence/absence of this acylation, and both were furthermore biotinylated when used in bio-layer interferometry (BLI) assays.

BLI was in fact our method of choice for biophysical analysis. Tips conjugated with streptavidin were employed, onto which the biotinylated Wnt peptides were immobilized. The different Dlp constructs and Dally were then used as the analyte tested in solution.

The data reported in **Figure 3.5 b** clearly show how a specific interaction can only be observed for the core protein of Dlp tested against the palmitoleoylated Wnt peptide. On the other hand, Dlp^{core} vs. non-palmitoleoylated Wnt peptide didn't show any signal, in line with the *in vivo* observation that the S239A mutation in Wingless, removing its lipidation, was sufficient to abrogate morphogen trapping by Dlp. Moreover, $Dally^{core}$ vs. palmitoleoylated Wnt peptide didn't show signal (and the same therefore holds also for the non-palmitoleoylated Wnt peptide, as the two peptide forms are identical except for the additional presence of the lipid in the palmitoleoylated one).

Additionally, no difference in the binding to lipidated peptide was observed between $Dlp^{core\ \Delta 400\text{-}437}$ and $Dlp^{core\ furinmut}$, with almost identical Kd (compare **Figure 3.5 c** with **Figure 3.5 b**). This ruled out the possible role of furin processing as determinant of the interaction[§§].

The interaction presented an affinity in the high nanomolar range (Kd ~ 950 nM, **Figure 3.5 d**). However, this has to be considered as an estimate and not as an exact Kd as we could not perfectly fit a 1:1 Langmuir model to the interaction, due to the impossibility of completely removing aspecific binding. The latter, in fact,

[§§] For this reason, in the remainder of the text, the protein is only referred as "Dlp core". The crystal structure presented later in the text was obtained with the furin mutant construct.

was a constant of the BLI tests attempted with glypicans, and was recognized as non-specific since it was present even when the glypicans were tested with tips without any immobilized ligand. Several trials at removing it were attempted using strategies suggested by the manufacturer, including the use of BSA and detergents (Tween-20): however, in the case of BSA and detergent concentrations below 0.05% Tween-20, they were not sufficient to remove non-specific binding; detergent concentrations above 0.05% were, but they also abrogated, as might be expected, the lipid-based interaction between Dlp and the Wnt palmitoleoylated peptide. Finally, a solution was found by running the same analyte series against tips with and without ligand immobilized, and using the tips without ligand as reference tips to subtract their signal from the total signal elicited in the presence of the immobilized Wnt peptides. This allowed us to obtain the curves shown in **Figure 3.5 b** and **c**, but the effect of the non-specific signal on the curve shapes disallowed a perfect fitting of the 1:1 interaction model, as noted.

In summary, the biophysical data obtained by BLI with pure proteins and peptides not only confirmed that a direct binding between the glypican core and Wnt occurs - and that it is centred on the Wnt lipid - but also pointed towards an interaction specific for a subset of glypicans only (Dlp but not Dally in the case of *Drosophila*). The observation was further corroborated and expanded by *in vivo* data from our collaborators. In fact, Gpc4 and Gpc6 (the two closest human homologues of Dlp) trapped extracellular Wg like Dlp, but the effect was not observed in the case of Gpc3 and Gpc5 (the two closest human homologues of Dally) (**Figure 3.6**).

Hence, the differential ability of glypicans to bind the Wnt lipid appeared to be a general feature of this class of proteins, and not simply a specificity of *Drosophila*. Encouraged by these combined biophysical and *in vivo* findings, we next wanted to uncover the structural basis for the newly discovered interaction.

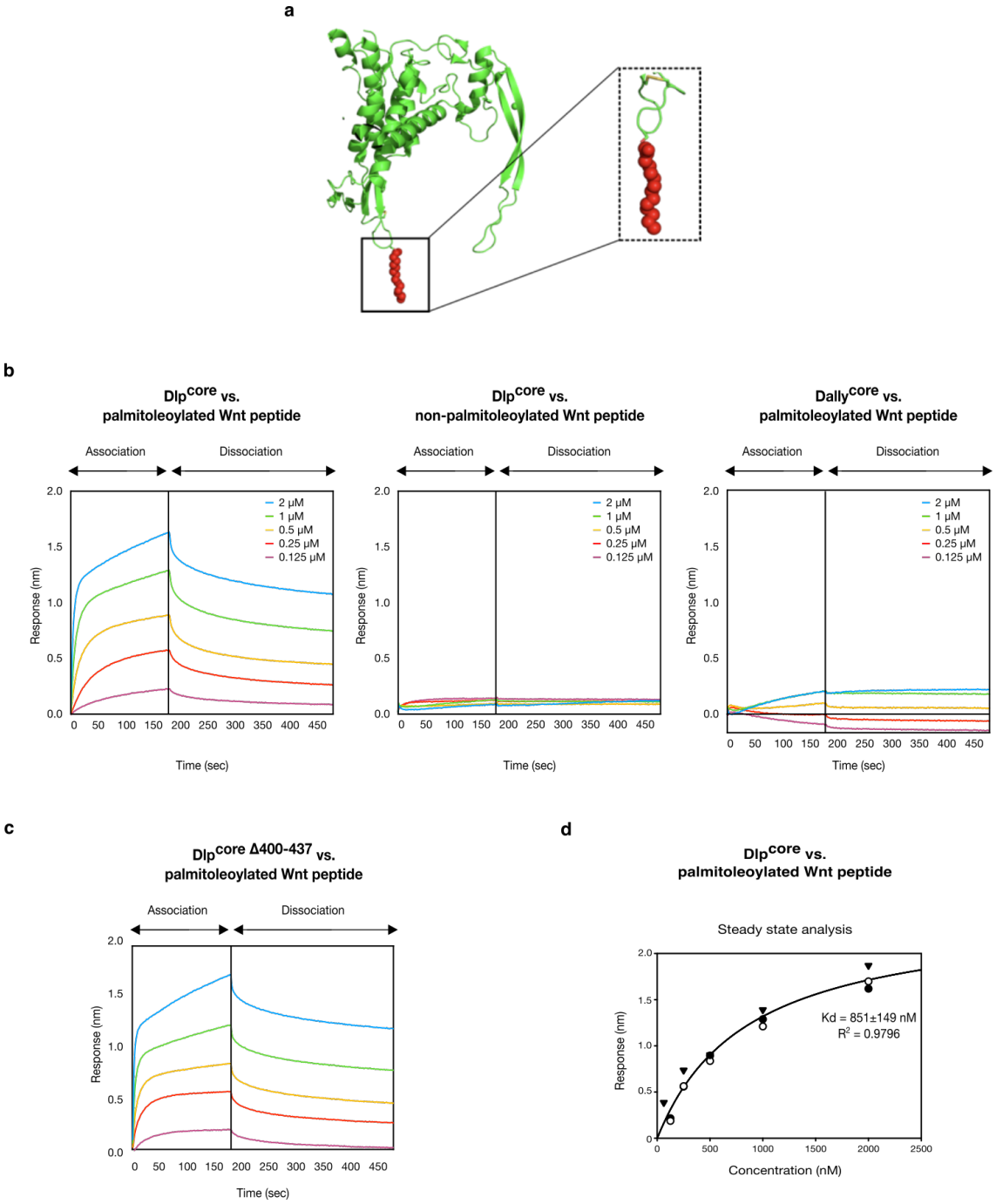


Figure 3.5. Biophysical analysis confirmed a direct interaction between Dlp^{core} and the Wnt lipid.
a) Model of the circular, disulphide stabilized hWnt7a peptide used in this study, mapped onto the XWnt8 structure (PDB ID: 4F0A). **b)** Representative reference-subtracted biolayer interferometry traces of Dlp^{core} and Dally^{core} tested against biosensors loaded with palmitoleoylated and non- palmitoleoylated Wnt peptide, showing how only Dlp^{core} vs. palmitoleoylated peptide presented a signal. The Dlp^{core} traces shown in this panel were obtained for the Dlp$^{core\ furinmut}$ construct. **c)** Representative reference-subtracted biolayer interferometry traces of Dlp$^{core\ 400-}$

[437] tested against biosensors loaded with palmitoleoylated Wnt peptide, eliciting a similar signal to $Dlp^{core\ furinmut}$. **d)** Steady-state analysis of the Dlp^{core} vs. palmitoleoylated hWnt7a peptide interaction measured by BLI. Kd is calculated from a global fit of three independent experiments and should be considered an estimate and not an exact value for the reasons detailed in the text. Panels b and d were derived and adapted from ref. [95].

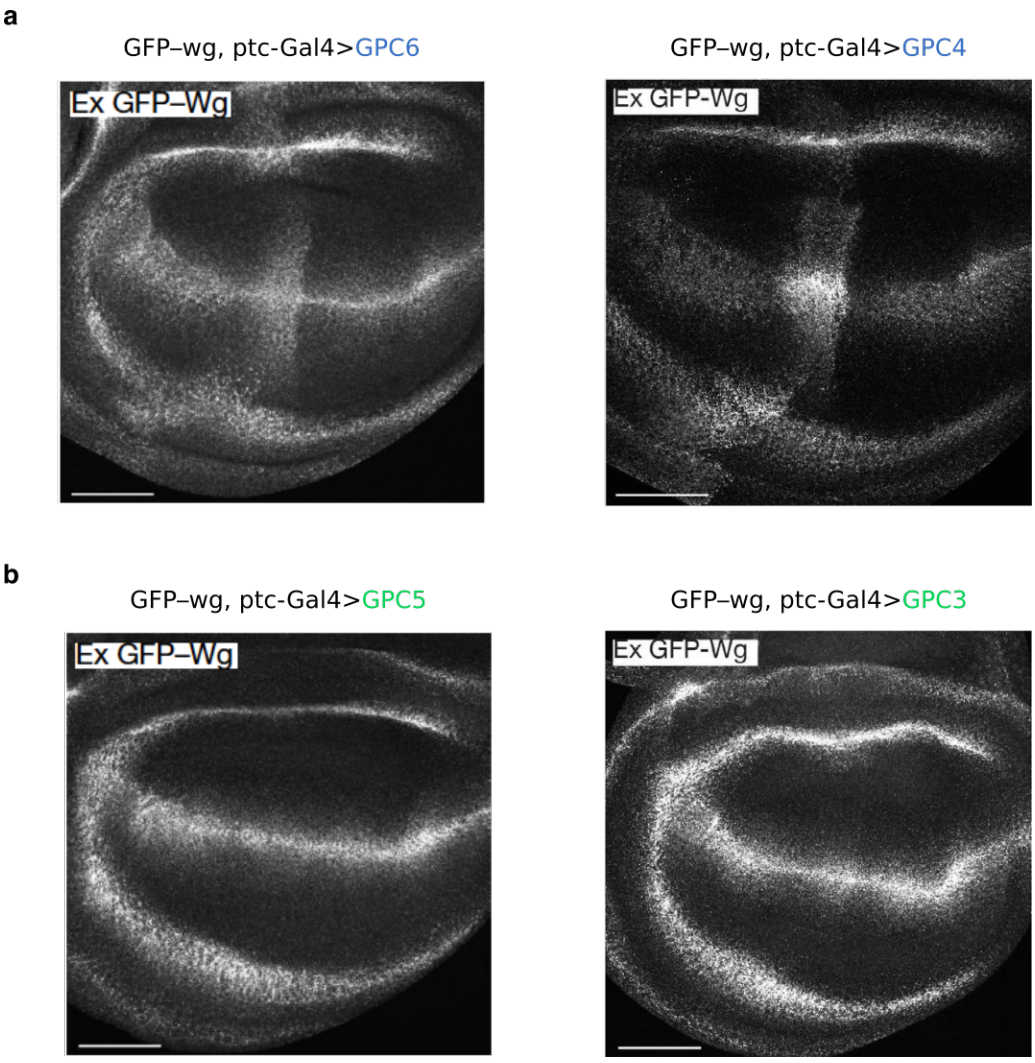


Figure 3.6. *In vivo* data confirmed the differential trapping activity between glypicans members.
a) Human homologues of Dlp (Gpc6 and Gpc4), expressed with *ptc-Gal4*, trapped GFP–Wingless at the cell surface, while that was not the case for **b)** Human homologues of Dally (Gpc5 and Gpc3). The panels were derived and adapted from ref. [95] and the data obtained and the figures prepared by Dr Ian McGough (Francis Crick Institute, London).

Crystal structure of *Drosophila* glypican Dlp in complex with a palmitoleoylated Wnt peptide

Initially, we analysed the previously reported crystal structure of the apo form of the Dlp^{core} [76] (see also **Figure 1.7 c**) to hunt for any pocket that might be able to accommodate the lipid of Wnt.

Intriguingly, structural comparison has suggested that the N-terminal region of glypicans is distantly related to the extracellular cysteine-rich domain of the Frizzled family of Wnt receptors[129], presenting a conserved set of six disulphide bonds disposed in a similar architecture to the five canonical disulphide bonds of the cysteine-rich domain - as confirmed not only by the apo Dlp^{core} structure[76] but also by the other available glypican structure, human Gpc1[78] (**Figure 3.7 a-f**). The presence of a CRD-like region was very interesting because, as noted in the **Introduction**, the available Wnt-Frizzled crystal structures have shown that this domain can accommodate the Wnt palmitoleate by forming a groove-like binding site.

However, superposition between the glypican and the Frizzled CRDs demonstrates how an analogy is untenable. First of all, the CRD is more compact and the putative binding groove shallower (**Figure 3.7 a-f**). Moreover, the glypican CRDs contain an additional α-helix that is located transversally to the entrance of the groove, thereby blocking lipid insertion (**Figure 3.7 g-h**). This holds for the structures of Dlp and hGPC1.

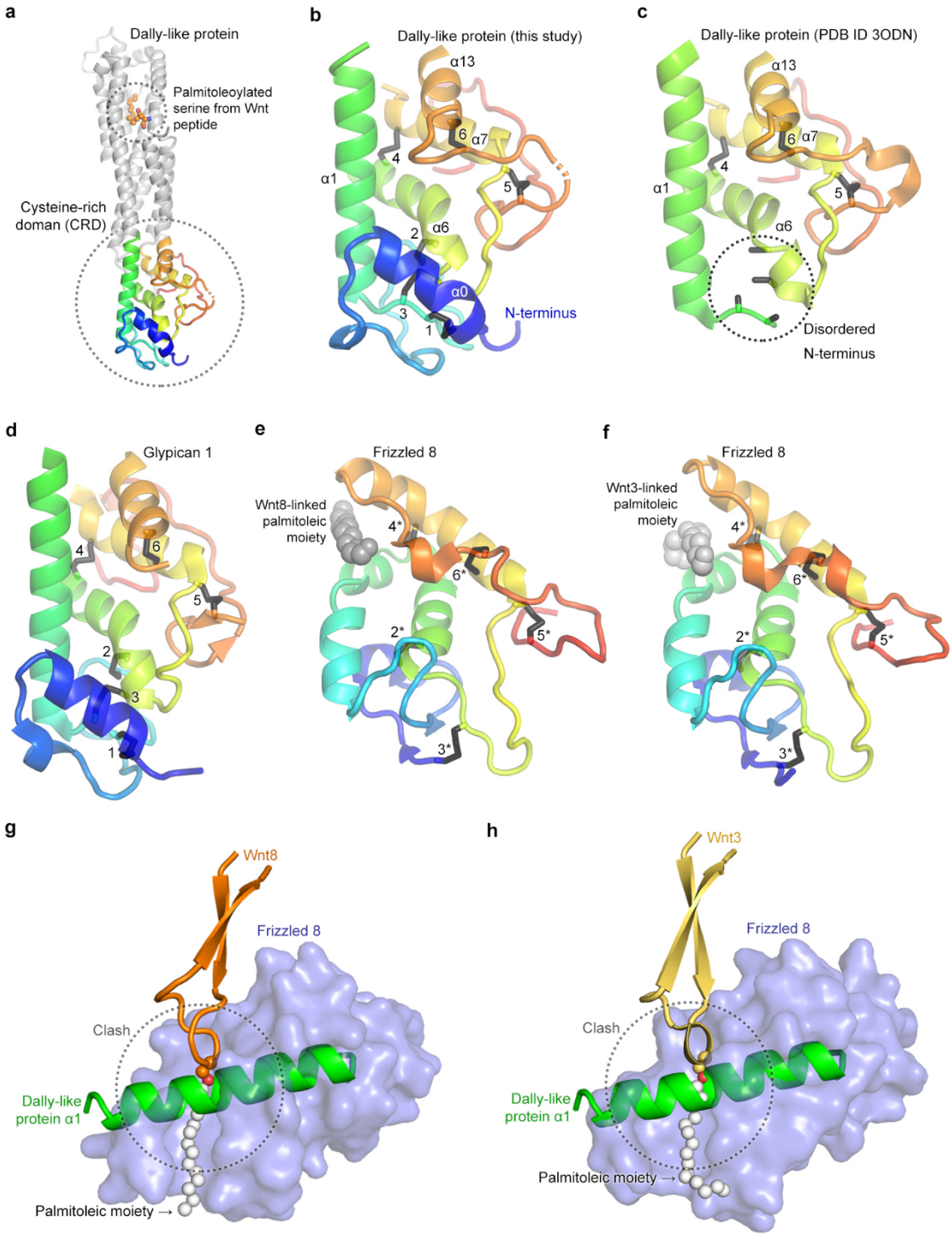


Figure 3.7. The glypican CRD is not the Wnt binding domain.
a-f) Conserved architecture and comparison of the CRDs of: **a,b)** *Drosophila* Dlp^{core} in complex with palmitoleoylated serine of Wnt peptide (PDB ID: 6XTZ). **c)** *Drosophila* apo Dlp^{core} (PDB ID: 3ODN); **d)** human GPC1 (PDB ID: 4YWT); **e)** mouse Frizzled 8 in complex with palmitoleic moiety from Xenopus Wnt8 (PDB ID: 4F0A); **f)** mouse Frizzled 8 in complex with palmitoleic moiety from human

Wnt3 (PDB code 6AHY). Evolutionarily conserved disulphides are shown in black and numbered. **g, h)** Superposition of Dlp^{core} CRD and mouse Frizzled 8 CRD bound to lipidated Wnt8 **(g)** or Wnt3 **(h)**, showing that a conserved helix of glypican CRD sterically hinders binding of the Wnt palmitoleate. The structure of *Drosophila* Dlp^{core} in complex with the palmitoleoylated serine of Wnt peptide **(a, b)**, that will be introduced later in the text, is already shown here to facilitate comparison. The figure was derived from ref. [95].

Therefore, the CRD-like domain of glypicans could not be the site of the elusive Wnt binding pocket we were looking for, and no other cavities could be predicted based on the apo Dlp^{core} structure alone. Thus, we decided to attempt the experimental determination of the Dlp^{core} structure in complex with the Wnt7a peptide already used in the biophysical assays, synthesized in a non-biotinylated form.

Although the protein had the propensity to form crystals after two weeks in a range of PEG-based conditions as precipitant and a varied set of pH, the crystals generally grew as thin plates diffracting to 3 Å resolution at best with a high degree of anisotropy in the c* dimension. After several rounds of optimization, PEG4000 was identified as the most consistent precipitant, and a two dimensional pH/precipitant screening matrix (see Materials and Methods) finally provided an isotropic crystal diffracting to 2.2 Å resolution, in 16.25% (w/v) PEG4000, 0.1 M MMT pH 9, with the protein used at a concentration of 3.5 mg/ml (**Figure 3.8**).

The second hurdle in obtaining the complex structure was to overcome the lipidated peptide's limited solubility, to ensure that enough peptide was available for the complex formation. From equimolar nominal concentrations (probably actually resulting in much lower peptide availability due to insolubility) we

progressively switched to increasing excess of peptide. In our optimized protocol, besides adding the peptide at the moment of the crystallization set-up, we further soaked a solution containing the Wnt7a peptide at a 1 mM concentration on the crystals after their formation, before harvesting them for the data collection (see **Materials and Methods** for details).

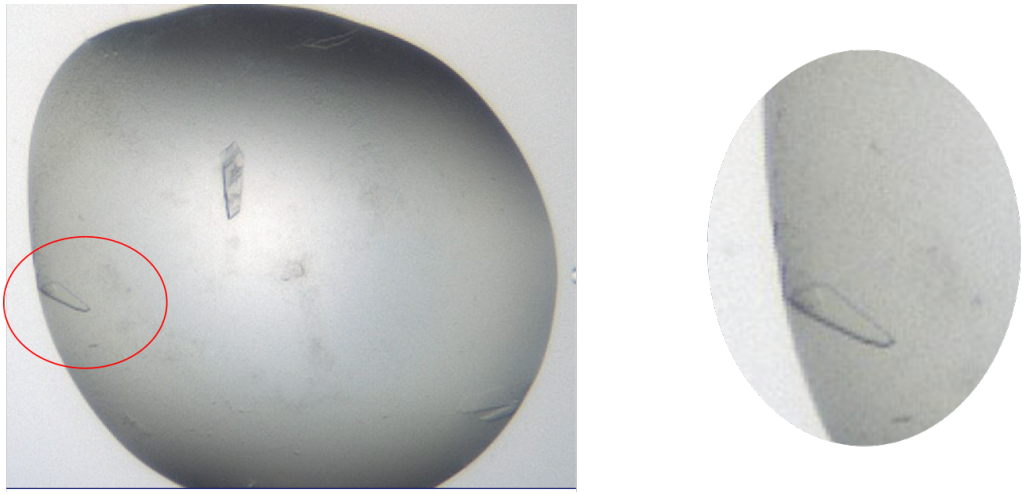


Figure 3.8. Highest diffracting crystal obtained for Dlp^{core} in complex with Wnt7a peptide.

As noted, these efforts finally led to the collection of a 2.2 Å resolution diffraction dataset, with the protein crystallizing in space group $P2_12_12_1$. The structure of Dlp^{core} in complex with the palmitoleoylated peptide (PDB ID: 6XTZ) was solved by molecular replacement, using the apo Dlp^{core} structure previously published[76] as the search model. The structure was then refined by coordinate, TLS and individual B-factor refinement in reciprocal space, alternated with model rebuilding and real-space refinement in COOT (see **Materials and Methods** for further details and **Table 3.1** for the data collection and refinement statistics). The

all-helical elongated cylindrical shape of the uncomplexed Dlp^{core} is conserved, with a global mainchain RMSD of 1.98 Å (for 355 equivalent Cα pairs) (**Figure 3.9 a**). Moreover, some additional portions of the Dlp^{core} N-terminal segment, not reported in the previous structure[76], could be visualised in the complex. These include an α-helix, which we named αH0 for consistency of helix nomenclature with the apo Dlp^{core} previously published (for secondary structure nomenclature see **Figure 3.10**), and a 30-residue unstructured loop, similar to that seen in the published structure of hGPC1 but not in the apo Dlp^{core} (compare **Figure 3.7 b, c,** and **d**). In particular, we could also visualize all the disulphide bonds present in the CRD-like domain of Dlp (one of them was not visible in the apo Dlp^{core} structure previously published), thus confirming experimentally that also the Dlp CRD contains the full set of canonical disulphide-bonds. However, even in this case, as predicted *in silico*, the region corresponding to the palmitoleate-binding groove of Fz CRD remained occluded by the transversal a-helix αH1, and, indeed, no electron density indicative of lipid-binding could be seen in this blocked groove.

On the other hand, crucially, additional electron density, consistent with a palmitoleate acyl chain, could be observed on the opposite side of Dlp^{core}, in the so-called C-lobe. There, a conformational rearrangement takes place in the complex structure: helix αH9 and αH10 tilt away from the opposing helix, αH1, to generate a substantial shift in position compared with the apo structure (~ 6 Å between apo and complex structures at P365). This opens up a deep hydrophobic tunnel, where well-ordered positive electron density could be

observed in the difference map[***]. This revealed the acyl chain, the ester bond with the associated serine sidechain and the serine mainchain of the palmitoleoylated peptide (**Figure 3.9 a** and **b**). The remainder of the peptide lacked a well-defined conformation in the complex structure. This is most likely due to flexibility of this portion as a consequence of the lack of direct interactions with the protein pocket, and is in line with our biophysical data supporting an interaction mainly driven by the lipid moiety of the peptide. The helix movement needed to generate the cavity also explains why the presence of the latter could not be predicted based on the apo structure only.

The main interactions between the lipidated peptide and the tunnel-like binding pocket are hydrophobic, and the cavity is therefore lined with hydrophobic residues contributed by six C-lobe α-helices (L164, L165, F172 of αH1; F191, L194, F195, L198 of αH2; F216 of αH3; L259, I263, F267 of αH5; I359, V363 of aH9 and L366 of αH10, see **Figure 3.9 c**). The sixteen-carbon (C16) acyl chain of the palmitoleate adopts an extended conformation resembling those reported for the Wnt-attached lipid in the FzCRD-Wnt complexes, but markedly different from that observed for lipid binding to Notum, which requires a kink in the palmitoleate (at the C9-C10 *cis* double bond). Notably, the palmitoleate only occupies the narrower, top part of the cavity generated by the conformational

[***] As noted above, before finally obtaining the 2.2 Å diffraction dataset, several other datasets at lower resolution (~ 3 Å) were collected and the structures solved with similar methods. Tantalising evidence of the Wnt peptide was already present in these lower resolution Dlp-Wnt peptide complex structures. To maximize the likelihood of peptide binding and improve the quality and clarity of the difference map in the binding tunnel, I progressively increased the peptide:protein ratio in favour of peptide excess (with additional, subsequent soaking). This provides the rationale for this methodological choice in the optimized protocol, leading to the solution of the 2.2 Å structure presented.

change in complexed Dlp^{core}, leaving substantial space in the bottom part of the cavity (**Figure 3.9 d**).

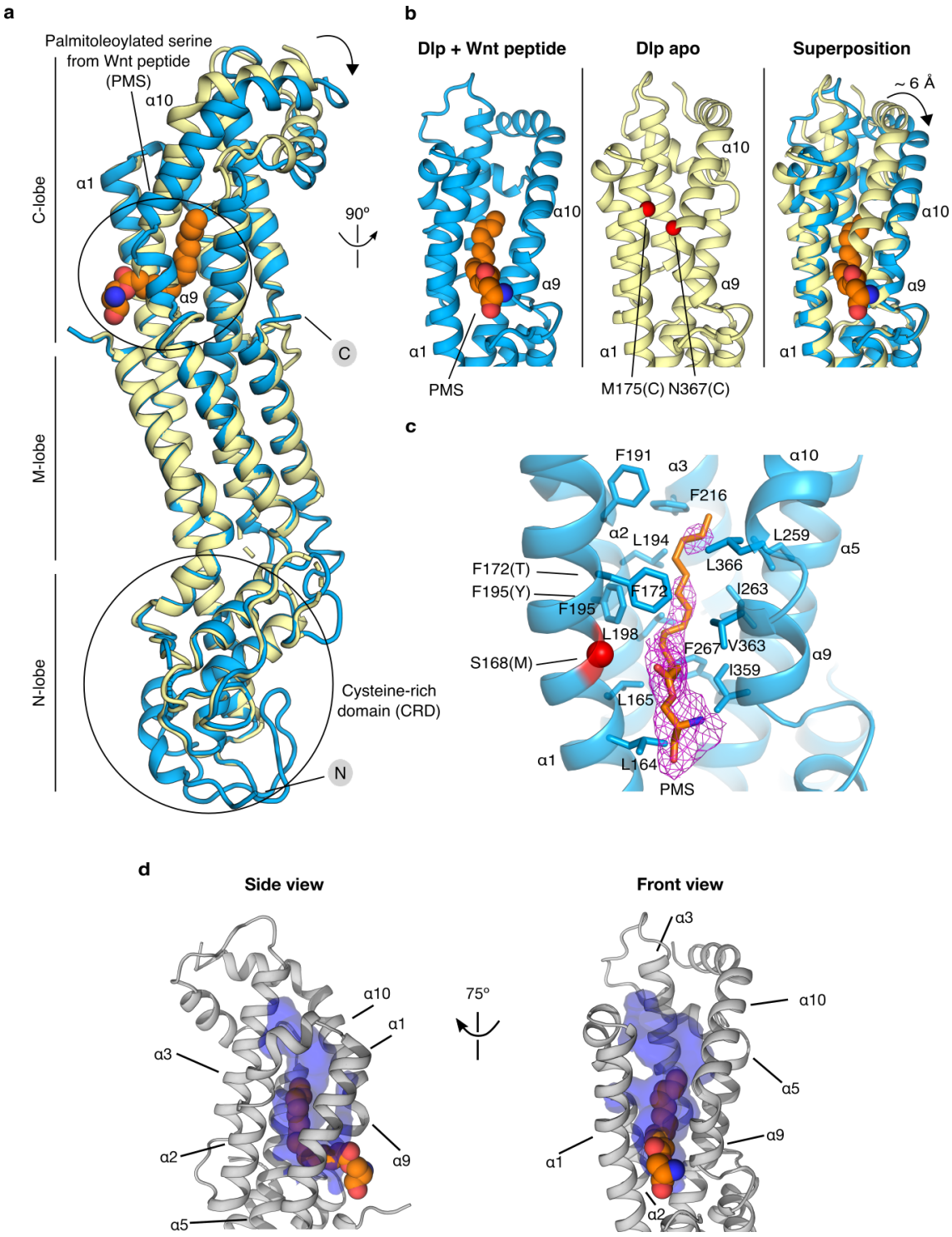


Figure 3.9. Structure of Dlp^{core} in complex with palmitoleoylated Wnt peptide.

a) Superposition between the Dlp^{core} apo structure (PDB ID: 3ODN, pale yellow) and Dlp^{core} in complex with human Wnt7a palmitoleoylated peptide (PDB ID: 6XTZ, marine). The palmitoleoylated serine (PMS) from the Wnt7a peptide is depicted in spheres in atomic colouring (C: orange, N: blue, O: red). N and C

termini are indicated. The black arrow indicates domain movement of the C lobe between apo and palmitoleoylated Wnt7 peptide-bound form. **b)** Close-up view of the Dlp C-lobe, revealing the conformational rearrangement of helices α9 and α10 opening the Wnt-binding pocket. M175 and N367 of Dlp mutated to cysteine residues to form a disulfide bridge to interfere with Wnt binding are indicated by red spheres. **c)** Close-up view of the lipid-binding pocket. Stick representation of residues lining the binding pocket (located within a 4.5 Å range from the Wnt peptide lipid for hydrophobic interactions). The red sphere indicates the site of one of the mutations predicted to interfere with Wnt binding (S168M). The positions of the F172T/F195Y mutations are also labelled. The Fo − Fc omit electron density map for the Wnt peptide palmitoleate is shown in purple, contoured at 2.5σ. **d)** Side and front view of the lipid-binding cavity of Dlp^{core} in complex with the Wnt palmitoleoylated peptide, showing the cavity extension beyond the end of the Wnt peptide acyl chain. The internal volume of the cavity is coloured in blue, with the palmitoleoylated serine from the Wnt peptide represented as spheres in atomic colouring (C: orange, N: blue, O: red). The panels were derived and adapted from ref. [95].

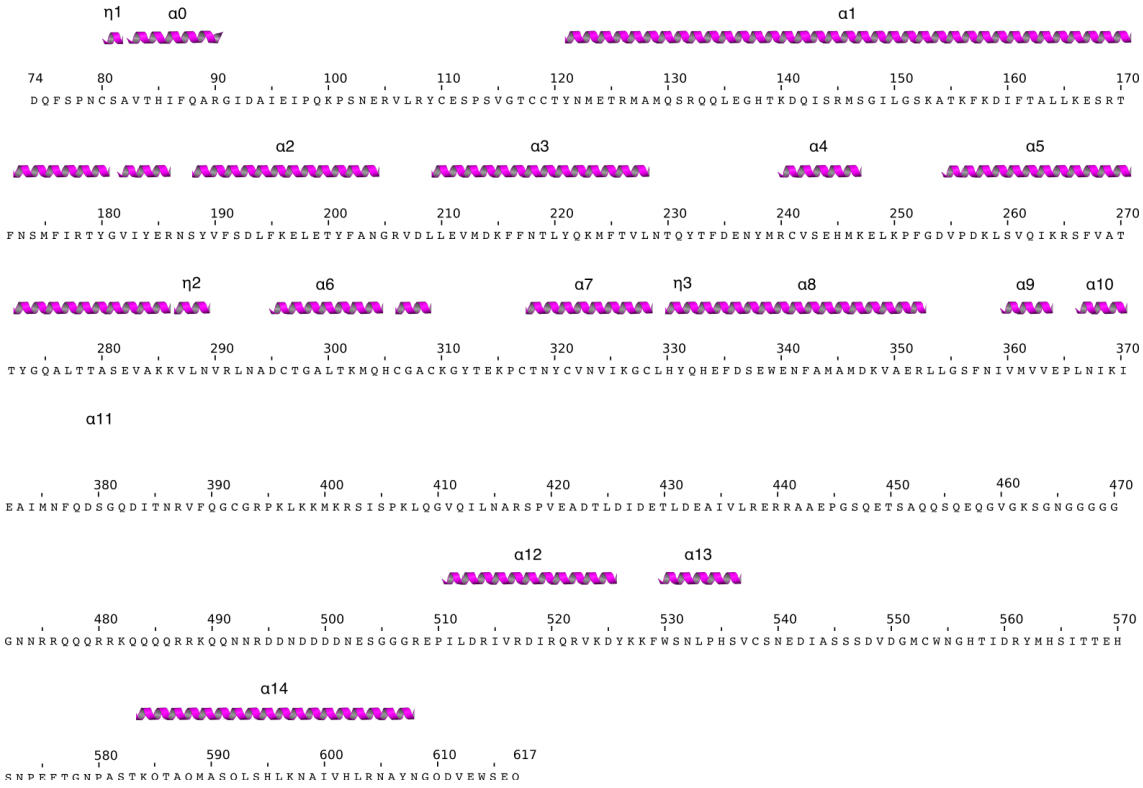


Figure 3.10. Secondary structure nomenclature for the Dlp^{core}-Wnt7a peptide complexed structure.
Sequence of Dlp^{core} construct from the complexed structure annotated with secondary structure elements. To facilitate comparison between the complexed and apo structure, secondary structure nomenclature of the complex reflects that of the previously published apo structure (PDB ID: 3ODN). α indicates α-helices, η indicates 3_{10}-helices. The figure was derived from ref. [95].

To avoid biasing the map, the lipidated peptide model (or, better, its visible portion described above) was added at the last rounds of refinement. Moreover, to rule out refinement artefacts, several different types of omit map were generated (**Figure 3.11**), including a simulated-annealing omit map and a feature-enhanced Polder map[130], all providing evidence for lipid-like electron density in the binding pocket. The rationale for fitting the electron density with the palmitoleoylated serine from the Wnt peptide instead of, for instance, chemicals present in the crystallization cocktail with similar shape, like PEG, came from multiple considerations: although crystallized in a different space group, PEG was also present in the crystallization cocktail of the published apo form and no conformational rearrangement and no lipid-like density is present in the latter; to fit the electron density, a fragment of PEG of the exact same dimensions as the palmitoleate would be required; the highly hydrophobic binding cavity is less compatible with a polar molecule like PEG than with a lipid; above all, all the omit maps showed, to various extents, evidence for continuous electron density not only for the lipidated serine, but also for additional portions of the peptidic part of the peptide. The Polder map is particularly clear in this sense (**Figure 3.11 d**). However, because even this map was not informative enough to undoubtedly assign the identity and fit the residues surrounding the serine on either side of the serine itself, we took a conservative approach and we decided to build the palmitoleoylated serine only (since the peptide is circular, the directionality for residue modelling cannot be inferred by sequence analysis alone).

An analysis of the crystal contacts was also performed, to evaluate if the crystal packing could have influenced the conformational rearrangement. Although one contact could be found in the C-lobe region, the thermodynamic-based evaluation of the interface conducted using the PISA web server (see **Materials and Methods**) showed how this is a minor interface (554 $Å^2$ out of 22260 $Å^2$ of accessible surface area), with a CSS (Complex Significance Score) of 0, indicating that it is an energetically very weak interface and therefore not expected to trigger a significant conformational rearrangement in the structure.

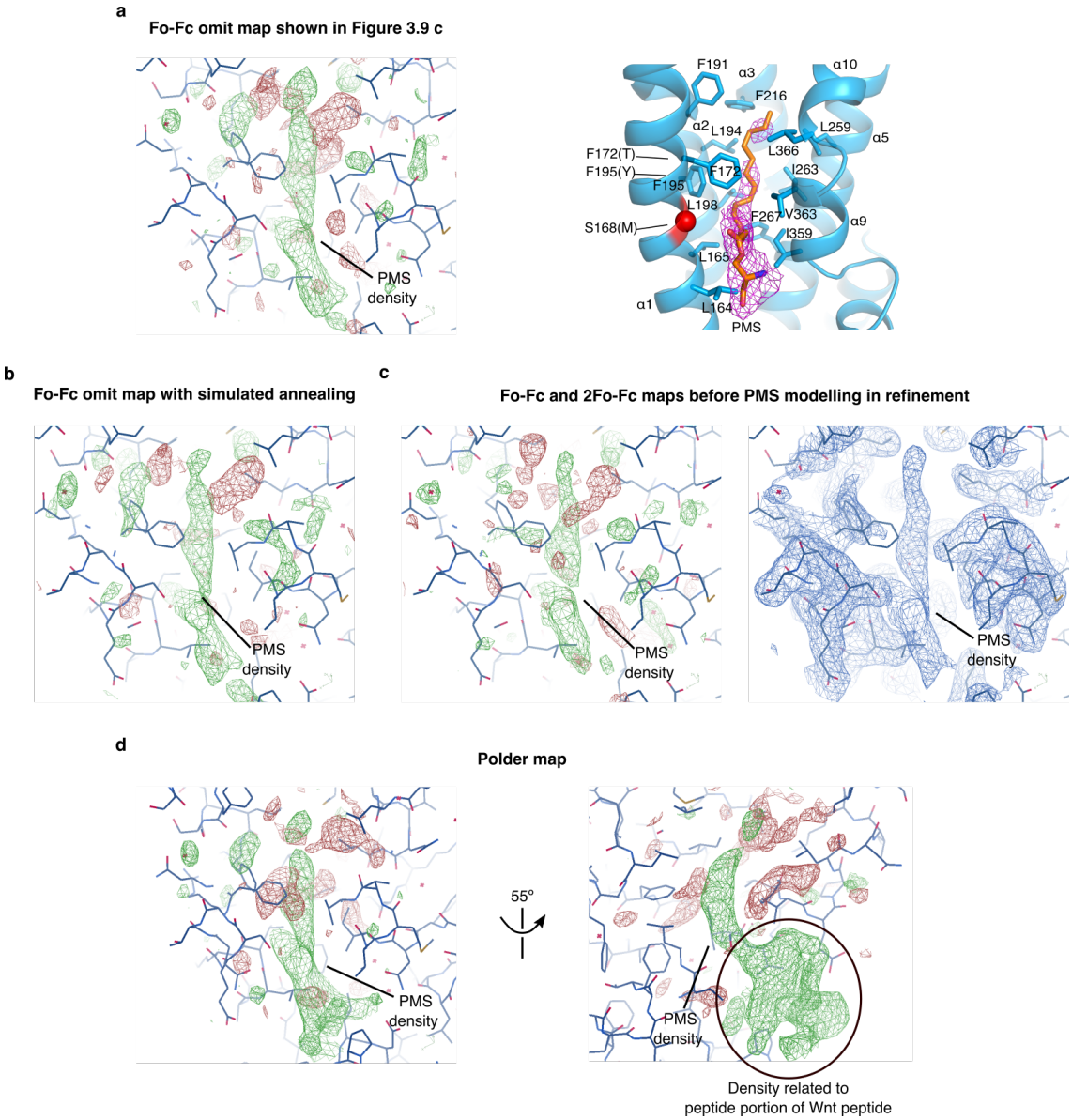


Figure 3.11. Several types of omit maps showed and supported the modelling of the electron density in the binding pocket as PMS from the palmitoleoylated Wnt peptide.
a) Coot-displayed Fo − Fc omit map (left) generated from the refined model after removal of PMS, as shown in **Figure 3.9 c**. To help orientation and comparison, **Figure 3.9 c** is duplicated here, displaying the same map in Pymol (right). The maps are contoured at the same level in the two programmes (±2.5σ). **b)** Coot-displayed Fo − Fc omit map (contoured at ±2.5σ) generated from the refined model, after removal of PMS, application of simulated annealing and one round of coordinate and B-factor refinement. **c)** Coot-displayed Fo − Fc map (left) and 2Fo − Fc map (right) of the last refinement iteration before modelling PMS in the electron density. As PMS was never modelled, these maps are unbiased. The Fo − Fc map and 2Fo − Fc map are contoured at ±2.5σ and 1σ, respectively. **d)** Polder map generated from the refined model using phenix.polder[130]. This omit

map, whose characteristic is to enhance weak electron density features in the omit region, shows additional electron density that can be assigned to the peptide portion of the palmitoleoylated peptide, therefore further supporting PMS modelling into the electron density. The map is contoured at ± 2.5σ. The figure was derived from ref. [95].

Table 3.1. Data collection and refinement statistics.

	Dlp-Wnt7a-peptide complex
PDB ID code	6XTZ
Ligand code	O18
Data collection	
X-ray source (Diamond)	I04
Wavelength (Å)	0.97954
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	63.7, 77.4, 118.8
α, β, γ (°)	90, 90, 90
Resolution (Å)	56.14-2.21 (2.25-2.21)
R_{merge}	0.176 (2.026)
CC 1/2	0.9983 (0.6847)
$I / \sigma I$	8 (1.1)
Completeness (%)	100 (100)
Redundancy	13.1 (11.2)
Refinement	
Resolution (Å)	56.14-2.21
No. reflections	30027 (2687)
R_{work} / R_{free}	0.2267/0.2553
No. atoms	
Protein	3395
Ligand/other	23/46
Water	82
B-factors (Å^2)	
Protein	62.93
Ligand	78.40
Water	53.29
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.38

Values in parentheses are for highest-resolution shell.

From structure to function: design of mutations to impair Wnt binding to Dlp^{core}

The finding of a tunnel-like cavity able to accommodate the lipid portion of the peptide provided a strong molecular explanation for the ability of Dlp^{core} to directly bind Wnt. However, we then wanted to rule out the possibility that this was due to the specific experimental conditions of the crystallization experiment and validate our findings in solution and *in vivo*.

To the aim, we took two approaches. First, we aimed to hamper lipid insertion into the cavity by introducing a bulky residue at the entrance; we selected a serine, Ser168, located at the mouth of the tunnel, but not involved in intermolecular contacts based on our structure, and we mutated it to a methionine (mutation S168M). Alternatively, we engineered a disulphide bridge to prevent αH10 and αH1 from coming apart and, hence, to lock Dlp in the closed, apo conformation: in this sense, we selected Met175 on αH1 and Asn367 on αH10 (mutation M175C/N367C). The two mutation sites are highlighted in **Figure 3.9 b** and **c**.

We initially tested the mutations biochemically, by designing an assay able to assess and quantify the potential reduction in the Dlp^{core} lipid-binding ability and, at the same time, to exclude confounding effects introduced by the mutations themselves (e.g artefacts due to protein misfolding or mutation-driven oligomerization). To the goal, we used Fluorescence-detection Size Exclusion Chromatography (FSEC), sampling the different purified mutants for their binding to a fluorescently-labelled lipid, namely bodipy 505/512-labelled palmitic acid

(C16:0). The lipid was chosen as the best commercially-available mimic of the Wnt lipid, and the intensity and co-elution of its fluorescent signal with the protein 280 nm absorbance peak were evaluated.

As expected, the binding to bodipy-palmitate was substantially reduced for both mutant proteins with respect to wild-type Dlp^{core} (**Figure 3.12 a**). The fact that the mutants showed an elution profile similar to the wild-type protein in the 280 nm traces also indicated that the mutations were not impacting on protein expression, folding or multimerization propensity. To further exclude any oligomerization issue, a Western Blot in reducing and non-reducing conditions was also performed for the M175C/N367C mutant, and the absence of any shift for the Dlp^{core} band between the two experimental set-ups confirmed that the disulphide bridge was formed intramolecularly as envisaged, and was not causing oligomerization (**Figure 3.12 b**).

Next, we asked our collaborators to confirm the data *in vivo*. Both $Dlp[S168M]^{\Delta GPI}$ and $Dlp[M175C/N367C]^{\Delta GPI}$ were less efficient than wild-type $Dlp^{\Delta GPI}$ at solubilising GFP-Wingless in conditioned medium (**Figure 3.12 c**). Moreover, when tested in trapping experiments, in particular the M175C/N367C mutant completely abrogated *in vivo* trapping (**Figure 3.12 d**).

These combined observations confirmed the *bona fide* identification of the Dlp^{core} C-lobe cavity as the Wnt binding module in Dlp, and suggested that the movement of αH10 relative to αH1, as observed in the crystal structure, is essential for tunnel formation and lipid-binding.

We also aimed at identifying the molecular determinants of the differential activity between the Dlp and Dally family of glypicans. As can be observed in the

sequence alignments in **Figure 3.13 a**, the hydrophobic chemical character of the residues interacting with the peptide lipid in the Dlp family is broadly conserved also in the Dally family, apart from two main differences: a hydrophobic-to-polar substitution at residue F172 and F195 of Dlp (a threonine and a tyrosine, respectively, in Dally), and a putative glycosylation site introduced by the F to T change at position 172 of Dlp, which generate the consensus N-glycosylation sequence Asn–X–Ser/Thr in Dally.

Therefore, we generated two further Dlp mutants (F172T/F195Y, see also **Figure 3.9 c** mapping the mutations onto the pocket**,** and R169E/T170N/Q171M/F172T, the latter introducing the putative Dally glycosylation site in Dlp), and we also engineered a Dally mutant (E149R/N150T/M151Q/T152F), to check if removing the Dally glycosylation site was enough to provide Dally with lipid-binding ability. We directly performed the test *in vivo*, as this was the set-up closest to the physiological situation. The two Dlp mutants reduced trapping but not completely, while the Dally mutation did not confer lipid-binding activity to Dally (**Figure 3.13 b**). These results thus suggested that the difference in Wnt-binding activity between Dlp and Dally cannot be solely ascribed to residues at the entrance or within the tunnel.

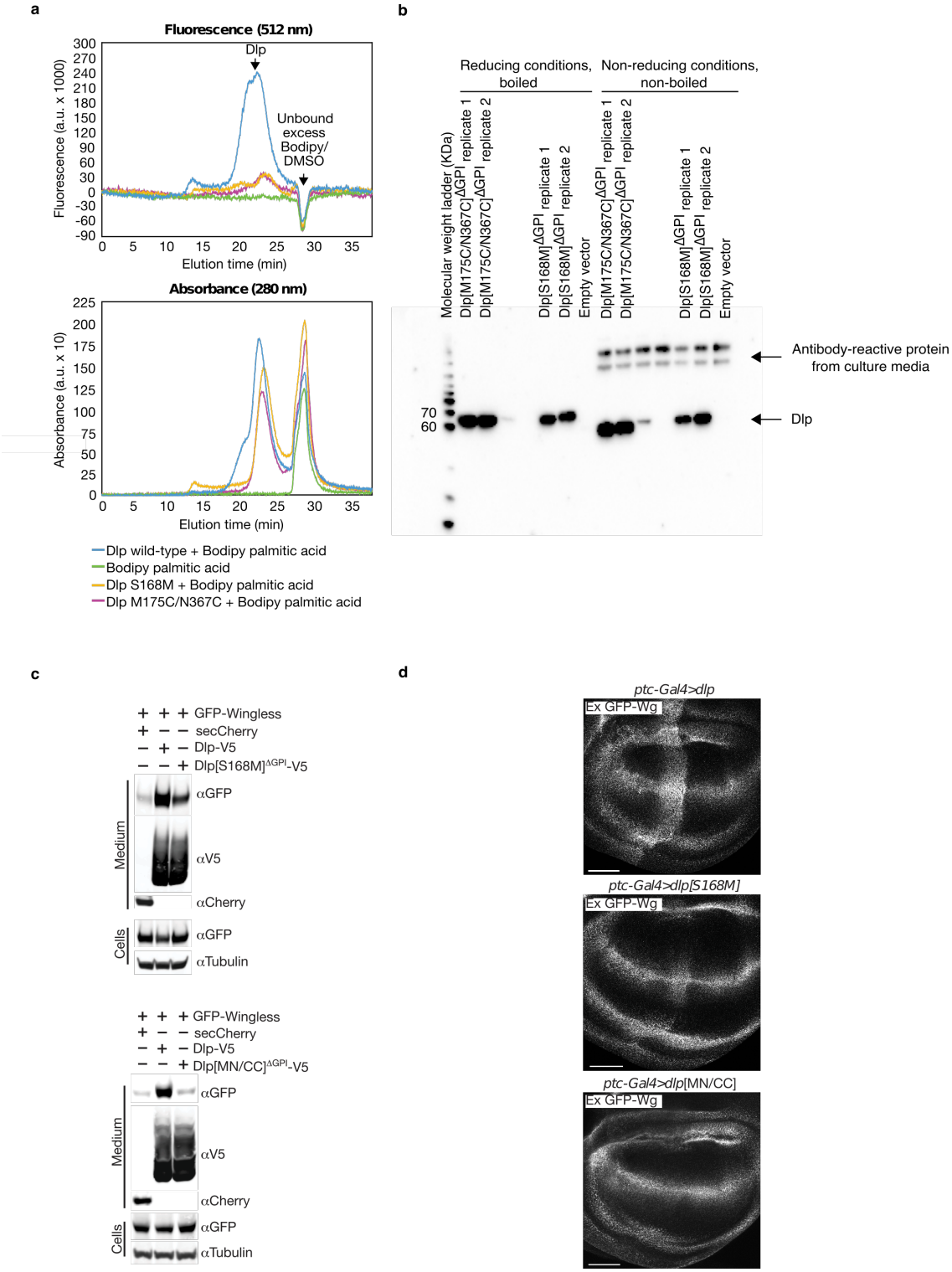


Figure 3.12. Structure-guided mutations impaired Dlp lipid-binding ability. **a)** Representative FSEC profiles of purified wild-type Dlp^{core} (blue) and the S168M (yellow), M175C/N367C (purple) mutants incubated with bodipy-labelled palmitic acid (C16:0), showing a reduction of fluorescence emission of mutants with respect to wild-type. This is suggestive of diminished binding affinity. The

fluorophore profile in absence of proteins (green) is presented as control. The experiments were repeated independently three times with similar results**. b)** Western Blot on supernatants from small-scale transfections of HEK293S GnTI- with different Dlp^{core} mutants. The absence of a shift for Dlp[M175C/N367C]$^{\Delta GPI}$ between reducing and non-reducing conditions suggests that the engineered disulphide bond is formed intramolecularly as envisaged. The additional antibody-reactive band in non-reducing conditions, visible also in the cells transfected with the empty vector, derives from a contaminant protein present in the culture media. The unlabelled lanes correspond to other mutants tested but not pursued further. **c)** Solubilization assays of Wingless from serum-free media conditioned by S2 Wingless-expressing cells. Dlp(S168M)$^{\Delta GPI}$ presents a reduced recovery, whereas Dlp(M175C/N367C)$^{\Delta GPI}$ has virtually no solubilization activity. The experiments were repeated independently three times with similar results. **d)** Trapping assays in wing discs expressing Dlp–HA, Dlp(S168M)–HA and Dlp(M175C/N367C)–HA in a stripe transversal to the Wingless expression, showing reduced trapping for Dlp(S168M) and no trapping for Dlp(M175C/N367C). The figure was derived from ref. [95]. For panels c and d, the data were obtained and the figures prepared by Dr Ian McGough (Francis Crick Institute, London).

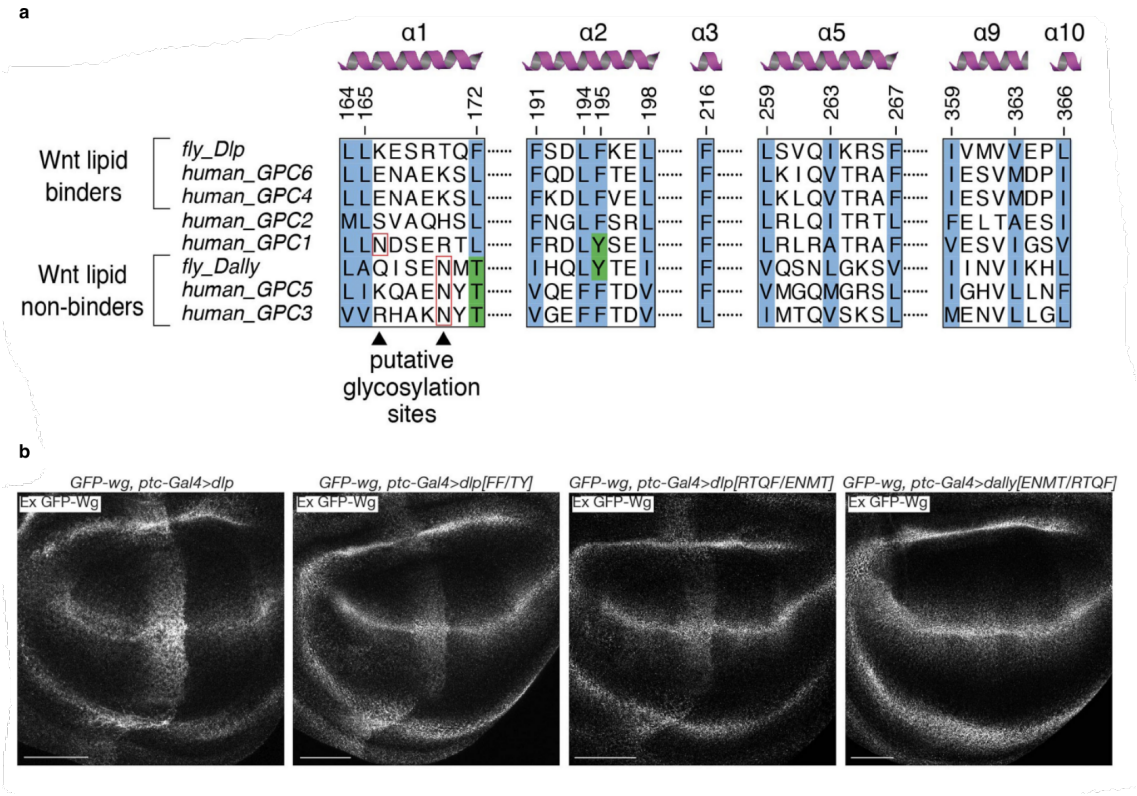


Figure 3.13. Mutational analysis showed that the difference between Dlp and Dally is not fully attributable to the cavity environment.
a) Sequence alignment of the Dlp lipid-interacting residues with the correspondent residues of *Drosophila* Dally and human glypicans, coloured according to the Clustal X scheme (blue, hydrophobics; green, polar). The

alignments were generated JALVIEW (v. 2.10.5)[131]. **b)** Trapping assays in wing discs expressing Dlp (F172T/F195Y), Dlp (R169E/T170N/Q171M/F172T) and Dally (E149R/N150T/M151Q/T152F; ENMT/RTQF, removing the Dally putative glycosylation site and replacing it with the correspondent Dlp sequence) in a stripe transversal to the Wingless expression, showing reduction but no complete loss of trapping for Dlp (F172T/F195Y), Dlp (R169E/T170N/Q171M/F172T), and no gain in lipid binding activity for Dally (E149R/N150T/M151Q/T152F). The panels were derived from ref. [95]. For panel b the data were obtained and the figure prepared by Dr Ian McGough (Francis Crick Institute, London).

Discussion

What we have learnt

Inherent to the definition of a protein as morphogen is its ability to form gradients that, in a concentration-dependent manner, induce direct or indirect specific cellular responses at a distance from the source. Accordingly, Wnts have been demonstrated to act over several cell diameters from the producing cells. But how can Wnt spread in the extracellular space and reach distant targets if its evolutionary conserved acylation with a palmitoleic acid reduces its solubility?

In the work presented here, in collaboration with the Vincent lab for the functional side of our studies, we first evaluated in the *Drosophila* wing imaginal disc the main models proposed in the literature to overcome this “Wnt solubility problem”, finding no supporting experimental evidence for them.

We then looked for alternative Wnt interactors in the same tissue and found that members of the glypican family of heparan sulphate proteoglycans appeared to bind the morphogen. In particular, the interaction seemed driven not only by the glypican long sugar chains, but also, more specifically, by their protein core.

Using purified core proteins of *Drosophila* glypicans and Wnt-mimicking palmitoleoylated peptides, we first confirmed biophysically by BLI, that a direct binding was indeed happening, with an estimated affinity in the sub-micromolar range. Moreover, the interaction was specific for a subset of glypicans (*Drosophila* Dlp and the related human homologues Gpc4 and Gpc6, but not *Drosophila* Dally and human Gpc3 and Gpc5) and the Wnt lipidation was crucial to the binding, since the same interacting glypicans assessed against a non-palmitoleoylated version of the Wnt peptide resulted in no binding signal.

Solving the structure of Dlp^{core} in complex with the palmitoleoylated Wnt peptide at 2.2 Å, we could explain that the lipid-binding activity of Dlp derives by a conformational rearrangement taking place in the C-lobe of Dlp^{core}, creating a deep hydrophobic tunnel to accommodate the palmitoleate.

Mutagenesis analysis - aimed for instance at preventing the conformational rearrangement with an artificial disulphide bond locking the structure in the closed conformation - validated the structural finding and confirmed the key role of this tunnel in accommodating the Wnt lipid in solution and *in vivo* as well.

This allowed us to suggest a model where the Dlp family of glypicans can spatially and temporarily regulate the activation of Wnt signalling thanks to the lipid-binding ability, and fulfil the role of the elusive Wnt partner that allows Wnt to signal at a distance.

Glypicans in Wnt morphogen diffusion

Shielding the lipid of Wnt, Dlp and the related glypicans can, indeed, act as a reservoir of signalling-competent morphogen, that can be readily accessed and

handed over to other Wnt receptors upon formation of suitable local concentrations of the latter. The relative abundance of and affinities/avidities of Dlp vs. other receptors likely decides the binding outcome (**Figure 3.14**). For example, in the case of Frizzled, when this key Wnt receptor is present at the cell surface in higher quantities than Dlp, the morphogen would preferentially interact with Frizzled, due to its higher affinity for Wnt; in this circumstance, Dlp would act as a co-receptor. In the case of higher concentration of Dlp vs Frizzled, instead, the glypican would keep the Wnt lipid in its pocket, delaying the activation of the pathway until suitable conditions are generated and the balance is shifted. This is in line with previous observations suggesting that Wnts transfer from Dlp-class glypicans to Frizzled receptors to initiate signalosome assembly[71].

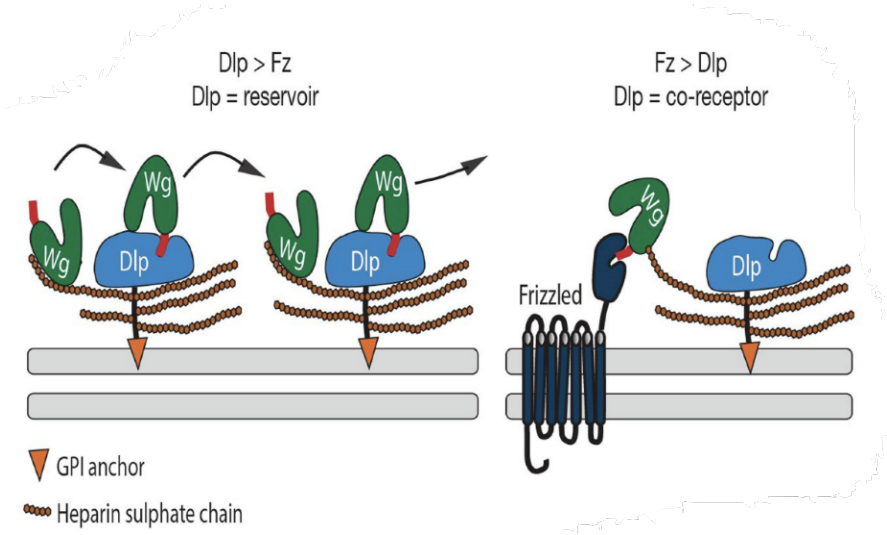


Figure 3.14. Model of Dlp functional interaction with Frizzled receptors. The figure was derived from ref. [95].

Remarkably, this mechanism can also allow a tighter control on the signalling cascade, (requiring Frizzled/other Wnt receptors to be present in suitable

quantities for the handover, and therefore preventing inadvertent pathway activation) and provide the cell with an additional layer of regulation, varying the expression levels of the Dlp class glypicans[132].

The proposed mechanism also fits well with the hypothesis that Wnt is not freely released in the extracellular space[49], and can reach distant targets thanks to cell migration, donor cell-target cell contacts and/or cell protrusions like cytonemes[64]. Indeed, cytonemes are coated with glypicans[133] and, in chick embryos, neural crest cells have been shown to carry and deliver Wnts at a distance in a Gpc4-dependent manner[134]. We were not able to directly test the cytonemes hypothesis in our studies due to the lack of suitable experimental genetics tools in *Drosophila*.

Despite not being the only lipid-binding protein in the extracellular space (as noted in the **Introduction**), glypicans of the Dlp class are particularly well suited to play a key role in Wnt transport, particularly along epithelia. First, their GPI anchor might facilitate rapid diffusion within the cell membrane while ensuring retention in the plane of the epithelium. Moreover, their dual mechanism of Wnt binding, by the sugar chains and by our newly discovered hydrophobic cavity, might facilitate the handover mechanism described above. The long and flexible HS chains (extending some 40-130 nm compared to the 11 nm axis of the cylinder-like protein core[78]) might in fact allow the donor Dlp to hold onto Wnt while the palmitoleate finds a suitable tunnel on a receiving Dlp molecule.

On the other hand, the Dally class of glypicans, only retaining the HS binding site but not the lipid interaction, is likely to act as an accessory, non-essential Wnt co-receptor.

It is in fact worth stressing once more that both classes of glypicans are able to act as co-receptors, directly binding by their HS chain a so far unidentified polar patch on the Wnt protein surface. In addition to that, the further binding site for the palmitoleate identified here confers the Dlp class of glypicans the additional lipid-shielding and, hence, proposed reservoir capacity discussed above.

In line with this, Dally is not essential for Wingless activity in *Drosophila*, but the phenotype of Dally and Dlp double mutants is stronger than that of the single Dlp mutant, suggesting a partial redundancy[71,75,135,136]; moreover, overexpressed Dlp$^{\Delta GAG}$ extends the Wingless gradient but less than full-length Dlp[75]. So, the heparan sulphate chains boost the overall affinity by providing a secondary site of interaction; it is however difficult to tell which site is engaged first during complex formation.

Open questions

Several questions remain to be answered. For example, the presence of the large, additional space in the binding tunnel unoccupied by the palmitoleic acid is intriguing. What is its function?

A possible explanation relates to a certain level of tunnel promiscuity in lipid binding that was suggested by some of our experiments. In fact, as noted in the structure description, the palmitoleate of the Wnt peptide sits in the binding pocket in an elongated shape; this shape can be assumed both by saturated or unsaturated fatty acids. In other words, the unsaturation is not a prerequisite to fit the cavity volume, as in the case of the shallow lipid-binding pocket of Notum,

for example – where this mechanism contributes to the specificity of recognition. It follows that even the saturated counterpart of palmitoleate, palmitic acid (C16:0), should be able to fit the tunnel. Indeed, this was confirmed by my FSEC experiments on the Dlp mutants (**Figure 3.12 a**), which used as probe a fluorescently-labelled saturated palmitic acid.

Therefore, the Dlp class of glypicans can represent a broader hub for morphogen trafficking, helping the solubilisation of further lipidated proteins, like Hedgehog, which presents a N-terminal palmitate and a C-terminal cholesterol.

Indeed, the original apo structure of Dlp was obtained in the context of studies aimed at clarifying the role of this glypican in Hedhehog signalling, following former reports of its requirement to ensure normal Hh pathway activation. These studies fell short of demonstrating a direct interaction between Dlp and Hedhehog, however the Hedgehog protein used in the paper was without the canonical lipid modifications, and the findings therefore do not exclude a lipid-based interaction.

It is tempting to reason that the additional space in the Dlp binding cavity observed in our structure might allow for the two Hh lipid moieties to bind Dlp – and the width of the cavity there compared to the top part of the tunnel makes the space particularly suitable for cholesterol.

However - as will be further dissected in the final **Discussion** chapter - the dual Hh lipidation, and the complex effects it has on the mechanisms of pathway activation, hamper a direct comparison with the findings obtained for Wnt, and make every extrapolation from the latter to interpret data on the Hh side too speculative. I limit myself to suggest that our structural data do not rule out, and

might indeed support, the possibility for a lipid-based, direct interaction between Dlp (and glypicans of the same sub-family) and Hedgehog.

The other important point, that will need further structural analysis in the future, relates to the drivers of the conformational rearrangement observed in the complex structure, and the determinants of the differential lipid-binding ability between the Dlp class and the Dally class of glypicans. In this sense, my work has demonstrated that, although differences at the mouth of the cavity (e.g. glycosylation) or inside the cavity (hydrophobic-to-polar substitutions) have an effect, they do not wholly explain the complete lack of lipid interaction of Dally^{core} compared to Dlp^{core}. The identification and characterization of further binding partners (for example, analysis of potential interactions with Frizzled receptors) could help to clarify this matter. It is for example worth highlighting that the helix αH1 transverses the entire Dlp structure from the N-lobe to the C-lobe, and this could facilitate the transmission of even distant conformational alterations to the cavity; some of these, generated by sequence features differentially present in Dlp but not in Dally, could then allow the shift of αH9 and αH10 relative to αH1. Relatedly, it would be interesting to assess the structural and functional role of the CRD in the N-lobe: despite this domain not being able to bind lipids due to occlusion of its cavity as discussed in the presentation of results, it might still play a role in Dlp homo-oligomerization or as an interaction surface with other proteins/CRDs.

Chapter 4 - Wnt intracellular transport and secretion: biochemical studies on the carrier Wntless

As noted in the **Introduction**, Wntless is the dedicated carrier receiving Wnt from Porcupine upon its palmitoleoylation in the endoplasmic reticulum, and escorting the morphogen in all its intracellular travel through the Golgi apparatus up to the plasma membrane of the producing cell. The Wnt-binding unit in Wntless was reported by a previous study to be in the first extracellular/intraluminal loop[†††], and this region was predicted to have a lipocalin-like fold. Moreover, the loop was shown to be capable of Wnt-binding even when recombinantly expressed in absence of the rest of the Wntless regions (although this finding relied on immunoprecipitation experiments from cell lysates rather than on a biophysical analysis with purified proteins).

Our attention thus initially focused on this loop, in order to solve its structure. potentially in complex with Wnt or with a Wnt surrogate as used for the glypican study. However, apart from the indicated report, minimal biochemical characterization of Wntless was present in the literature, and its sequence does not present any, even remote, relationship to a structurally characterised protein. Therefore, the project proved challenging from the very early stages of identification of the correct domain boundaries, and significant effort was required to even achieve protein expression.

[†††] For clarity, the term “loop” in the context of this chapter refers to the fact that the region is located between two transmembrane regions of an integral membrane protein, and does not indicate absence of secondary structure.

After a thorough screening, however, first an expressing construct was identified, and successively the yield and the quality of the purified protein underwent significant improvements. However, most of the recombinantly expressed Wls loop kept behaving unsatisfactorily in gel filtration, forming a series of oligomers unsuitable for crystallization. Some monomeric/dimeric protein could also be isolated, but failed to yield crystals.

This led me to question the definition of this Wnt-binding loop as a unit capable of independent folding and prompted me to switch the focus to the full-length transmembrane protein. Even in this case, however, despite significant efforts, the protein failed to show a satisfying behaviour, particularly in the expression scaling-up process.

Recently, however, a preliminary test of Wntless incorporation in detergent-free systems (SMALPs) has shown how this strategy could finally pave the way for successful expression, purification and downstream analysis of the protein.

The next paragraphs will therefore present the most relevant experimental turning points that allowed to continuously improve the expression and purification behaviour of Wntless, despite the ultimate lack of success in getting structural information. The breadth of the experimental conditions analysed and the trials performed can be appreciated by reference to the relevant **Materials and Methods** sections, appropriately recalled in the text.

The Wnt-binding 1st extracellular loop of Wntless produced as soluble recombinant protein: expression and purification attempts

The experimental work performed to increase the expression and the purity of the 1st extracellular loop of Wntless followed several steps, involving the use of different constructs and expression systems.

Initially, I focused on the identification of the correct domain boundaries of the Wnt-binding domain of human Wntless. As noted above, the very limited sequence homology of the Wnt carrier with other proteins in the protein databases hampered the use of three-dimensional homology modelling to assist with construct design: indeed, bioinformatic homology detection performed with web-based tools such as HHPred[110] failed to identify homologues other than members of the Wntless family itself. *De novo* structure prediction was also attempted (using for instance the web server Robetta[116]), but the quality of the output models, as assessed by the confidence score provided by the modelling program, was poor and unreliable for experimental design.

Therefore, construct design was mainly based on multiple sequence alignments between family orthologues and amino acid conservation analysis, helped by secondary structure predictions, disorder predictions, signal sequence prediction and selection of terminal residues most likely to favour construct solubility and crystallization according to published investigations on the topic[137] (see **Materials and Methods** for a list of the servers used). As shown in **Figure 4.1**, the 1st extracellular loop of Wntless contains four conserved cysteines, and is predicted to have a mainly β-sheet-based secondary structure.

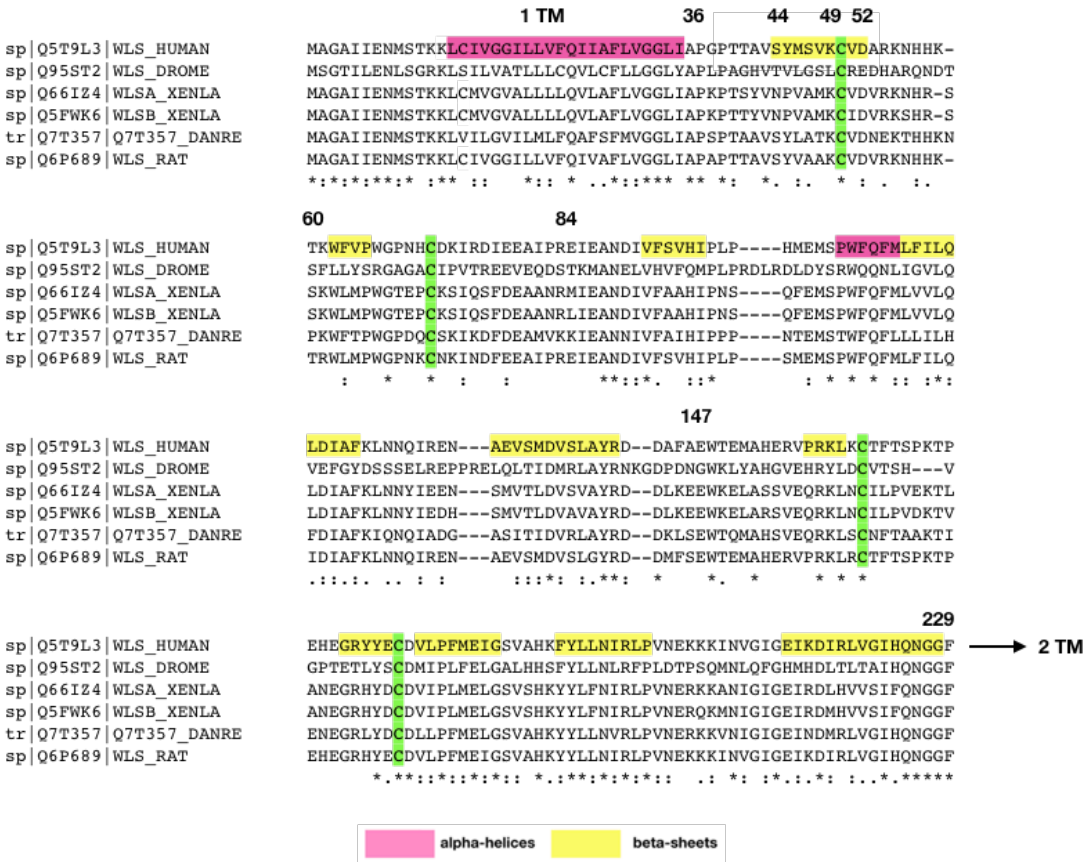


Figure 4.1. Multiple sequence alignment of the 1st extracellular loop of Wntless family members from different species.
The numbers indicate relevant residues used as construct boundaries in construct design, and refer to the numbering of the human sequence. Below the sequence, according to Clustal Omega scheme, the symbol * indicates conservation of residues, the symbol : conservation of strongly similar biochemical properties, the symbol . conservation of weakly similar biochemical properties. 1 TM = 1st transmembrane region; 2 TM = 2nd transmembrane region. Human = *Homo sapiens*; DROME = *Drosophila melanogaster*; XENLA = *Xenopus laevis*; DANRE = *Danio Rerio*; RAT = *Rattus norvegicus*.

14 constructs of the human loop with different boundaries were therefore designed around this conserved cysteine core (see **Table 2.3** in **Materials and Methods**) and initially tested in transiently transfected HEK293T cells, but they presented either very little or no secretion; Western Blot analysis of the pellets demonstrated cell retention of the protein, probably indicating misfolding (**Figure**

4.2). To assess if co-expression with the ligand could promote expression, the constructs were also transfected in commercially sourced Wnt3a-producing mouse fibroblast (L-cells), but this didn't improve the outcome.

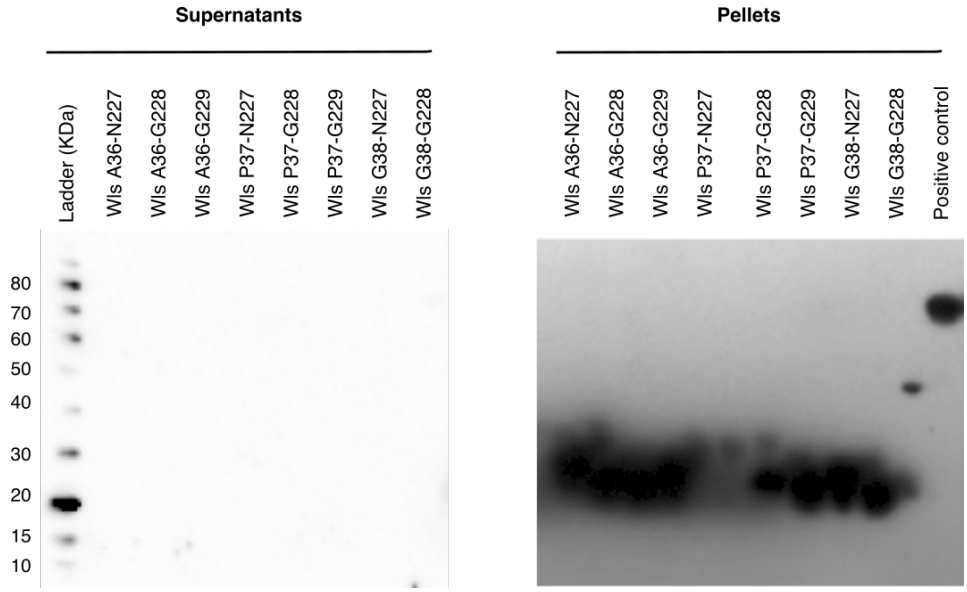


Figure 4.2. Human Wls 1st extracellular loop constructs were retained in the cell.
Representative Western Blot of some of the Wls loop constructs screened, tested upon small-scale transient transfection of HEK293T mammalian cells. The constructs were designed to be produced as soluble proteins, secreted in the culture media. The proteins were probed with a His-reactive anti-PentaHistidine primary antibody (Qiagen). The positive control is represented by mouse Semaphorin 6a (aa. 19-649, UNIPROT ID: O35464), provided by Dr Gergely Nagy (Division of Structural Biology, University of Oxford).

I hence decided to switch to a different expression system (*Drosophila* Schneider's S2 cells). The reasons for selection of S2 cells as expression system were multiple: the latter has already proven useful to express Wnt pathway proteins[27]; additionally, it is the same system used by our collaborators at the Crick, the Vincent group, which in addition to their analyses of glypicans also performed genetic studies on Wntless. To match even further their experimental

setup, I also used a plasmid encoding the *Drosophila* loop rather than the human. Therefore, I generated a stable S2 cell line constitutively expressing the complete *Drosophila* 1st extracellular loop (aa. 36-236); the cells were also co-transfected with a construct encoding *Xenopus laevis* Wnt8 under copper sulphate-controlled inducible expression. This was done to add versatility to the expression strategy (possibility of co-expression with the ligand), while the choice of XWnt8 was based on the successfully determined crystal structure of this Wnt family member in complex with Frizzled 8[27].

Small-scale testing showed that this construct expressed (**Figure 4.3 a**). However, upon large-scale expression and purification, combining a nickel-based affinity purification and gel filtration step, the vast majority of the protein aggregated, as evident from the size-exclusion chromatography elution profile (**Figure 4.3 b**).

In subsequent batches, several modifications of the initial purification protocols were performed, including addition of chromatographic steps (to better clean the protein before size-exclusion chromatography and to prevent possible aggregation with contaminating proteins), change of buffer, and use of additives such as L-arginine, glycerol, the non-detergent sulfobetaine NDSB-201 and the reducing agent DTT (see **Table 2.8** in **Materials and Methods** for details). Nevertheless, the tendency to aggregate still persisted and was not alleviated also in case of co-expression with the ligand. This, coupled to a detectable but not high expression of the protein, hindered progression to the set-up of crystallization trials.

During purification optimization, however, by applying increased concentrations of reducing agent (DTT), a shift towards a band of monomeric molecular weight in Western Blot analysis - from a series of bands likely representing different oligomers - was observed (**Figure 4.3 c**). This hinted at a disulphide-driven oligomerization; indeed, a literature report[138], although not peer-reviewed, was found to confirm the observation and to suggest the possibility that this behaviour was driven by the first two cysteines present in the loop.

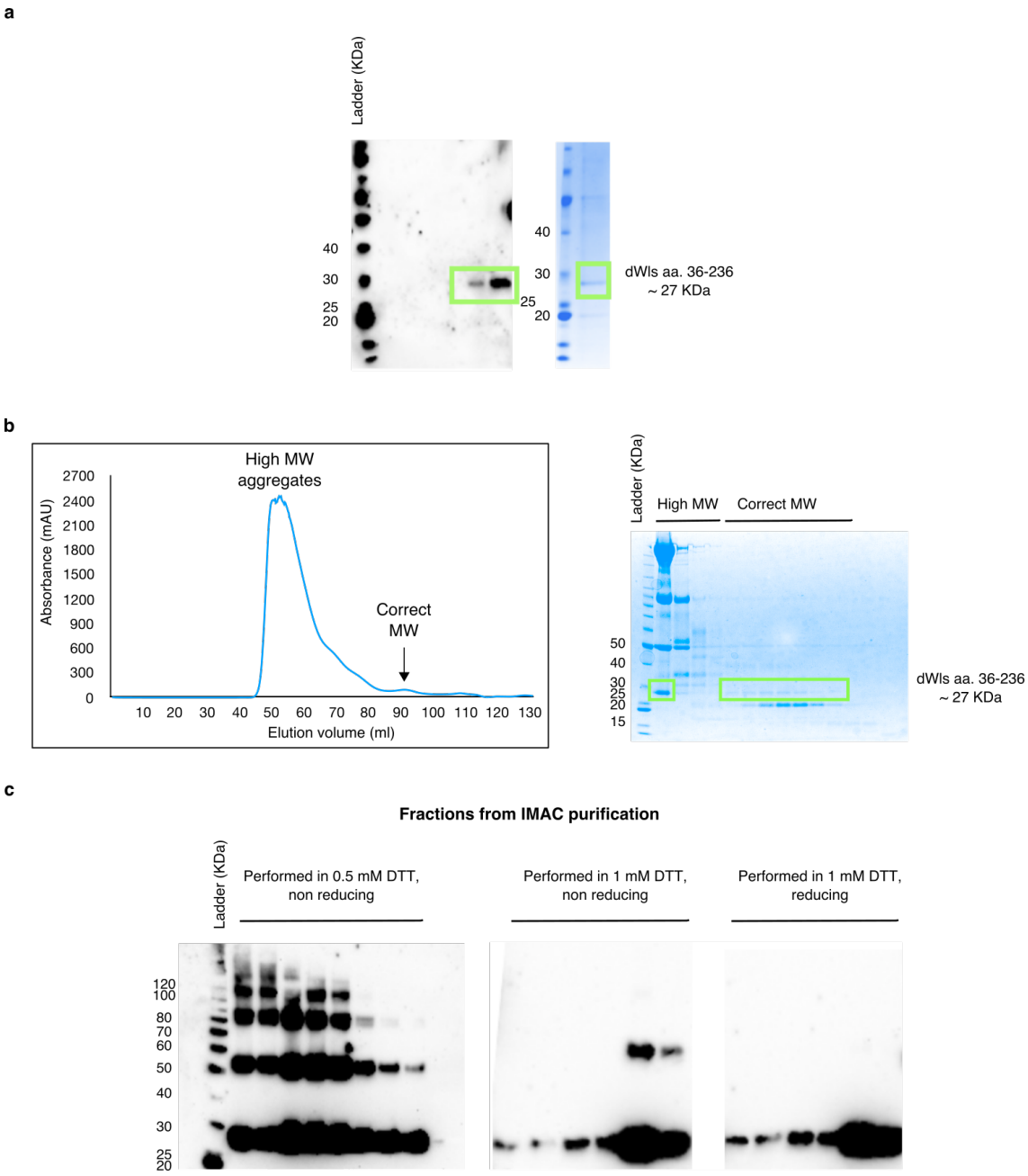


Figure 4.3. A *Drosophila* Wls loop construct could be expressed in *Drosophila* cells, but the protein aggregated.
a) Anti-His WB and SDS-PAGE on supernatants from small-scale expression. **b)** Large-scale expression: representative size-exclusion chromatography (HiLoad 16_600 Superdex 200 pg, GE Healthcare) elution profile and SDS-PAGE in reducing conditions, showing how most of the protein elutes with the high-molecular-weight peak, suggesting protein aggregation. **c)** Anti-His WB on fractions from IMAC purification performed with the addition of increasing concentrations of reducing agents in the purification buffer. The reduction in the high-molecular-weight bands in non-reducing conditions with increased DTT concentration, and in reducing conditions, suggests a disulphide-driven oligomerization.

This prompted me to reconsider my initial construct design strategy, which was based on the idea that all the four cysteines of the loop were involved in internal disulphide bonding, and therefore crucial to protein folding. I therefore decided to test further loop boundaries by differentially truncating cysteines out of the constructs (see **Table 2.5** in **Materials and Methods**), going back to the human Wls and to our standard mammalian expression system, HEK293, which was easier to transfect transiently and amenable to a quicker screening.

In this case, differently from our initial set of constructs tested, a human Wls loop region extending from aa. 84 to 229 - hence removing the first two cysteines and extending up to the predicted start of the 2nd transmembrane region (**Figure 4.4 a**) - showed very good expression, in particular when produced as a fusion protein with a C-terminal monoVenus fluorescent protein (**Figure 4.4 b**). Once again however, despite expressing, the construct showed an aggregating behaviour (**Figure 4.4 c**). Even in this case, extensive optimization of purification was attempted, employing the toolkit already described for the *Drosophila loop*, but introducing additional trials such as the use of detergents and tests of buffers with a wider pH range (see **Table 2.9** in **Materials and Methods**), but still without success in rescuing or preventing aggregation.

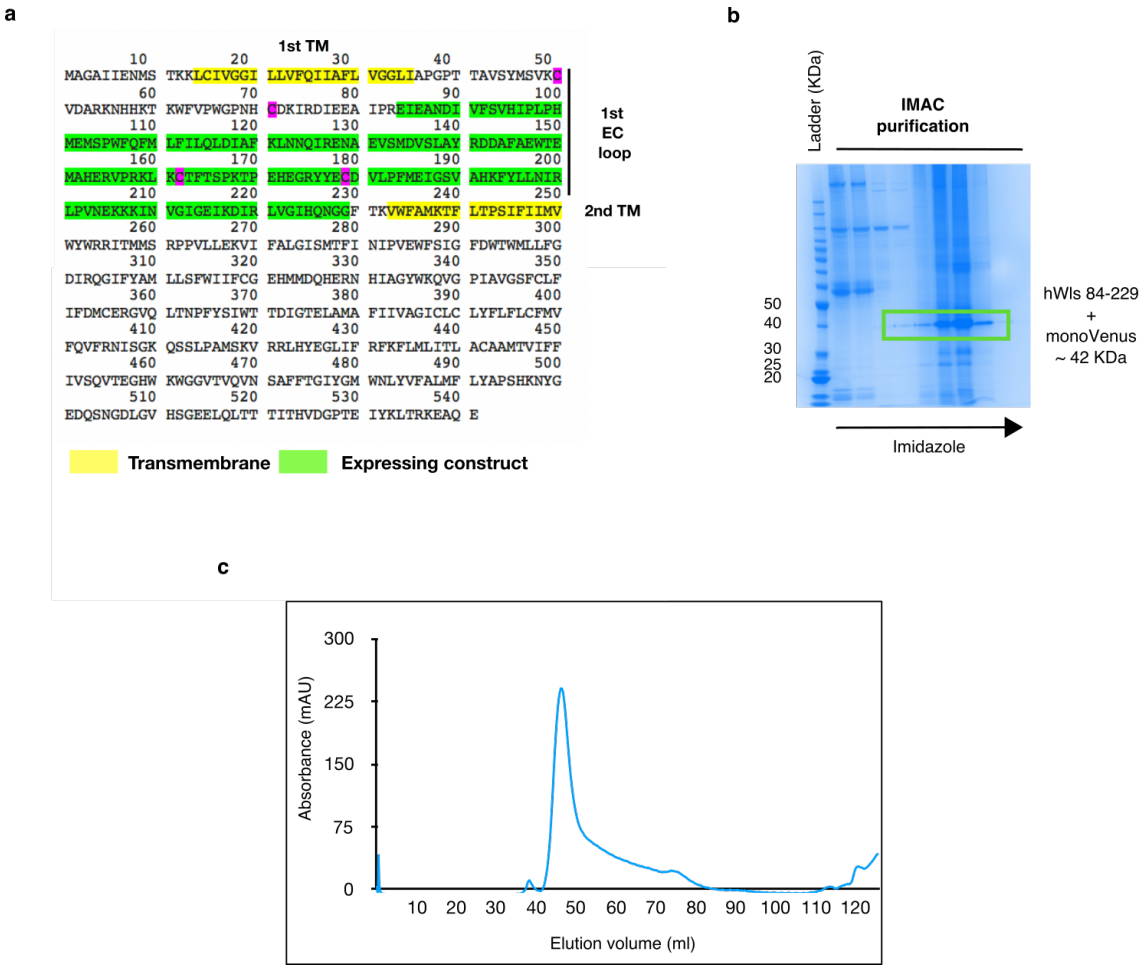


Figure 4.4. Human Wls loop cysteine truncations promoted expression in mammalian cells, but the protein still aggregated.
a) Overview of construct boundaries of the expressing construct mapped on the sequence of human Wls; 1 TM = 1st transmembrane region; 2 TM = 2nd transmembrane region; EC = extracellular. **b)** Large-scale expression: SDS-PAGE on representative fractions from IMAC purification, showing elution of the human Wls loop fusion construct with increasing imidazole concentration. **c)** Representative size-exclusion chromatography (HiLoad 16_600 Superdex 200 pg, GE Healthcare) elution profile of the SEC performed on the pooled IMAC fractions, indicating how most of the protein aggregates.

Nevertheless, two major achievements were obtained by the design of this new construct: the identification of suitable construct boundaries for expression, and the solution to the issue of low yield, that was a drawback of the insect expression

trials. Of course, aggregation - probably deriving from protein instability - remained however a significant obstacle to crystallization.

I therefore wondered if a tighter control of protein expression - like an inducible system - would help the protein behaviour, for example allowing better intracellular processing and folding of the protein.

Therefore, a 3c protease-cleavable Wls84-229-monoVenus fusion - the same construct described above used in transient transfection - was stably integrated by lentiviral infection in GnTI- TetR HEK293 cells, allowing doxycycline-activated induction of the expression. The tag was also changed from the His-tag, used so far, to a Twin-Strep tag, to reduce, from the early stages of the purification procedure, contamination coming from non specifically-binding proteins.

In this system, despite batch-to-batch variation, the protein generally showed a much higher degree of purity, and the gel filtration elution profile changed from the single aggregation peak of the constitutive, transient transfection, to a more diversified shape - probably representing different conformations/oligomeric states of the protein - with additional peaks and with the protein present in several elution fractions, including those corresponding to a monomeric and dimeric molecular weight (**Figure 4.5**). Since the yield and purity were now sufficient, the latter fractions were used to set up crystallization screenings, which have however so far failed to generate crystals.

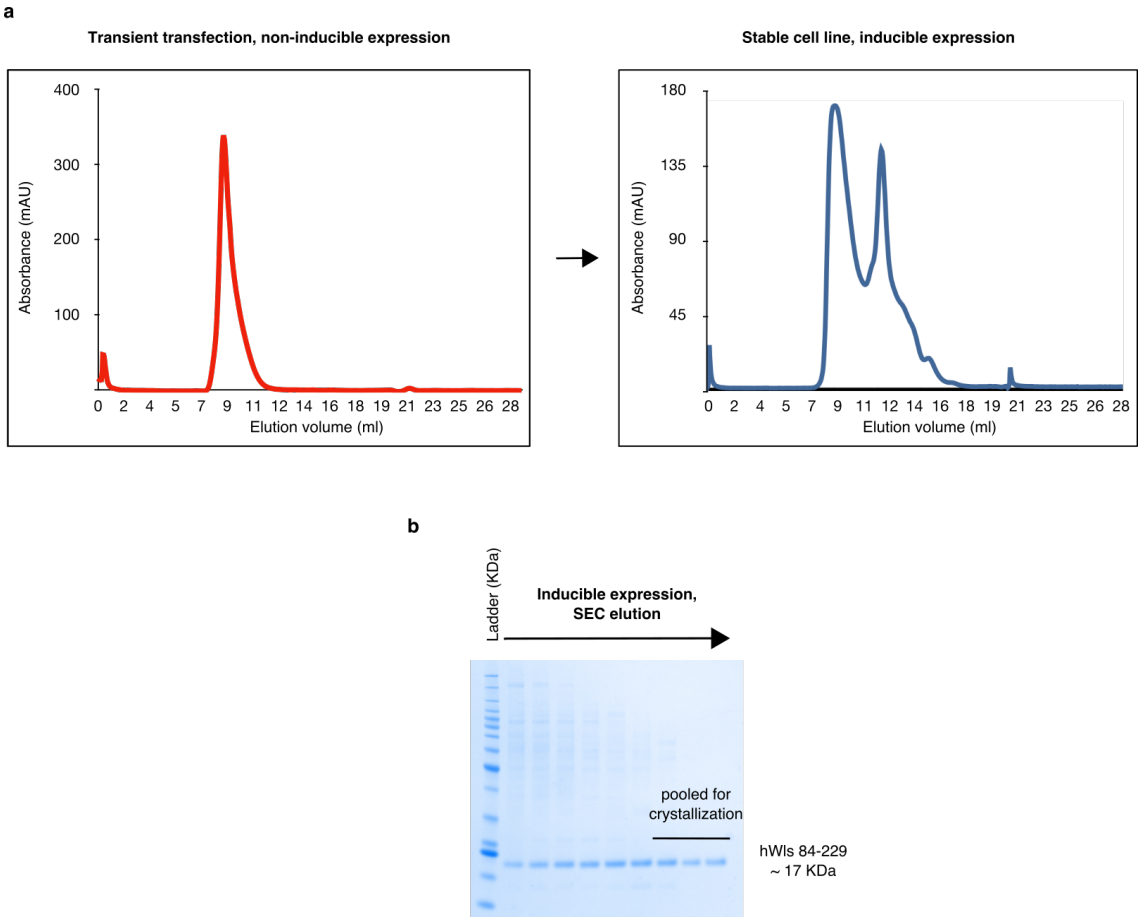


Figure 4.5. Switch to inducible expression improved the purification behaviour of human Wls loop.
a) Comparison between size-exclusion chromatography (Superdex 200 Increase 10_300 GL, GE Healthcare) elution profiles of human Wls aa. 84-229 expressed upon transient transfection of HEK293S GnTI- cells (left) and upon induction of expression in stably transduced HEK293S GnTI- TetR cells (right), after 3c cleavage and removal of tag. **b)** SDS-PAGE of representative SEC fractions from peaks of the right chromatogram of panel a, demonstrating a higher degree of purity of the protein, and its presence even in fractions corresponding to lower molecular weights close to the monomeric size (pooled for crystallization).

Full-length transmembrane Wntless: expression and purification attempts.

Despite the constant improvement in the expression and purification of the loop, finally yielding some material for crystallization trials, the outcome was thus not

optimal. Even in the case of the optimized inducible construct, indeed, a significant part of the protein still suffered from aggregation, and even the low-molecular-weight, unaggregated sample recovered for crystallization did not show an ideal chromatographic profile, possibly explaining why the screening attempts were unsuccessful in yielding crystals.

Hence, I started thinking that, in contrast to what was proposed in the literature, the 1st extracellular Wls loop was not able to fold and function in isolation, and that a potential reason for the apparent loop instability and aggregation propensity was indeed the absence of the native context of the other loops and regions of the full-length protein - possibly interacting with the 1st loop, and stabilizing its structure.

Therefore, I decided to shift my attention to the production of full-length human Wntless. 7 constructs, exploring different combinations and truncations of transmembrane regions and loops (see **Table 2.6** in **Materials and Methods**), were transfected in inducible GnTI- TetR HEK293, given the promising results obtained for the 1st extracellular loop with these cells. Additionally, the proteins were produced with a cleavable C-terminal monoVenus fusion partner, to facilitate screening by fluorescence-detection size exclusion chromatography (FSEC). A combination of the non-ionic detergent n-dodecyl-β-d-maltoside (DDM) and the cholesterol ester cholesterol hemisuccinate (CHS), at a 5:1 w/w ratio, was employed as detergent of choice both for protein extraction from the plasma membrane and for downstream FSEC analysis. This is in fact one of the most successful detergent compositions used to solve membrane protein structures according to analyses of the PDB database[139], and is indeed the

standard detergent used in membrane protein studies in our laboratory, based on positive results obtained with Frizzled receptors.

Out of the 7 constructs screened, 2 presented a good level of expression and a promising chromatographic behaviour in FSEC in small-scale: namely, Wls 1-495 (encompassing all the transmembrane regions but excluding the disordered C-terminal tail) and the complete full-length 1-541 (**Figure 4.6 a**).

With the aim of crystallization, Wls 1-495 underwent scaling-up first. In contrast to the small-scale screening, however, the procedure was problematic since at the first step of purification, streptavidin-based affinity: both the yield and the purity of the recovered protein were in fact low, reflecting a significant non-specific binding of contaminant proteins. Subsequent size-exclusion chromatography confirmed very low yield, and additionally showed a different chromatographic behaviour compared to the satisfyingly monodisperse peak obtained in small-scale tests (**Figure 4.6 b**).

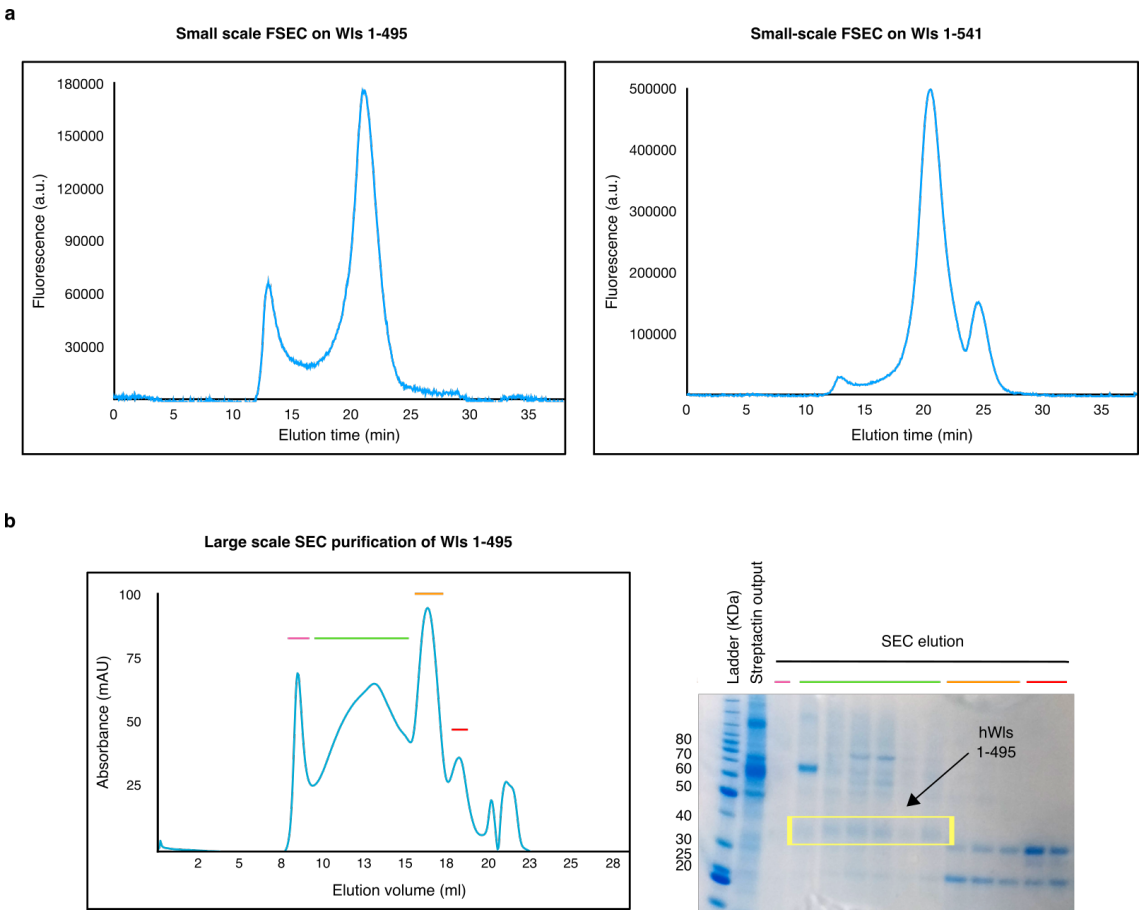


Figure 4.6. Human Wls transmembrane constructs well-behaved in small-scale test, but not upon expression scaling up.
a) Small-scale FSEC on human Wls aa. 1-495 and human Wls aa. 1-541, after induction of expression for 48 hours at 37 °C with 1 ug/ml doxycycline. The run was performed on a Superose 6 Increase 3.2/300 column. **b)** Scaling-up of the expression of human Wls aa. 1-495. As described in **Materials and Methods**, the protein was first loaded on a Strepactin column, underwent on-column 3c-cleavage, and then was run on a Superose 6 Increase 10/300 GL column (GE Healthcare). The SEC elution profile and the related SDS-PAGE are shown. The yellow box highlights the cleaved eluted Wls aa. 1-495, as confirmed by mass spectrometry analysis.

In the case of membrane proteins, such behaviour is not uncommon, and might often depend on an inappropriate choice of the detergent employed, resulting in protein instability. For this reason, a detergent screening was performed by FSEC in small-scale, testing 32 detergent conditions belonging to different detergent classes and with different ratios of detergent:CHS (see **Table 2.7** in **Materials**

and Methods). The screen identified three hits that improved the behaviour of the protein with respect to DDM: Cymal-6 (a cyclohexyl-lalkyl-β-D-maltoside) and the so-called NG detergents LMNG (Lauryl Maltose Neopentyl Glycol) and DMNG (Decyl Maltose Neopentyl Glycol), a recently developed class presenting a quaternary central carbon which allows the detergent to have two hydrophilic heads and two lipophilic tails, promoting protein stability. For all the three hits, the best ratio with CHS was found to be 10:1 (detergent: CHS, w/w) (**Figure 4.7**).

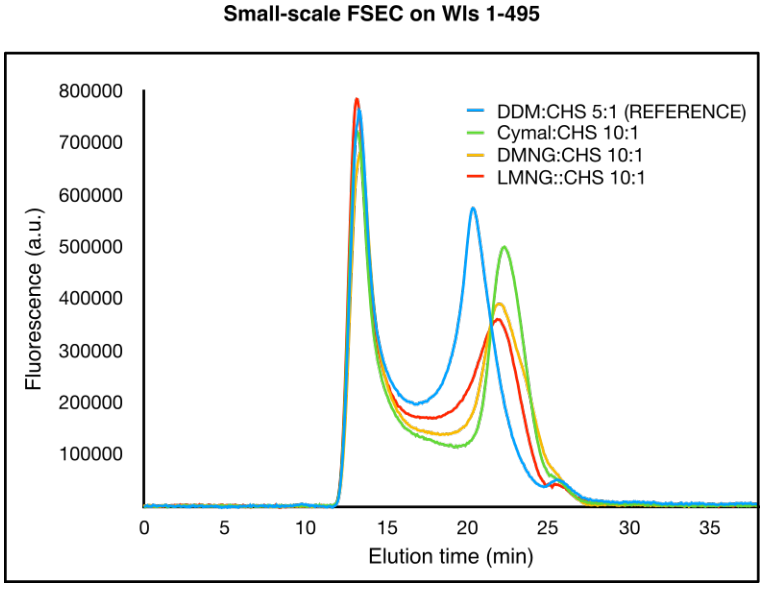


Figure 4.7. Detergent screening by FSEC on human Wls aa. 1-495.
Out the 32 conditions tested, those shown improved the protein behaviour as reflected by the elution profile compared to reference (DDM:CHS 5:1).

Nevertheless, when the protein production was scaled up, the outcome didn't improve even in the presence of these detergents. Similar lack of success also held for the other expressing construct, Wls 1-541.

For this reason, and due to the increased focus dedicated to the other two, more successful, projects presented in this thesis (the glypicans and Notum), the work on Wntless was mainly stopped at this point.

Isolation of human Wls transmembrane protein in SMALPs

Recently, however, a newly developed system for detergent-free membrane protein extraction has been introduced in our lab. This is based on polymer-bounded lipid nanodiscs, called styrene-maleic acid lipid particles or SMALPs. Upon addition of styrene and maleic acid to biological membranes, in fact, the polymers solubilize the membrane protein by forming an annulus surrounding a central lipidic core, which conserves the bilayer organization and the native lipid composition[140]. This creates an environment which retains most of the physical and biochemical features of the plasma membrane, including protein-lipid interactions which might play an important role in protein stabilization and are not mimicked by the micellar organization of detergents.

Extraction of Wls 1-541 was therefore attempted using SMA particles and a preliminary scaling-up in this case showed very promising results. As can in fact be seen in **Figure 4.8**, and differing strongly from previous trials, the protein is almost pure after the first streptavidin affinity purification step. Although further characterization was not possible due to time constraints, and the outcome will of course need to be confirmed and reproduced, this is a very relevant progress which can represent a turning point in the structural characterization of this important Wnt carrier.

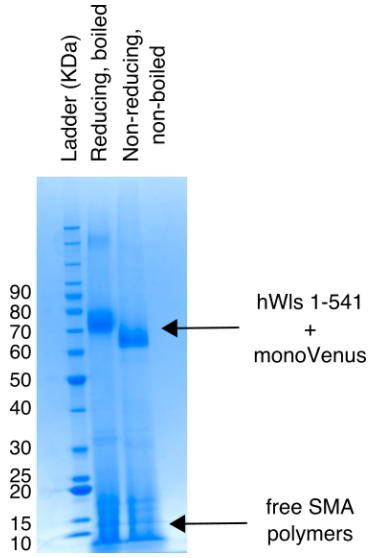


Figure 4.8. Extraction of human Wls aa. 1-541 in SMALPs improved protein purity and yield recovered.
SDS-PAGE of SMALP-extracted protein after elution from Streptactin column. In this case, 3c cleavage was not performed and the protein is still fused to monoVenus. The low-molecular-weight streak corresponds to free SMA polymers interacting with the Coomassie dye, according to ref. [119].

Discussion

Despite the extensive troubleshooting required, both the recombinant expression of the putative Wnt-binding unit of Wls - the 1st extracellular/intraluminal loop – and that of the full-length integral membrane protein have seen continuous improvement over the course of the experimental work.

From no expression of the loop, a thorough assessment has allowed me to first identify likely domain boundaries, to in turn progressively optimize the yield, and to a minor extent, the chromatographic behaviour of the sample, and even set up crystallization screens with the purified protein. Particularly interesting is the fact that, among a significant variety of constructs assessed with very different length and boundaries in the same expression system, the only construct expressing to

a satisfying extent was the one truncating two out of four cysteines present in the loop. This might in fact argue against the most obvious hypothesis that these cysteines are involved in disulphide bridges between themselves, and might instead suggest that at least some of them (the first two) mediate intermolecular rather than intramolecular contacts. This is line both with the results of some of the experiments presented here (namely, the fact that increasing concentrations of reducing agents appear to partially rescue the aggregation propensity of the protein) and with some reports from the literature. On the other hand, the possibility can't of course be ruled out that this aggregation is just due to non-specific disulphide-driven oligomerization promoted by an improperly folded protein. In this case, the apparent beneficial effect of reducing agents and of cysteine truncation would just be artefactual and would not reflect any underlying functional properties of the first two cysteines of the loop. In support of this second view, it should be considered for instance that, even after truncation of these cysteines, most – although not all – of the protein still aggregated in subsequent purification attempts.

In either case, however, having tested such a wide range of constructs and experimental conditions without success in fully rescuing the aggregating behaviour, it looks likely that the loop would anyway need some additional contacts/interaction/structural support coming from its native environment to properly fold or to be stabilized, and might not be functional in isolation. For this reason, future work aimed at getting structural information on Wntless might benefit more from a focus on the expression of the integral membrane protein rather than on its 1st extracellular loop.

In this sense, the fact that the full-length protein can be purified in polymer-based nanoparticles like SMALPs is both very promising and intriguing. It might in fact suggest that native lipid-protein interactions (which are not faithfully replicated in detergents, hence the aggregation upon detergent-based extraction) are crucial to promoting protein stability, and that the behaviour of the protein might be influenced by specific membrane lipid composition. It has in fact been proposed that the preferential incorporation of an unsaturated rather than saturated lipid could restrict Wnt adsorption to membranes with an appropriate curvature and, therefore, phospholipid composition[31]. If this same composition also affected Wls stability, this would provide an additional layer of regulation to Wnt transport and localized pathway activation.

For future work, the first step would of course be a better characterization of the polymer-isolated membrane protein: while the stark difference in purity after streptavidin purification between detergent-solubilized and SMA-solubilized Wls is encouraging, assessment of the oligomeric state would be needed to rule out that the protein aggregates even in this case. To this aim, analytical ultracentrifugation (AUC) might be the most appropriate technique, since, in the case of SMALPs, potential interactions with the resin might hamper appropriate determination of the molecular mass of the polymer-protein complex by size-exclusion chromatography[119]. Inserting a gel-filtration chromatographic step in the purification protocol might anyway prove useful to separate free and excess SMA and lipids from the protein-containing SMALPs, ensuring a more homogenous sample and reducing artefacts in downstream analysis based on UV spectroscopy methods (SMA in fact absorb at 254 and 280 nM[119]). Negative

stain electron microscopy is another method that could help in characterizing SMALPs and evaluating the particle aggregation state[140], and would even provide an initial insight into the feasibility of obtaining high-resolution structural information on Wls by cryo-electron microscopy.

In this sense, it is conceivable that, even if the extraction of apo Wls in SMALPs is successfully confirmed, its EM structural characterization will still prove challenging, due, among other factors, to the small molecular weight of the protein. However, the definition of an optimized purification protocol for the full-length Wls aa. 1-541 in particular offers two distinct advantages: first, this construct retains the C-terminal interaction tail relevant for binding to proteins of the retromer complex, and complex formation and co-purification strategies can be devised; moreover, the polymer-embedded Wls can be used to generate antibodies which can help EM imaging and data analysis.

In parallel, it will be worth trying the SMALP protocol for the other transmembrane construct, i.e. Wls 1-495 lacking the C-terminal disordered tail. Reports in the literature suggest in fact that transfer from SMALPs to lipid cubic phase (LCP) is possible[141], and might represent a viable strategy for membrane protein *in meso* crystallization.

In the shorter term, however, a successfully purified and non-aggregating Wls could already be used to investigate some important biological questions related to the protein mechanism of action. Since the SMALPs-embedded proteins can be treated as soluble globular proteins, the sample can in fact undergo biophysical protein-protein interaction studies to evaluate Wls interaction partners and elucidate Wls' role in the context of the Wnt handover process, as will be

more extensively discussed in the final **Discussion** chapter. Likewise, the postulated role of pH changes in Wnt binding and release can be assessed, for example by using an FSEC setup similar to that employed in the case of the glypican mutants, where the ability of Wls to bind the fluorescent palmitic acid probe could be evaluated in different running buffers with a varied pH range.

Addendum: discussion points from the structure of human Wls-human Wnt8a

Very recently, in the final phase of thesis drafting, a cryo-EM structure of human Wls in complex with its ligand, human Wnt8a, has been solved and published[142]. This section will thus give an overview of the main features of the complex to, in turn, retrospectively analyse some results presented in this chapter.

The Wntless structure is composed of eight transmembrane regions present in the mature protein (**Figure 4.9 a**), with both the N- and C-termini protruding towards the cytosolic side of the membrane. This rules out the hypothesis, postulated in the literature, of the first transmembrane region being a cleaved signal sequence - something that appeared already evident from the studies presented in this chapter on the full-length transmembrane protein, where the expressing constructs indeed included the first transmembrane helix. Interestingly, TM regions 2 to 8 present a structural arrangement similar to the seven-helix core of GPCRs, and Wls shows in particular a very close homology with the transmembrane portion of human Fz4 reported in ref. [143]. Further studies will be needed to investigate this important observation and evaluate whether the

homology with GPCRs might have some functional role, or if the activity is carried out (mainly) in a G protein-independent way, as for Frizzleds.

As already noted in this chapter, in addition to the transmembrane moiety, Wntless presents a long, structured 1st extracellular/luminal loop (called "luminal domain" in the paper), which is relevant for the interaction with Wnt. Different from what was predicted[46], this loop is however not the lipid-binding unit, but instead extensively binds only the protein part of Wnt. Its fold is composed of eight antiparallel β-strands, and the four cysteines present in this region all form intra-domain disulphide bonds, and not intermolecular disulphides, as was hypothesized at a certain point during the experimental work on Wntless presented in this chapter (the topic will be further discussed below).

The PAM moiety of Wnt sits instead in a deep hydrophobic tunnel located in the transmembrane core of Wntless, in a cavity that is open both on the luminal/extracellular side and on the lipid bilayer (**Figure 4.9 b**). This reinforces the suspicion, expressed above, that the protein might depend on the surrounding lipid environment for its stability. Indeed, electric potential consistent with glycerophospholipids used for the reconstitution in nanodiscs (POPG) - the method used for protein purification - is visible in the cryo-EM map in the lipid-binding pocket, close to the palmitoleic acid (**Figure 4.9 b**).

Finally, in comparison to the Wnt-Fz complexes shown in the- Introduction, the Wnt in the Wls-bound structure displays significant flexibility in particular at the level of hairpin 3 (the "index" finger, see **Figure 4.9 c**), and this conformational freedom might be important for the cargo offload and the passage of the morphogen to other receptors.

It is very interesting to evaluate the experimental findings presented in this chapter in light of the structure of the Wls-Wnt complex.

It is indeed worth noting that the boundaries of the initial set of constructs designed for the 1st extracellular loop (see **Figure 4.1** and **Materials and Methods Table 2.3**) correctly encompassed all the eight secondary structure elements forming the β-sandwich fold of the domain. The fact that no such constructs expressed (see **Figure 4.2**) in HEK293 cells (the same expression system used for the solution of the Wnt-Wls structure in the paper) reinforces the idea that the loop might not be capable of proper folding in isolation. The rest of the protein might provide stabilization either by shielding hydrophobic patches (**Figure 4.9 d**) or by anchoring the loop in a proper orientation to promote the correct formation of the two disulphide bonds stabilizing the fold (an improper intramolecular or intermolecular disulphide bond formation might explain the beneficial effect observed with reducing agents in **Figure 4.3 c**). Still on the theme of cysteines, it is therefore remarkable that a construct truncated for one of the two disulphide bridges was secreted in the media and represented the most promising construct for the 1st extracellular loop (see **Figure 4.5**). On the other hand, the truncated region mainly included an unstructured set of amino acids crucial for the interaction with Wnt, but not for the core folding of the β-sandwich. Actually, the first β-strand was also truncated, but it's worth remembering that this construct was expressed as a fusion partner with the β-strand-rich monoVenus, which could have worked as a stabilizing unit surrogating the first β-strand of Wntless.

A further important element for the stabilization of the loop is represented by the Wnt ligand. In the paper, the authors indeed explicitly stress the importance of the co-expression of Wnt and Wls to ensure correct transport of both proteins to the cell surface, and their ability to express hWnt8a in HEK cells is a great result, given the known difficulties in producing the morphogen in mammalian cells in quantities suitable for structural studies, as explained in the **-** Introduction.

The need for the morphogen might also clarify why the only expressing construct of the 1st loop in isolation (and without a fusion partner) was obtained in insect cells by co-expression of the protein with *Xenopus* Wnt8. The limited amount of morphogen that I was able to express, its limited solubility and, possibly, the limited affinity of Wnt for Wls in the absence of the lipid-binding cavity in Wls, impacted however on the yield I could recover. Furthermore, as noted above, it is anyway likely that the construct was suffering from misfolding because of the lack of the transmembrane region, motivating its aggregation propensity.

Given the apparent importance of Wnt in promoting the expression of Wls, it is a significant achievement that, using SMALPs, I am instead able to purify Wls even in the absence of the morphogen, in an uncomplexed form. If the protein purified in such a way will prove to be fully functional and well-behaved upon deeper characterization, it will represent a very powerful tool for biophysical studies of the Wls interaction with other Wnt receptors, as outlined in the final **Discussion** chapter. The SMALP method presented in this chapter could therefore represent a strong addition to the protocol developed in the paper to get the EM structure of the Wnt-Wls complex.

In this sense, it is also relevant to highlight that the latter complex was reconstituted in nanodiscs, and this requires preliminary solubilization of the proteins in detergents. As already noted, the SMALPs procedure, conversely, completely removes the need for detergents, therefore allowing the study of the interaction of Wntless with the constitutive native lipid environment and membrane composition.

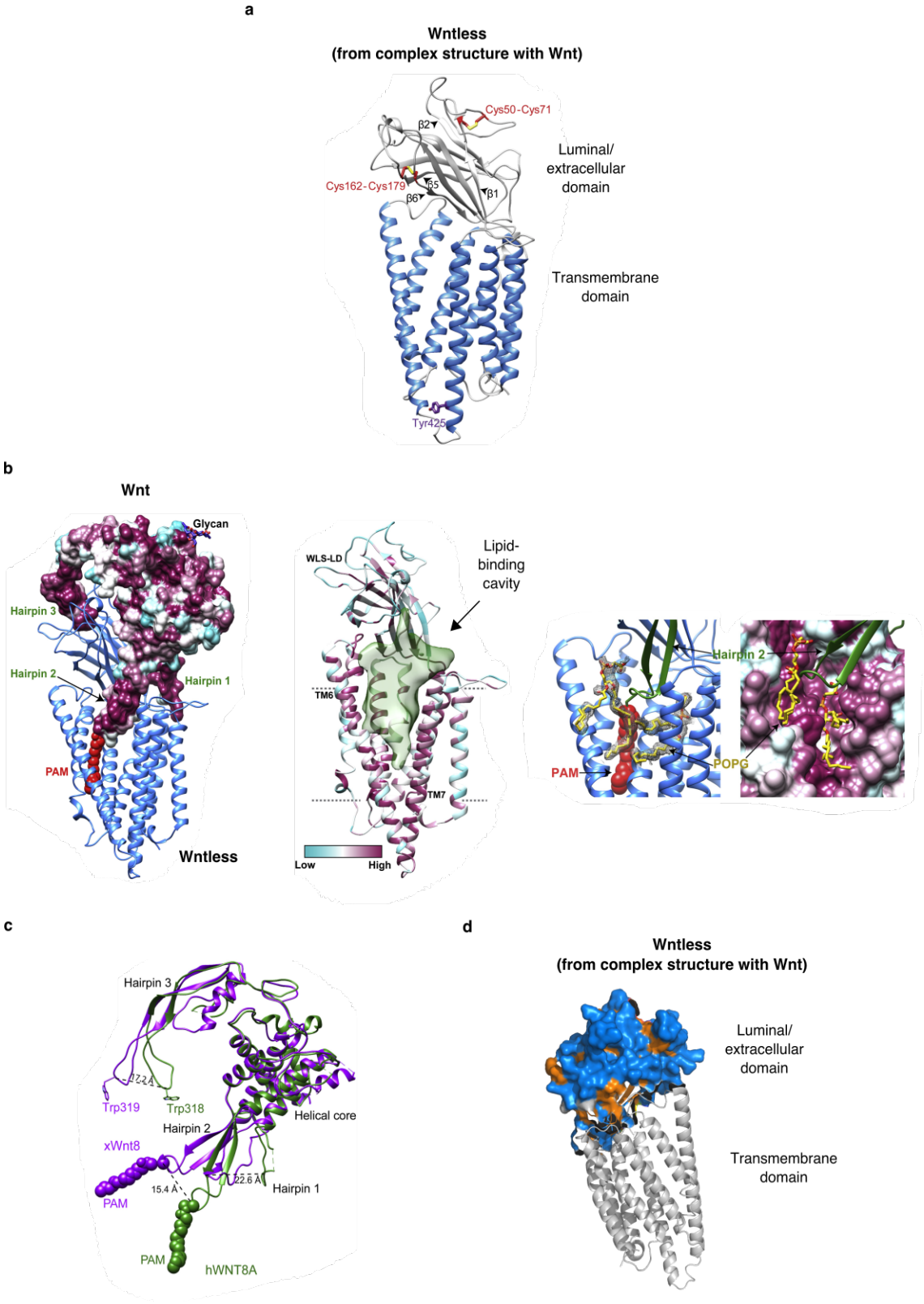


Figure 4.9. Structure of human Wntless in complex with human Wnt8a. **a)** Structure of human Wntless extracted from the complexed structure, with the disulphide bonds formed in the first extracellular/luminal loop shown. **b)** Complex structure with Wnt, showing the position of the lipid-binding cavity

accommodating the palmitoleic moiety of Wnt (PAM). Wls is coloured in marine in the left panel, with Wnt coloured by conservation. Wls is instead coloured by conservation in the central panel. The left panel shows the electric potential map and the positioning of the POPG lipids in the Wnt-binding cavity. **c)** Superposition between human Wnt8a from the Wnt-Wls complex structure and Xenopus Wnt8 from the Wnt-Fz complex structure from ref.[27]. **d)** Hydrophobic patches on the surface of the 1st extracellular/luminal loop of Wntless. The hydrophobic residues are coloured in orange, with the rest of the loop depicted in marine. The transmembrane domain is shown in grey. Panels a, b, and c are derived from ref.[142].

Chapter 5 - Targeting the Wnt extracellular modulators

Background and rationale for targeting the Wnt pathway in Alzheimer's disease

The Wnt pathway in the adult brain

In addition to its long-established role in neural development, evidence is accumulating about an important involvement of Wnt signalling in postnatal/adult brain function as well[144]. Wnt ligand expression has in fact been reported in all the major components of mice adult cerebral cortex, including the neocortex, the thalamus and the hippocampus, with layer-restricted and region-restricted expression of some Wnts - e.g. Wnt2b, Wnt5a and Wnt7b in different regions of the neocortex[145]; in parallel, post-natal cerebral expression of several Frizzled receptors has also been shown[145]. Functionally, this has been associated with a role for Wnts in the maintenance of the mature synaptic architecture both in the pre-synaptic[146] and in the post-synaptic compartments[147], in the regulation of synaptic plasticity[148], in axon pathfinding[149] and in dendritic spine growth and structure[150]. The synaptic maintenance and synaptic plasticity seem to be controlled by a delicate equilibrium between canonical and non-canonical Wnt pathways, with both arms implicated in the early and late phase of long-term potentiation (LTP) – a crucial neurophysiological correlate of increased synaptic strength. Conversely, a reduction in the canonical, β-catenin dependant pathway

signalling, with a concomitant increase in non-canonical Planar Cell Polarity signalling, appears to be implicated in synapse disassembly[151].

Wnt pathway in the ageing brain and role in Alzheimer's disease

In the ageing brain, reduced expression of several Wnts and intracellular Wnt mediators have been described both in the human and rodent brain, and reduced Wnt activation has been associated with synaptic loss, a key feature underlying the cognitive decline present in neurodegenerative diseases like Alzheimer's Disease (AD)[151]. Conversely, upregulation of extracellular Wnt negative modulators like Dkk1[152] (which acts by binding Lrp5/6 and antagonizing signalling activation) and sFRP[153] (a potential Wnt-interacting molecule, see **Introduction**) has been reported. Moreover, conditional deletion of *Lrp6* led to dendritic spine loss, memory alterations and neuroinflammation in the APP/PS1‡‡‡ mouse model of Alzheimer's disease[154].

In addition to synaptic loss, the Wnt pathway has been associated with AD by two further mechanisms: direct regulation of the formation of pathological mediators of the disease, and contributions to changes in the integrity of the blood-brain-barrier (BBB).

With respect to the first aspect, the already cited conditional knock-out of *Lrp6* in the AD mouse model led to increased β-amyloid production via increased processing of the Amyloid Precursor Protein (APP), whereas activation of the Wnt

‡‡‡ Harbouring mutations both in the Amyloid Precursor Protein and the Presenilin 1 genes, resulting in increased β-amyloid expression levels and deposition.

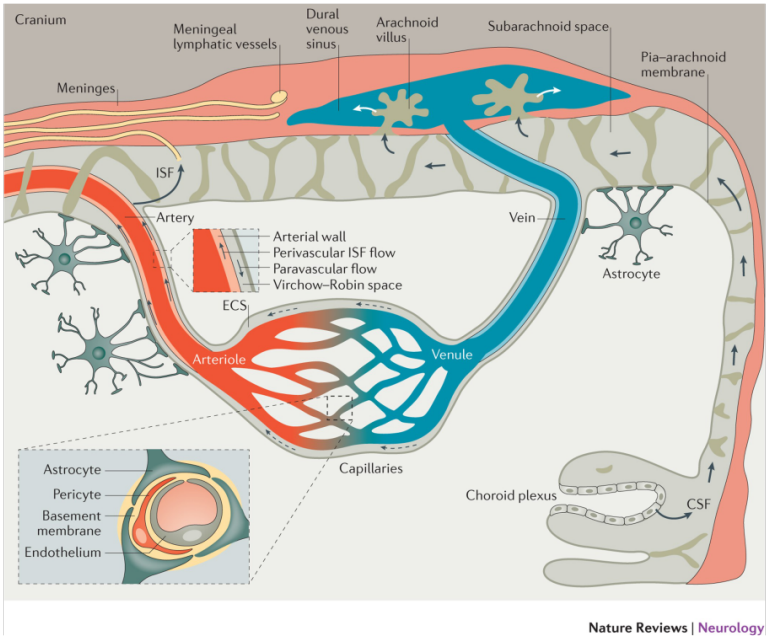
intracellular negative regulator GSK-3 has been shown to hyperphosphorylate the microtubule-associated protein tau, promoting the formation of neurofibrillary tangles[155]. Moreover, the Wnt pathway has been reported to control the expression of the Wnt neural target gene REST, a neuronal transcription factor induced in the normal ageing human brain, which elicits transcription of stress response genes and therefore protects neurons from oxidative stress and amyloid toxicity: reduced Wnt activation in turn determines reduced REST levels, with impaired protective ability[156].

Regarding the blood-brain-barrier, this is a continuous endothelial membrane formed by an endothelial monolayer sealed by tight junctions between cells and supported and wrapped by mural cells (pericytes and vascular smooth muscle cells) and astrocyte end-feet. This architecture and the close connection between endothelial cells hamper the passive transport of most substances from the blood to the brain, therefore protecting the brain interstitial fluid and the nervous milieu from potential contact with toxic or harmful molecules. As a consequence, the BBB controls the specialized transport and regulated entry of a plethora of molecules into the brain (**Figure 5.1 a**); preservation of its function is therefore crucial[157].

Cerebral vessel pathology is a key hallmark of AD[158], and the initial vascular insult has been suggested to lead to alterations in the integrity of the BBB. This in turn worsens the vascular dysfunction, ultimately causing neurodegeneration through a series of mechanisms (**Figure 5.1 b**) including impaired glucose transport, infiltration of leukocytes and inflammatory mediators in the brain, and impaired removal of toxic species (reduced activity of the multidrug transporter P-

glycoprotein 1 causing decreased β-amyloid clearance). These alterations have been observed in AD patients both by neuroimaging studies and in post-mortem histopathological analyses[157].

a



b

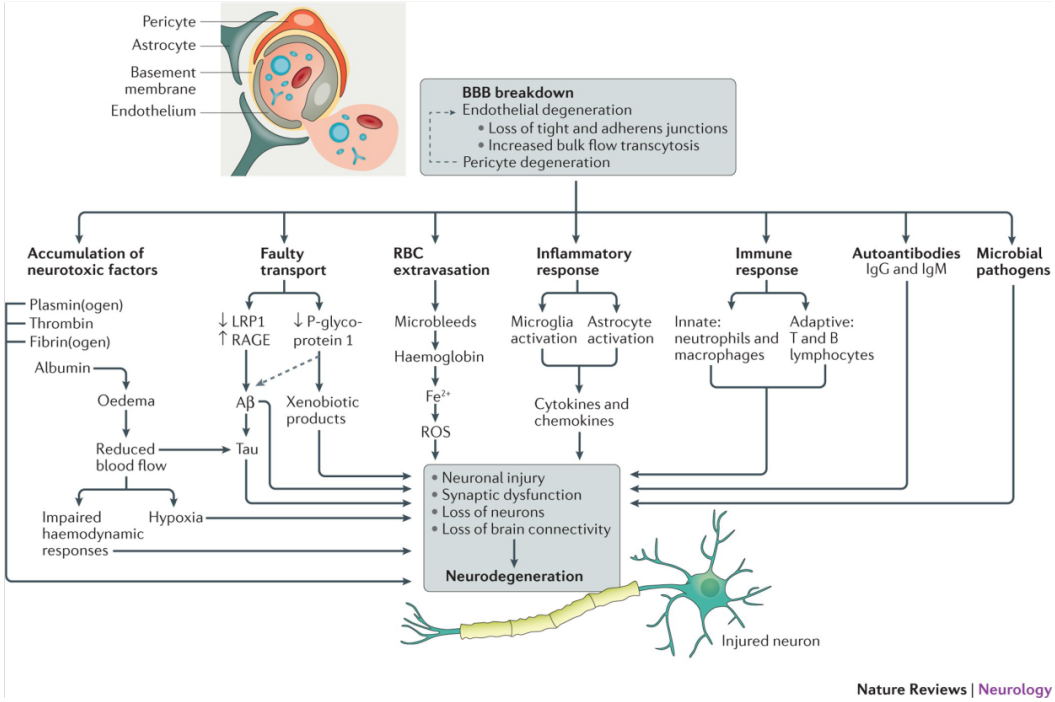


Figure 5.1. The blood-brain-barrier (BBB) and its involvement in neurodegeneration.
a) Architecture of the BBB. ISF: interstitial fluid; CFS: cerebrospinal fluid; ECS: extracellular spaces. **b)** Mechanisms of neural damage following BBB breakdown. RBC: red blood cell; ROS: reactive oxygen species; LRP1: LDL receptor-related protein 1; RAGE: receptor for advanced glycosylation end products. The panels are derived from ref. [157].

Wnt7a and Wnt7b have been demonstrated to regulate BBB maintenance forming a complex with, and signalling through, Reck (a GPI-anchored protein) and Gpr124 (an orphan G-protein-coupled receptor)[159]. Claudin-1 and claudin-3, two major tight junctions connecting the endothelial cells of the barrier, are also under the direct transcriptional control of canonical Wnt pathway[160]. Finally, Wnt also promotes the expression of the glucose transporter GLUT1 and the already cited P-glycoprotein 1[161].

Due to the summarized involvement of the Wnt pathway in many aspects of brain function and neurodegeneration, restoration and boosting of a pathologically reduced signalling appears therefore a promising therapeutic avenue for the treatment of Alzheimer's Disease.

Targeting Notum

Some attempts at targeting Wnt pathway in Alzheimer's Disease have already been performed using GSK-3 inhibitors; however, the results have been so far unsatisfactory, mainly due to selectivity issues related to the multiplicity of functions and targets of GSK-3, which are not only specific to the Wnt pathway[161]. Therefore, the development of drugs against targets able to more selectively reactivate the Wnt pathway is particularly compelling.

From this perspective, Notum might represent a very attractive candidate. First, as already described (see **Introduction**), it is a Wnt pathway-dedicated modulator. In addition, its steric specificity for unsaturated fatty acids, and the dependence of its enzymatic activity on a catalytic triad (Ser 232, Asp 340 and His 389), whose transition state intermediate oxyanion hole is stabilized by several additional residue amides (Gly126, Gly127, Trp128, Ala233), in the presence of a well-defined hydrophobic pocket with a narrow but flexible entry mouth, make the protein very attractive from a druggability point of view. Moreover, Notum is required for brain development in vertebrates[162], and was found among the top 20 up-regulated genes after short-term (4h) Wnt3 stimulation in rat hippocampal neurons[163].

On these premises, a drug development program has been organized by our lab and a consortium of laboratories composed of medicinal chemists and cell biologists, from the Francis Crick Institute London (Vincent lab) and the University College London (Salinas lab, Fish lab, Whiting lab). The parallel aims of the Consortium are to validate the potential of Notum as target for AD therapy - by a better molecular understanding of its underlying biology in the disease - and to design potent brain-penetrant drugs through iterations of structure-guided drug design and optimization.

Expression, purification and crystallization of Notum

Notum was expressed and purified according to the published protocol devised in our lab (see **Materials and Methods** and ref. [85]). Briefly, the protein was

expressed by transient transfection of GnTI-deficient HEK293S cells, and initially purified by nickel-based affinity chromatography, followed by polishing by size-exclusion chromatography, whose purest fractions were collected for crystallization (**Figure 5.2 a** presents a representative example).

The major focus of the crystallization experiments was aimed at obtaining crystals reproducibly and consistently diffracting at high resolution, allowing subsequent soaking of the different classes of inhibitors designed during the structure-based drug development campaign. To this goal, several conditions previously shown to harbour apo- and Wnt-peptide bound Notum crystals were screened and optimized, with the best crystals growing in an interval of precipitant between 1.3-1.5 M of ammonium sulphate and a pH of 3.9-4.1 in citric acid buffer, in a time frame between 1 and 3 weeks (see **Figure 5.2 b** for representative examples). In these conditions, the crystals usually diffracted at a resolution between 1.2 and 1.6 Å.

With respect to inhibitor soaking, for each compound different intervals between crystal soaking and harvesting were tested (see **Materials and Methods**), with the best results generally obtained upon overnight soaking.

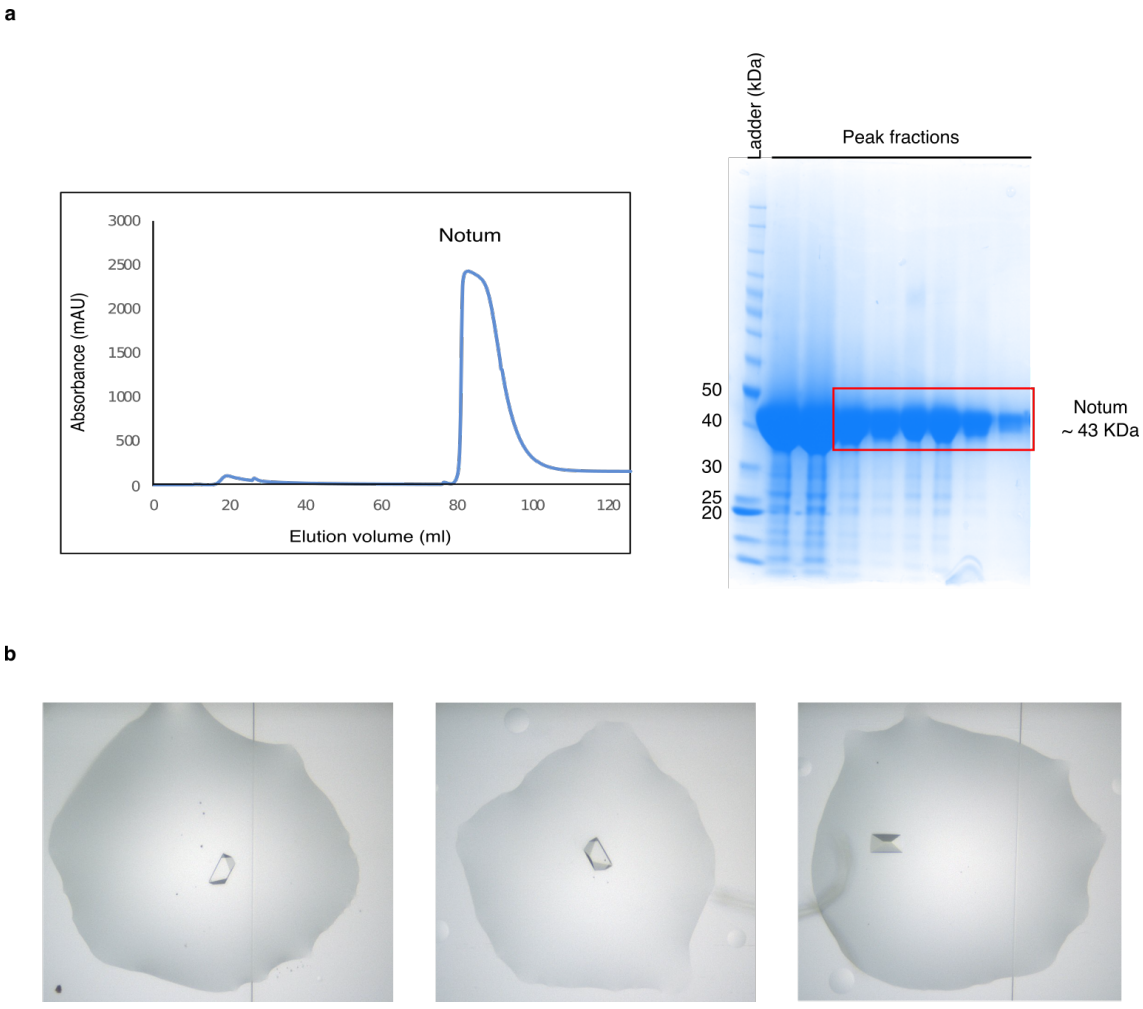


Figure 5.2. Representative purification and crystallization of Notum. **a)** Size-exclusion chromatography (HiLoad 16_600 Superdex 200 pg, GE Healthcare) elution profile and SDS-PAGE of the peak fractions. In the red box the fraction collected for crystallization experiments. **b)** Representative crystals.

From fragment screening to lead candidate: overview of series development and representative crystal structures

Initial scaffolds for drug design were previously identified by our lab using a fragment-screening, high-throughput automated approach at Diamond Light Source (performed by Dr Yuguang Zhao) (**Figure 5.3**). This was based on a

library of 776 “poised” fragments[164], i.e. fragments that can be synthesized following a well-characterised chemical reaction, therefore allowing a quick and robust subsequent exploration of the chemical space around them for the development of analogues and derivatives. This initial screening yielded 61 compounds identified as binders, which were then further tested to confirm Notum inhibition by a biochemical assay using octanoyloxypyrene-1,3,6-trisulfonate (OPTS). OPTS is a molecule acting as substrate for esterases and lipases, releasing a fluorescent signal upon hydrolysis[165]: the reduction in fluorescence in the presence of Notum inhibitors was used to measure the compound IC_{50} (i.e. the inhibitor concentration needed to inhibit 50% of the target activity).

The process led to the identification of six confirmed hit series amenable to further development: 4 however showed a plateau in potency, while two (based on a phenoxyamidic and triazolic chemical backbone) were selected for structure-based drug optimization (**Figure 5.3**). Subsequently, a computational virtual screening of a custom library of 4350 carboxylic acid fragments yielded 20 additional hits with $IC_{50} < 25$ μM, which were also subject to crystallographic analysis.

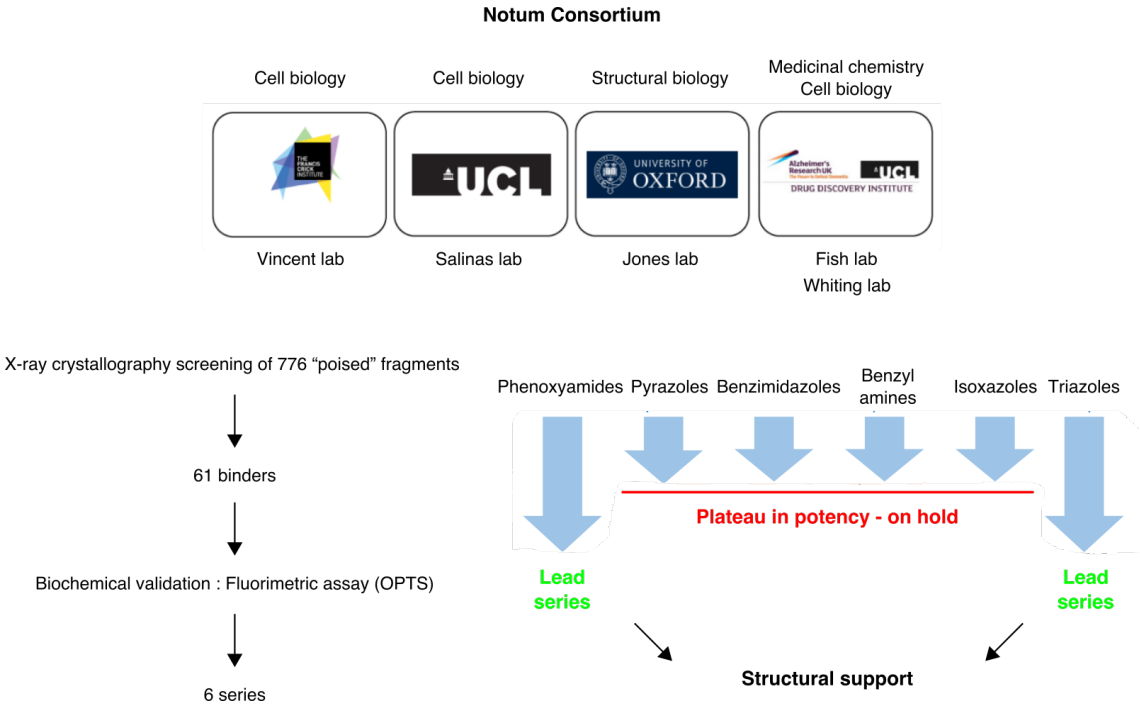


Figure 5.3. Overview of fragment screening and chemical series development for Notum inhibitors.

For each of these three series, the following paragraphs will outline the drug development effort presenting a subset of the most potent and representative examples of Notum structures in complex with inhibitors, out of the 67 compounds and 42 structures in whose screening I was personally involved at different stages of crystal growth, inhibitor soaking, diffraction data collection, structural data analysis and presentation, and out of the 14 crystal structures among them that I personally solved and refined (listed in Appendix **Table 7.1**).

Due to the very series-specific relationships existing between the chemical modifications designed and their effects on compound potency, it is difficult to isolate the description of the improvements in the structure-activity-relationship (SAR) from its crystallographic explanation, and reserve the latter to a distinct discussion chapter. For this reason, in the following paragraphs, the drug

optimization process and their rationale suggested by the x-ray structures will be presented together for each series. The subsequent final discussion section will then focus on the properties making our best compound a promising drug for Alzheimer's treatment, coupled with a broader analysis of the opportunities and drawbacks of pharmacologically targeting the Wnt pathway.

Phenoxyamides[§§§]

One of the main challenges in the development of suitable compounds for AD, and neurological targeting in general, is the capacity of the drug to penetrate the blood-brain-barrier, due to the very low permeability of the BBB to transcellular passive diffusion, as previously outlined.

In this sense, the 2-(2-methylphenoxy)-N-(pyridine-3-yl)acetamide scaffold from the initial fragment screening hit for the phenoxyamide series (**Figure 5.4 a**) appeared a very good candidate for further development, thanks to very good physicochemical properties already present in the non-functionalized initial hit (e.g. cLogP = 1.7, LLE = 2.8, CNS MPO = 5.6/6)[****].

[§§§] The work presented in this chapter has been published in the following article which I co-authored: ref. 166 Atkinson, B. N. *et al.* Discovery of 2-phenoxyacetamides as inhibitors of the Wnt-depalmitoleating enzyme NOTUM from an X-ray fragment screen. *Medchemcomm* **10**, 1361-1369, doi:10.1039/c9md00096h (2019). The medicinal chemistry studies have been performed by the laboratory of Prof. Paul Fish (UCL London) and the lab members cited therein. Some figure panels are derived from the article, as well as some sentences of the text.

[****] cLogP (the logarithm of the octanol/water partition coefficient) is a measure of drug lipophilicity and one of the main determinants of drug distribution into the brain, with cLogP values <2 suggested to be optimal. However, the parameter, particularly in the presence of highly hydrophobic binding pockets, should be evaluated together with other indicators including lipophilic ligand efficiency (LLE) – whose formula is pIC_{50} - cLogP, correlating lipophilicity per unit of *in vitro* potency - to rule out false positives in the optimization process showing increased potency as a simple result of increased lipophilicity. The latter situation is in fact detrimental for downstream pharmacology as it enhances the risk of off-target side effects. Moreover, given the peculiarities of the central nervous system (CNS), a multiparameter scoring system called CNS

Chemistry improvement of the drug in the hit-to-lead process was focused, in turn, on variation of the phenoxy ring, of the linker and of the amide heterocycles (see **Figure 5.4 a**).

Several substituents (e.g. Me, F, CF3, Cl, OMe, CN) were tested to modify the initial methyl group in the ortho- position 2 of the phenoxy ring, moving it to meta- or para- positions, however the potency profile of the inhibitors always favoured either the absence of any substituents, or the presence of a methyl moiety in ortho- as in the initial fragment hit, or anyway small lipophilic groups/atoms in that position (Cl, F, CF_3). The methyl moiety was kept in further drug optimization.

Regarding the linker, the SAR favoured the maintenance of the initial O-based linker (O > NH > S > CH2 preference in terms of potency).

With respect to the N-acetamide portion, the best combination of potency and physicochemical properties was instead obtained by using benzo-fused 6,5 and benzo-fused 6,6 heterocycles.

MPO (MultiParameter Optimization), which, among other scoring factors, includes cLogP, is also gaining traction, with a score of 6 ranked as highly desirable and a predictor of successful CNS drug design. See: ref. 167 Rankovic, Z. CNS drug design: balancing physicochemical properties for optimal brain exposure. *J Med Chem* **58**, 2584-2608, doi:10.1021/jm501535r (2015).

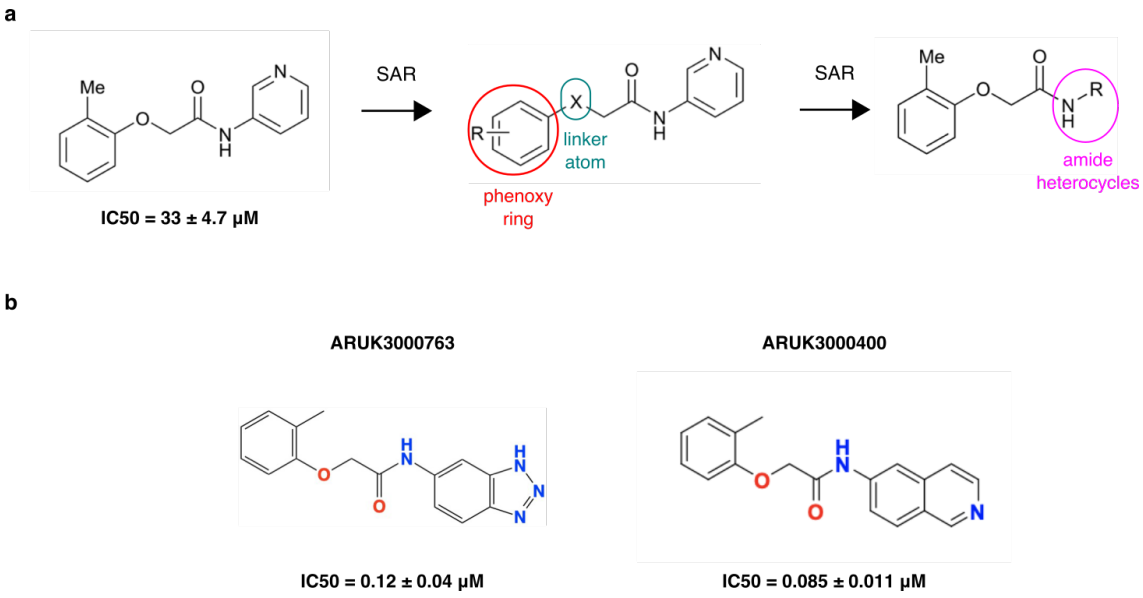


Figure 5.4. Phenoxyamide series medicinal chemistry optimization.
a) Initial fragment hit and SAR performed on different parts of the compound. **b)** Lead compounds developed for the series, whose crystal structures are presented. Panel a is adapted from ref. [166].

The outlined process led to the development of the benzotriazole ARUK3000763 (6,5 heterocycle) and the isoquinoline ARUK3000400 (6,6 heterocycle) as preferred compounds for the series (**Figure 5.4 b**), with a major increase in potency with respect to the initial hit from the fragment screening (12 ± 4 nM for the benzotriazole and 85 ± 11 nM for the isoquinoline, respectively, vs. 33 ± 4.7 µM of the initial hit fragment). Their crystal structures (**Figure 5.5**, PBD ID 6R8Q for the benzotriazole, 6R8R for the isoquinoline) recapitulate the binding mode of the series, providing a justification for many of the features observed during the drug optimization process. **Table 5.1** reports the data collection and refinement statistics for the two structures.

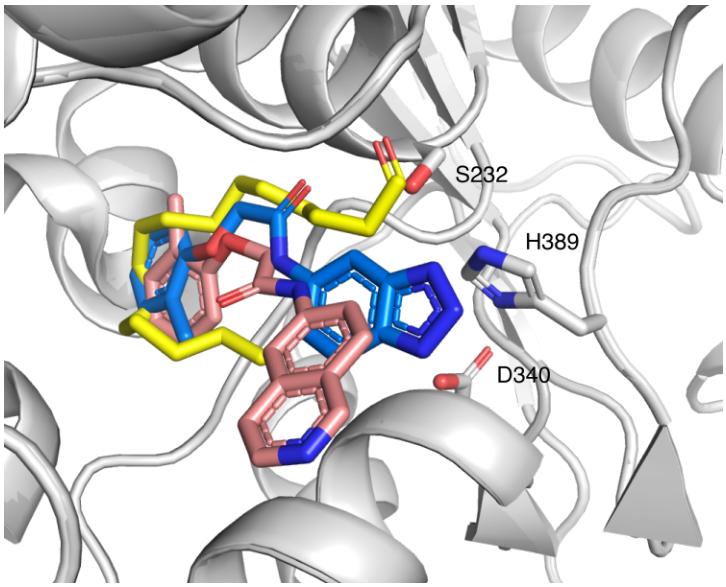


Figure 5.5. Overview of the structures of ARUK3000763 (marine) and ARUK3000400 (wheat).
To help orientation, the residues forming the catalytic triad are depicted as sticks. For comparison, the Wnt palmitoleate is also shown in yellow.

A key feature of both compounds is the formation of a strong set of hydrophobic and van der Waals interactions with many of the same residues that interact with the Wnt palmitoleate inside the cavity (e.g. Tyr129, Thr239, Pro287, Ile291, Phe319, Phe320, Val346), coupled to an optimally-oriented π-stacking with Phe268. A striking difference with respect to previously reported Notum structures, both apo and bound to the Wnt palmitoleate, is the additional π-stacking with Trp128, which relocates the residue relative to its native position by 4 Å in the case of the benzotriazole and by 6 Å in the case of the isoquinoline (**Figure 5.6 b** and **Figure 5.7 b**, respectively).

The abundance of hydrophobic interactions probably also explains why the initial fragment hit already presented good inhibitory properties, and the boost provided by the compound-Trp128 ring stack interactions gives a rationale for the added

potency provided by the fused benzo-rings with respect to single-ring heterocycles also tested during the optimization campaign. Potentially, the required twist of Trp128 from its apo position might also account for the generally low success rate of the soaking procedure that we observed for this series with respect to other series in the same experimental conditions: most series members have this additional energetic barrier to overcome in their binding, but the few that successfully rearrange Trp128 are then able to strongly exert their inhibitory effect with consequent high potency.

Both for the benzotriazole and the isoquinoline, the aromatic ring already fills much of the cavity (**Figure 5.6 a** and **Figure 5.7 a**), so that limited space between the compounds and the bottom of the shallow pocket is available to be occupied by substituents. This explains from a molecular point of view why functional groups located in meta- or in para- were poorly tolerated - due to the additional risk of clashes with residues lining the cavity - and why even the absence of substituents in the phenoxy ring didn't impact on compound potency compared with the presence of functional groups.

With respect to the SAR of the linker region, as the crystal structure of the benzotriazole complex shows, the oxygen is at an optimal distance and orientation to engage in an internal hydrogen bond with the linker amide group, keeping the compound in an energetically favourable conformation, suggesting why the O-based linker was preferred (**Figure 5.6 a**). Conversely, the isoquinoline loses this interaction as it undergoes a 180° rotation around the amide bond; however, in this case, the carbonyl oxygen of the linker is in a

favourable position to engage in polar interactions with Tyr129, possibly through a coordinated water (**Figure 5.7 c**).

Despite the described good chemical properties, unfortunately, when tested for metabolic stability by a microsomal stability assay the compounds showed rapid metabolization in a NADPH-independent manner. It was possible to overcome the issue by modification of the O-linker between the phenoxy ring and the N-acetamide parts, but not without a significant loss in potency, and this therefore stopped further development of this series.

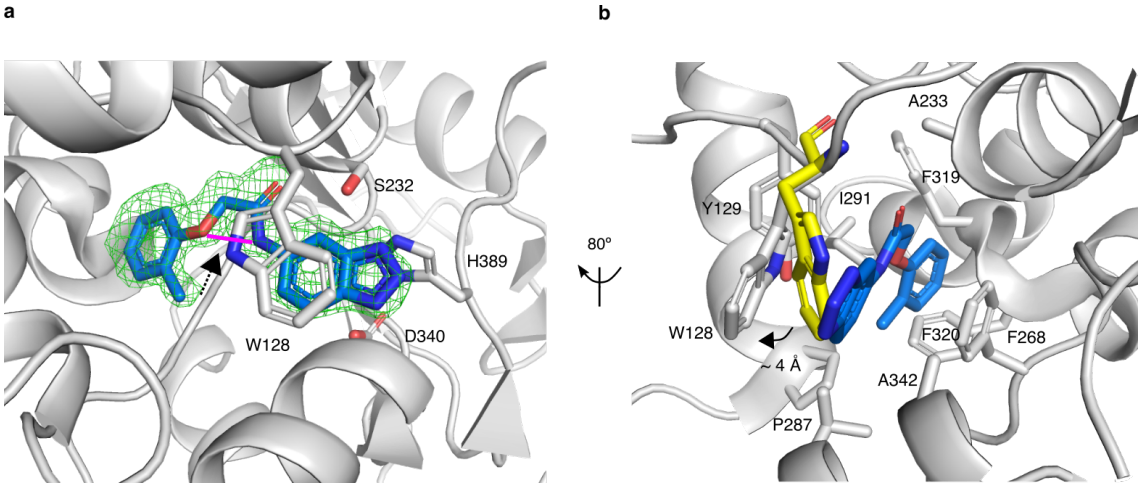


Figure 5.6. Structure of Notum in complex with ARUK3000763.
a) F_o-F_c omit map for ARUK3000763 and rationale for O-based linker preference in SAR studies. The omit map is contoured at ±2.5 σ. To help orientation, the residues forming the catalytic triad and W128 are shown as sticks. The purple line, indicated by the arrow, highlights the internal hydrogen bond between the O-atom linker and the amide group. **b)** Interactions of ARUK3000763 (marine) with the binding pocket and rearrangement of W128 with respect to the apo- and palmitoleate-bound conformation of Notum. The arrow shows the movement of W128 in the inhibitor structure (grey) vs. the original position of the tryptophan in the palmitoleate-bound structure (yellow). W128 of apo-Notum is not shown, as it occupies the same position of W128 of the palmitoleate-bound structure.

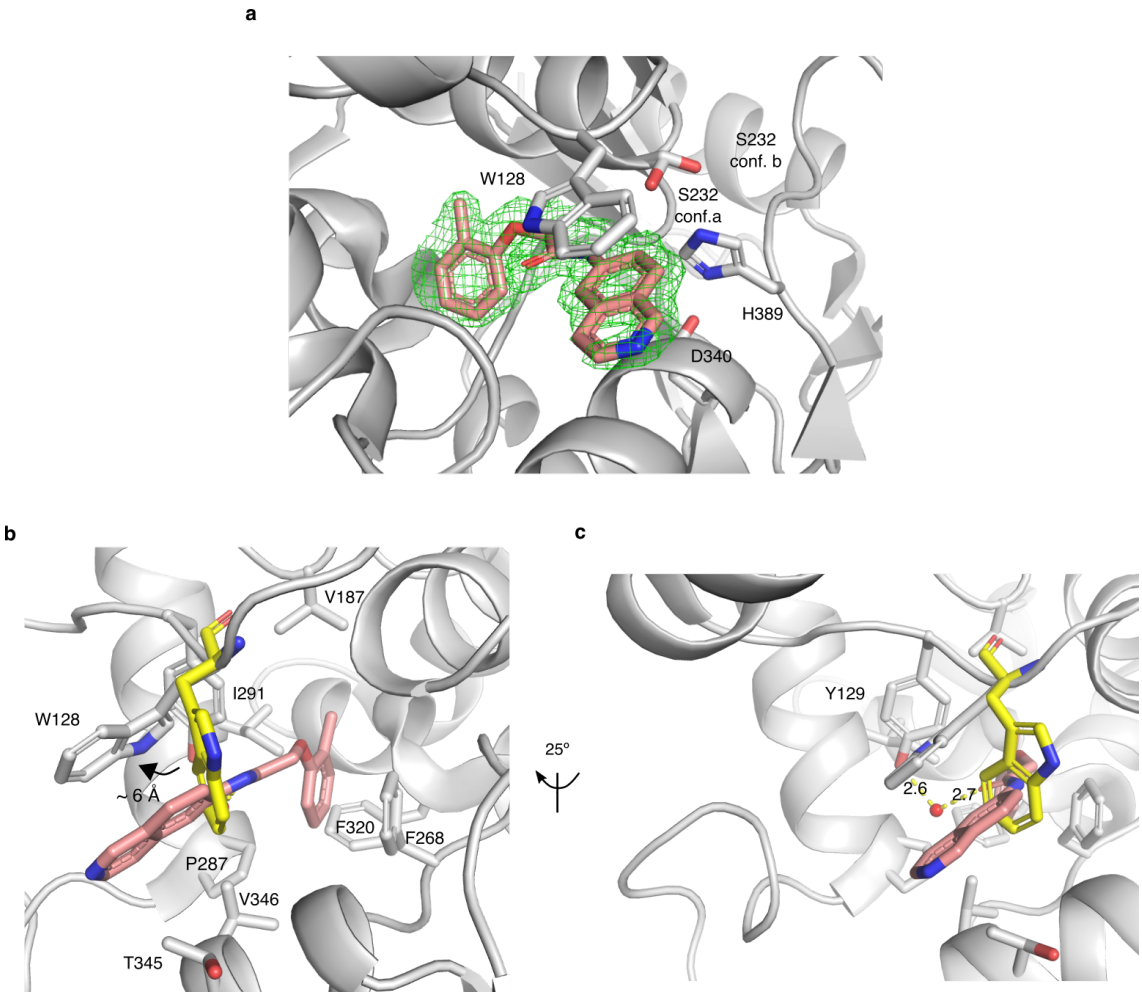


Figure 5.7. Structure of Notum in complex with ARUK3000400.
a) F_o-F_c omit map for ARUK3000400, contoured at ±2.5 σ. To help orientation, the residues forming the catalytic triad and W128 are shown as sticks. **b)** Interactions of ARUK3000400 (wheat) with the binding pocket and rearrangement of W128 with respect to the apo- and palmitoleate-bound conformation of Notum. The arrow shows the movement of W128 in the inhibitor structure (grey) vs. the original position of the tryptophan in the palmitoleate-bound structure (yellow). W128 of apo-Notum is not shown, as it occupies the same position of W128 of the palmitoleate-bound structure. **c)** Coordinated water (red sphere) favouring the interaction between Y129 and ARUK3000400 (wheat). The polar interactions are shown by yellow dashed lines, labelled with the distance in Å.

Table 5.1. Data collection and refinement statistics – phenoxyamides

	Notum_ ARUK3000763	Notum_ ARUK3000400
PDB ID code	6R8Q	6R8R
Ligand code	JV5	JV8
Data collection		
X-ray source (Diamond)	I24	I04-1
Wavelength (Å)	0.96863	0.91587
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	60.1, 71.3, 78.3	60.3, 71.7, 78.1
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	52.69 - 1.50 (1.52 - 1.50)	24.2 - 1.27 (1.29 - 1.27)
$I / \sigma I$	9.6 (1.7)	13.7 (1.1)
CC 1/2	0.998 (0.610)	0.999 (0.520)
Completeness (%)	99.9 (95.8)	100 (100)
Redundancy	12.5 (12.6)	13.0 (13.2)
Refinement		
Resolution (Å)	52.69 - 1.5	24.20 – 1.27
No. reflections	54513 (5288)	89931 (8890)
R_{work} / R_{free}	0.155/0.195	0.148/0.177
No. atoms		
Protein	2906	2968
Ligand/other	21/74	22/114
Water	202	184
B-factors (Å^2)		
All atoms	21	23
Ligand	19	28
R.m.s. deviations		
Bond lengths (Å)	0.008	0.013
Bond angles (°)	1.0	1.2

Values in parentheses are for highest-resolution shell.

Heteroaryl compounds††††

The second lead series explored was based on a triazolic hit (see **Figure 5.8**).

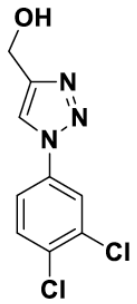


Figure 5.8. Initial hit from fragment screening for the heteroaryl series.

Modification of the triazolic portion to improve the initial IC_{50} was attempted with several different heteroaryl headgroups, that could be classified into 3 main classes: oxadiazoles (i.e. with two nitrogen substituents and one oxygen substituent in the heteroatomic ring), C-linked triazoles (i.e., with three nitrogen substituents in the five-member aromatic ring) and N-linked triazoles (i.e., with three nitrogen substituents in the five-member aromatic ring, one of which binds the phenyl moiety). Each class was then tested with further substituents of different size emerging from the headgroup (spanning from no substitution at all, to substitution with small groups, like -OH or -SH, up to much bigger functionalization with additional heteroaromatic rings).

†††† The results of the work presented in this chapter has been patented in the patent titled "Compounds" (patent application n. PCT/EP2019/073177, filing date 30.08.2019), of which I am a co-inventor. The medicinal chemistry and *in vivo* pharmacology studies have been performed by the laboratory of Prof. Paul Fish (UCL London) and the lab members cited therein.

In general, the structure-activity relationship analysis tended to favour smaller substituents emerging from the headgroup, with hydroxyoxadiazoles or aminooxadiazoles representing examples of subclasses of oxadiazoles reaching high potency levels (among those tested in structural studies, as high as 3 nM for hydroxyoxadiazoles and as high as 11 nM for aminooxadiazoles).

Here, in particular, three of the most potent aminooxadiazoles structures screened and solved will be presented (compounds ARUK3001213, ARUK3001233 and ARUK3001244). First, they recapitulate and are therefore well suited to summarize some general features of the binding mode of the heteroaryl series of compounds. Moreover, the high similarity of their chemical structure, apart from the heteroaryl substituent group, determines a very similar mode of binding in the pocket (see superposition in **Figure 5.9 b**), allowing me to dissect, rationalize and best exemplify the SAR observation that smaller groups are preferred as headgroup substituents. **Table 5.2** reports the data collection and refinement statistics for the three structures.

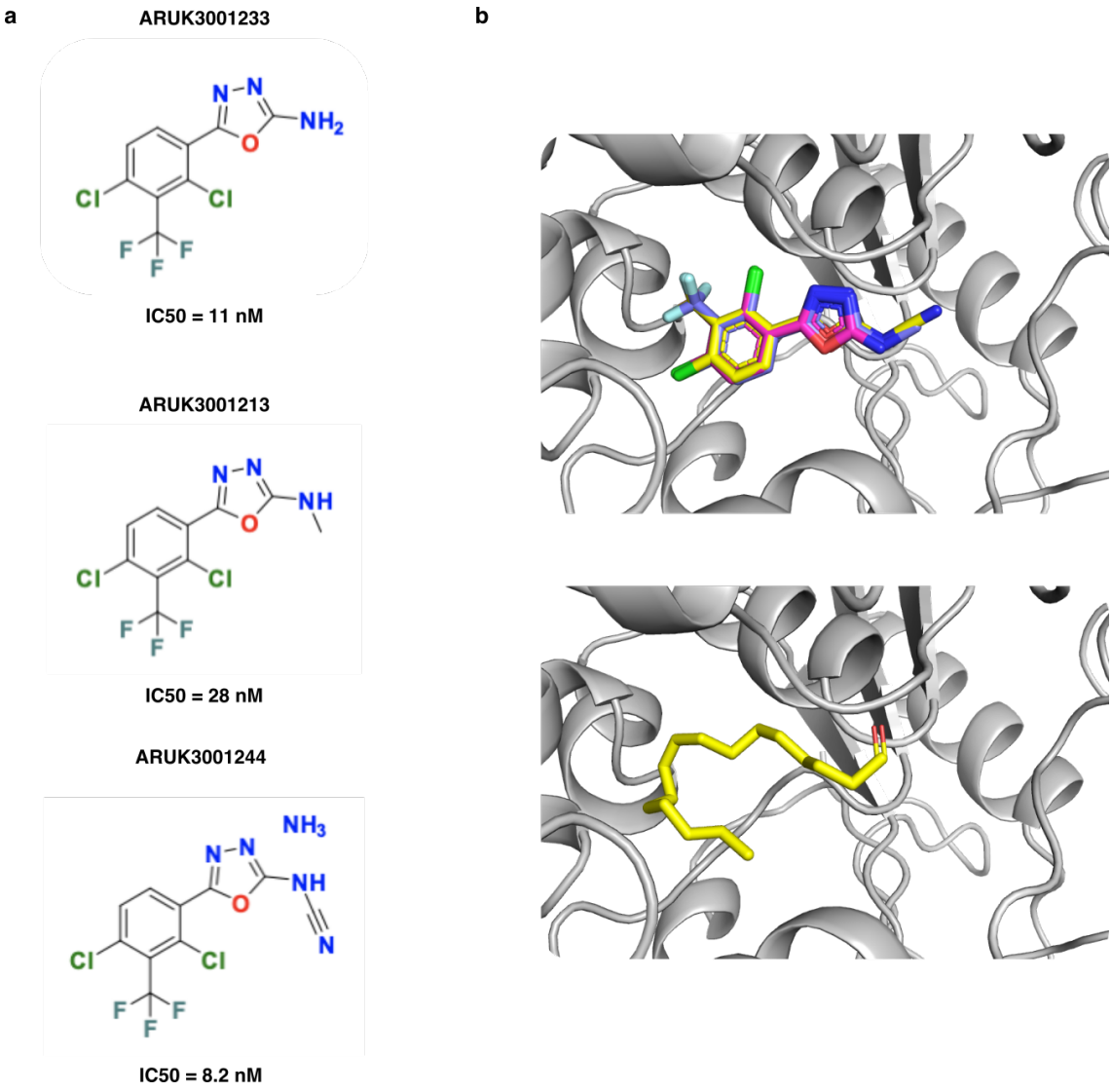


Figure 5.9. Overview of the aminooxadiazoles ARUK3001213, 1233 and 1244.
a) Chemical structures. **b)** Top panel: superposition of the crystal structures of the three inhibitors (ARUK3001213: slate; ARUK3001233: magenta; ARUK3001244: yellow) in complex with Notum. Bottom panel: for comparison, the Wnt palmitoleate in complex with Notum is shown in yellow.

Table 5.2. Data collection and refinement statistics – heteroaryl compounds

	Notum_ ARUK3001213	Notum_ ARUK3001233	Notum_ ARUK3001244
Data collection			
X-ray source (Diamond)	I04	I04	I04
Wavelength (Å)	0.97945	0.97945	0.97945
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	59.75, 71.43, 77.94	59.70, 71.74, 78.12	59.97, 72.11, 78.38
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	71.43 - 1.39 (1.41 - 1.39)	59.70 - 1.42 (1.44 - 1.42)	72.11 - 1.44 (1.46 - 1.44)
$I / \sigma I$	12.72 (1.04)	15.55 (1.10)	18.47 (1.01)
CC 1/2	0.9996 (0.6008)	0.9996 (0.6298)	0.9997 (0.4815)
Completeness (%)	100.00 (99.15)	99.95 (99.68)	99.07 (97.75)
Redundancy	12.47 (9.18)	13.0 (13.2)	13.16 (10.90)
Refinement			
Resolution (Å)	52.66 - 1.39	47.41 - 1.46	53.07 - 1.44
No. reflections	67567 (6592)	58655 (5738)	61534 (6001)
R_{work} / R_{free}	0.1778/0.2026	0.1665/0.2055	0.1900/0.2078
No. atoms			
Protein	2975	2896	2829
Ligand/other	19/37	18/56	20/19
Water	266	207	202
B-factors (Å^2)			
All atoms	28.49	32.02	33.86
Ligand	24.55	37.07	28.87
R.m.s. deviations			
Bond lengths (Å)	0.005	0.008	0.008
Bond angles (°)	0.82	1	1.03

Values in parentheses are for highest-resolution shell.

In terms of general features of binding, the compounds in this series tend to have in common strong interactions at the mouth of the lipid-binding cavity forced by the heteroatomic moiety: these are driven both by π-stacking with Trp128 and by hydrogen bonding with main-chain carboxyl groups, the catalytic Ser232 or coordinated waters, depending on the compound - the abundance of different possible interactions might also explain why both the triazole and oxadiazole heteroatomic scaffolds of this series could reach a high level of potency.

With respect instead to the SAR of the headgroup substituent and why smaller groups are preferred, the shorter substitution (unmodified aminooxadiazole, ARUK3001233, **Figure 5.10**) put the -NH2 group in a very favourable network of hydrogen bonding with two of the Notum catalytic residues (Ser232 and His389 through a coordinated water).

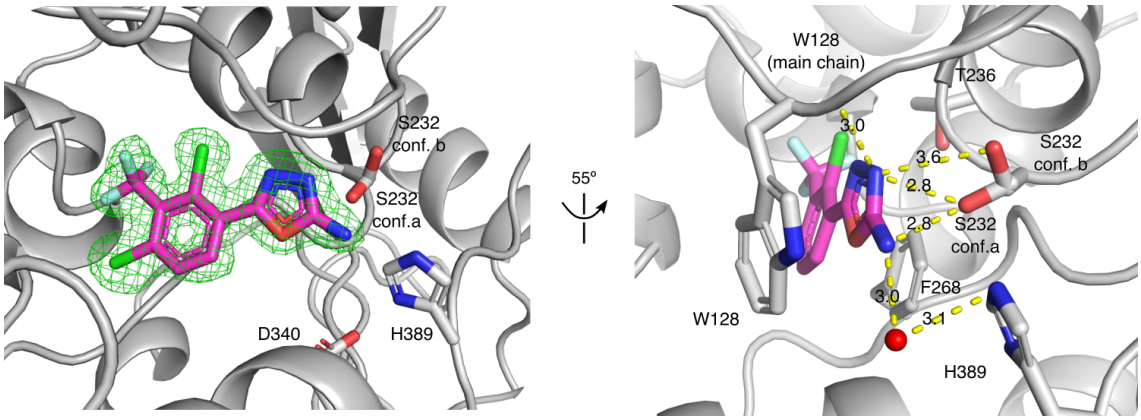


Figure 5.10. Structure of Notum in complex with ARUK3001233.
Left panel: F_o-F_c omit map for ARUK3001233, contoured at ±3 σ. To help orientation, the residues forming the catalytic triad are shown as sticks. Right panel: interactions of ARUK3001233 with the binding pocket. The polar interactions are shown by yellow dashed lines, labelled with the distance in Å. The red sphere represents a coordinated water.

Addition of a methyl group (ARUK3001213, **Figure 5.11**) still conserves a strong interaction with Trp128 (both in terms of π-stacking, and of a suitable distance for formation of hydrogen bonding with the main chain), but pushes away Ser232, at a suboptimal hydrogen bonding distance (3.5 Å); in addition, the indirect interaction with His389 is lost, potentially explaining the slight loss in potency.

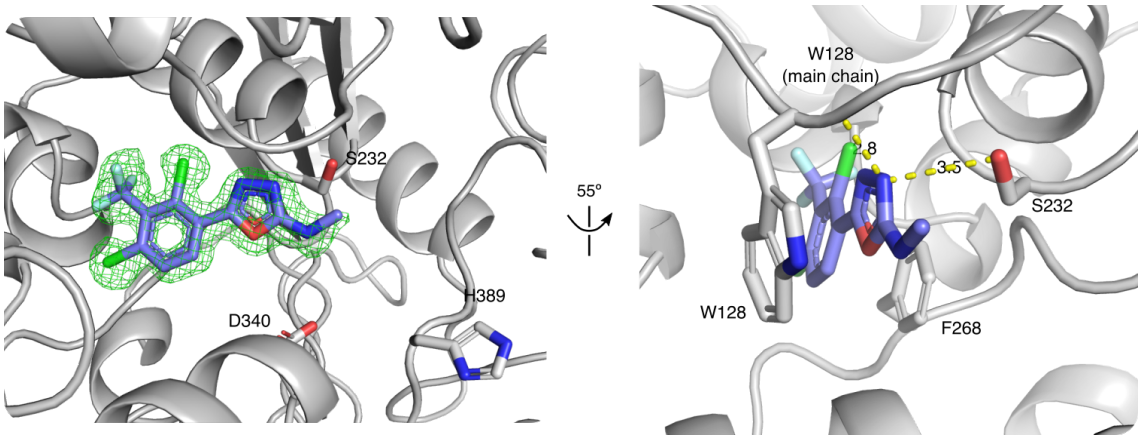


Figure 5.11. Structure of Notum in complex with ARUK3001213.
Left panel: F_o-F_c omit map for ARUK3001213, contoured at ±3 σ. To help orientation, the residues forming the catalytic triad are shown as sticks. Right panel: interactions of ARUK3001213 with the binding pocket. The polar interactions are shown by yellow dashed lines, labelled with the distance in Å.

Larger functionalization was almost always poorly tolerated, apart from some select situations exemplified by the cyanamide (ARUK3001244, **Figure 5.12**): in this case, even though the serine (Ser232) is still pushed away, the group regains the polar contact with His389.

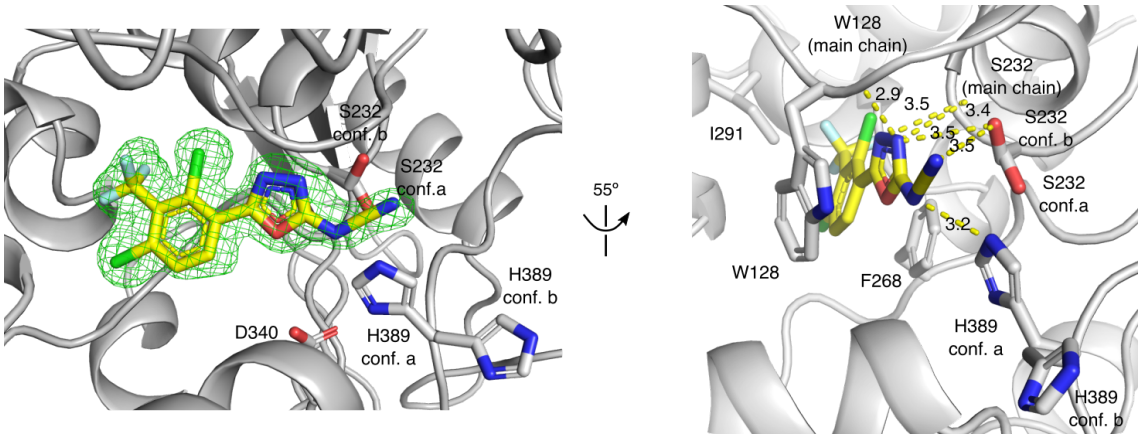


Figure 5.12. Structure of Notum in complex with ARUK3001244.
Left panel: F_o-F_c omit map for ARUK3001244, contoured at ±3 σ. To help orientation, the residues forming the catalytic triad are shown as sticks. Right panel: interactions of ARUK3001244 with the binding pocket. The polar interactions are shown by yellow dashed lines, labelled with the distance in Å.

However, the latter residue is located on a flexible loop, taking two conformations in this structure - indicative of variable positioning - and generally different orientations in the different Notum-inhibitor structures (**Figure 5.13**). This variability (reflected in the difficulty to design large interacting groups tolerated without significant loss in potency) might explain the limited success of large functionalization of the headgroup. A further complication in this sense relates to the fact that the mouth of the binding pocket - though flexible and able to adapt to the different inhibitors - is anyway narrow and therefore likely to incur steric clashes with bigger groups (as shown for the loss of optimal Ser232 orientation for ARUK3001244 and ARUK3001213 with respect to the smaller ARUK3001233). Indeed, first, the flexibility of the loop makes it difficult to devise strategies to clamp it effectively; second, a gain in interactions outside the pocket might come at the price of loss of key interactions with the catalytic residues for steric reasons.

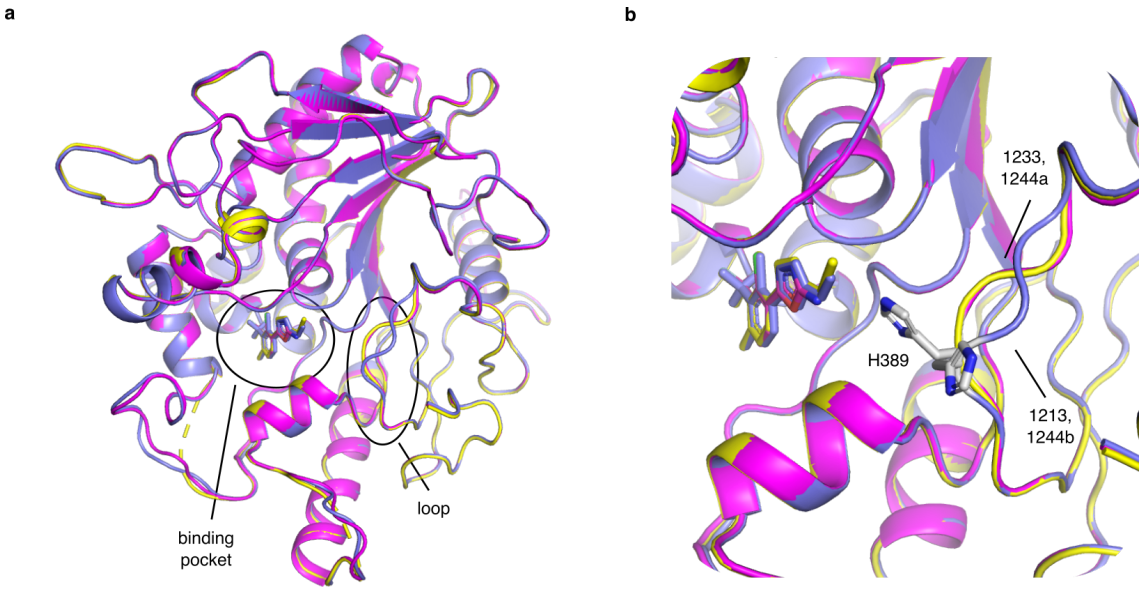


Figure 5.13. Variable positioning of the β11-β12 loop containing the catalytic H389.
a) Localization of the loop in the Notum structure with respect to the binding pocket. **b)** Close-up view and superposition between the structures of ARUK3001213 (slate), ARUK3001233 (magenta) and ARUK3001244 (yellow). The two alternative conformations taken by H389 in these structures are also shown. Numbering and nomenclature of the loop is based on ref. [85].

With respect to the non-heteroatomic portion inside the binding pocket, this series presented a lower number of interactions with cavity residues compared to the phenoxyamides, with only a conserved π-stacking with Phe268 common to most compounds, and further weak, compound-dependent, van der Waals interactions - allowing a certain degree of flexibility to accommodate even small substitutions in para-, as in the case of the shown aminooxadiazoles.

Maintenance of an optimal orientation for π-stacking with Phe268 was therefore the most relevant factor to take into account for the optimization of the part of the compounds sitting inside the binding pocket. Moreover, this constraint might also represent an additional element contributing to the lower tolerability of large functionalization at the headgroup, since, for some compounds tested, the gain

of interactions at the cavity mouth came at the cost of a tilt of the phenyl ring and hence a non-optimal stacking with Phe268.

The combination of these structural insights with physicochemical considerations aimed at improving the penetration into the blood-brain-barrier (e.g. acting on the drug charge) allowed our Consortium to develop our lead compound, ARUK3001185 (structure not shown for intellectual property reasons), starting from a triazole scaffold. The drug-like molecule has high potency (IC_{50} = 6.5 nM), optimal physicochemical properties (e.g. CNS MPO = 5.6/6), good solubility, permeability and metabolic stability, and, above all, when orally administered to mice, an excellent central nervous system penetration (**Figure 5.14**).

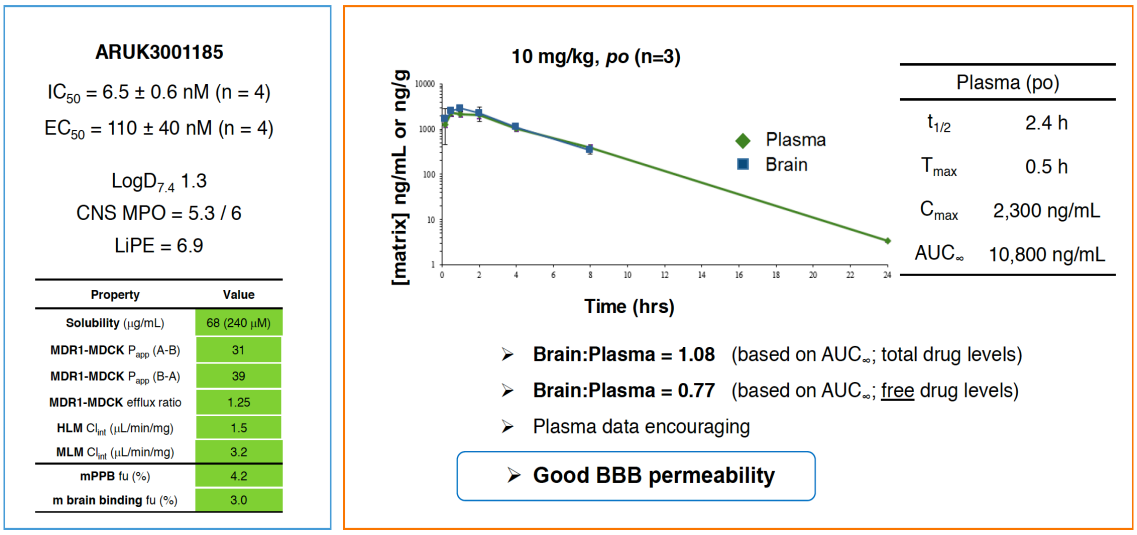


Figure 5.14. Physicochemical and *in vivo* mouse pharmacokinetics properties of lead compound ARUK3001185.
EC_{50} describes the half-maximal effective concentration of the compound inducing Wnt signalling reactivation in luciferase-based TCF/LEF cellular assays. $LogD_{7.4}$ represents the octanol/water distribution coefficient for ionizable compounds at pH 7.4. LiPE is equivalent to LLE (lipophilic ligand efficiency) defined above. The MDR1-MDCK is a permeability assay using Madin Darby canine kidney (MDCK) cells overexpressing the *MDR1* gene, encoding for the multidrug transporter P-glycoprotein1, and evaluates whether the compound is a P-glycoprotein substrate. HLM and MLM are microsomal metabolic stability tests.

mPPB and m. brain binding represent the percentage of drug bound to plasma and brain proteins, respectively. $t_{1/2}$ is the compound half-life upon *in vivo* administration, T_{max} is the time at which the maximum plasma concentration is achieved, C_{max} is the maximum plasma concentration observed, and AUC_{∞} (Area Under Curve) indicates the total drug exposure across time. (Data and figure reproduced as a courtesy of Prof. Paul Fish, UCL London).

Carboxylic acid compounds‡‡‡‡

As noted before, in addition to the aforementioned fragment screening performed using the DSI-poised Diamond Library, a library of custom acid fragments was also screened, first virtually and the best hits then validated experimentally. The goal was, on the one hand, to provide a backup series of compounds with high potency in addition to the patented lead series, on the other hand to identify alternative scaffolds specific and effective against the target, with good CNS penetration. Indeed, the identification of structurally diverse chemotypes all able to inhibit Notum but with different secondary pharmacology (i.e. with different not target-related effects) could represent a powerful tool in downstream target validation and off-target effects identification.

The screening identified in particular pyrrole-3-carboxylic acids and pyrrolidine-3-carboxylic acids as scaffolds able to efficiently inhibit Notum: these were used as a starting point for further chemical optimization (**Figure 5.15 a**).

‡‡‡‡ The work presented in this chapter has been published in the following article which I co-authored: ref. 168 Mahy, W. *et al.* Screening of a Custom-Designed Acid Fragment Library Identifies 1-Phenylpyrroles and 1-Phenylpyrrolidines as Inhibitors of Notum Carboxylesterase Activity. *J Med Chem* **63**, 9464-9483, doi:10.1021/acs.jmedchem.0c00660 (2020).. The medicinal chemistry studies have been performed by the laboratory of Prof. Paul Fish (UCL London) and the lab members cited therein. Some figure panels are derived from the article, as well as some sentences of the text.

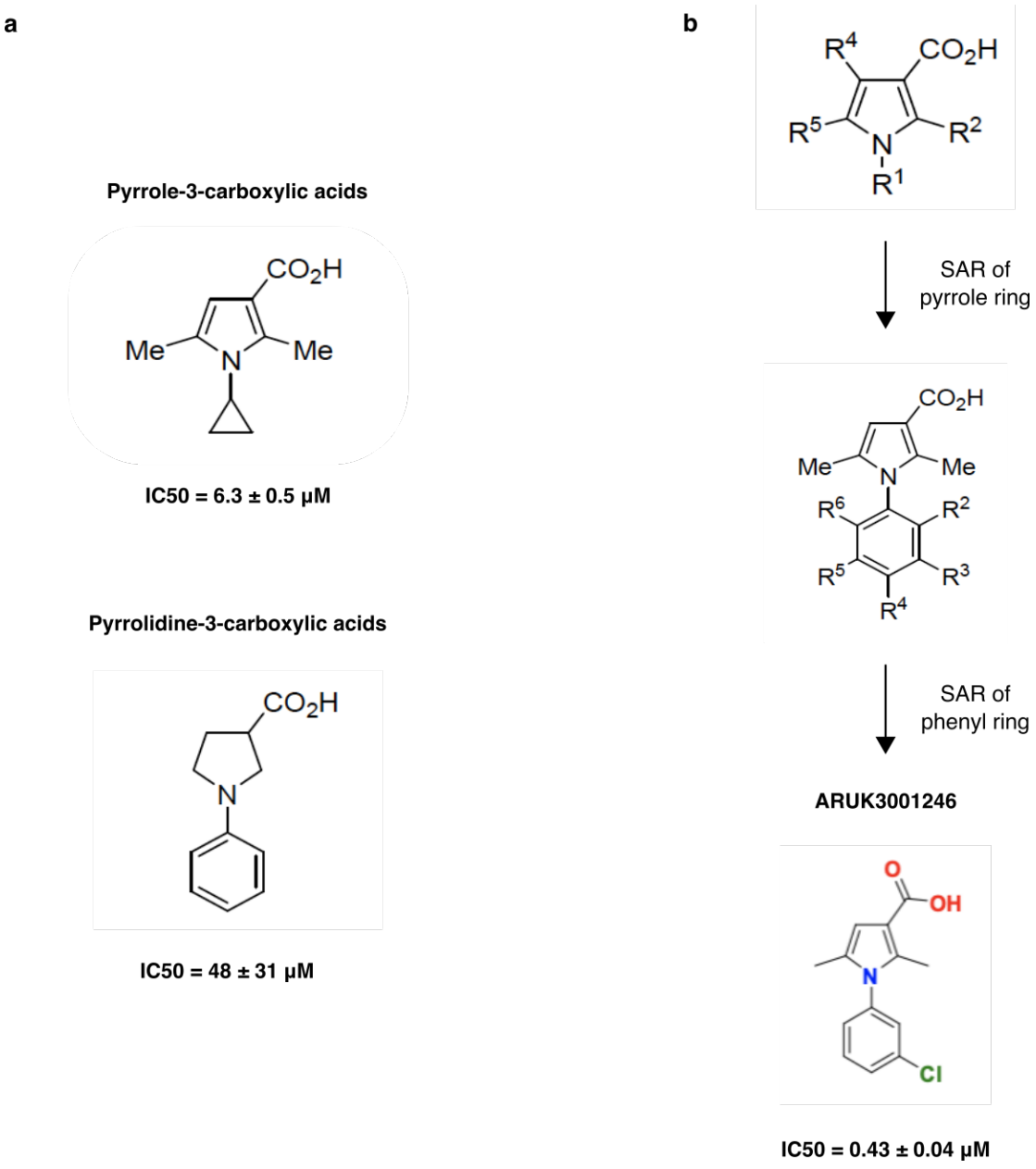


Figure 5.15. Carboxylic acid hits and pyrrole-3 scaffold medicinal chemistry optimization.
a) Initial hits. **b)** SAR performed on the pyrrole-3-carboxylic acid scaffold and lead compounds developed for the series, whose crystal structure is presented and discussed. Panel a and part of panel b are adapted from ref. [168].

The attention initially focused on the pyrrole-based compounds due to their higher inhibitory activity of Notum. First, modifications of the pyrrole ring were tested,

identifying 2,5-dimethyl-1-phenylpyrrole as the preferred substituent. Second, decoration of the phenyl moiety was performed, with the best hit in terms of potency being a Cl atom in position 3 (meta-) of the phenyl ring (**Figure 5.15 b**), although even a 4-Cl (para-) was well tolerated. This is in line with what was observed in the previous series: para- is generally not favoured because of the shallowness of the binding pocket causing steric clashes with the back of the cavity (see phenoxyamides), but small atoms in para- can be tolerated (see aminooxadiazoles). This led to the definition of the lead compound ARUK3001246, whose crystal structure bound to Notum was solved (PDB ID: 6YXI, **Figure 5.16**). **Table 5.3** reports the data collection and refinement statistics for the structure.

Superposition with the aminooxadiazoles presented in the previous paragraph shows how the compound sits deeper in the binding pocket, in line with the strong lipophilicity of its chemical structure, even pushing further out α-helix E, with a shift of around 3 Å at its N-terminal Pro287, compared for example to ARUK3001213 (**Figure 5.16 b** and **c**, secondary structure nomenclature according to ref. [85], see also **Figure 1.8 b**). Moreover, interestingly, this also causes the nearby β7-αE loop to be completely disordered in this crystal structure (**Figure 5.16 c**): the loop is part of what forms the "lid domain" of the α/β hydrolase fold[169], and might be involved in sealing off the binding pocket (being for example in a "closed conformation" in the Wnt lipid-bound Notum structure). It might be speculated that interfering with the loop movement might be another mechanism of inhibition of ARUK3001246 – although in general this a pretty flexible and difficult to build loop in different classes of Notum-binding inhibitors and binders

(see for example also the published structures of Notum-bound melatonin and melatonin-related molecules[170]), and it can't therefore be excluded that this is an artefact of the soaking procedure rather than a specific feature of this compound. With respect to cavity occupancy, the drug occupies a small volume, explaining why, as already noted above, the phenyl ring could also accommodate substitution in para without significant loss of potency.

The interaction of the compound with the binding pocket is mainly driven by hydrophobic forces; the acidic moiety, which is positioned deep within the pocket, does not engage directly with the residues at the mouth of the cavity, but only indirectly through a coordinated water (**Figure 5.16 d**). This provides an explanation for the observed SAR of these compounds, which consistently tracked with increased lipophilicity. In fact, the most active inhibitor of this chemotype also presented an undesirably high cLogP (5.5).

Table 5.3. Data collection and refinement statistics – carboxylic acids

	Notum_ ARUK3001246
PDB ID code	6YXI
Ligand code	PZ8
Data collection	
X-ray source (Diamond)	I04
Wavelength (Å)	0.979499
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	59.52, 72.16, 77.93
α, β, γ (°)	90, 90, 90
Resolution (Å)	59.52 - 1.34 (1.36-1.34)
$I / \sigma I$	17.6 (1)
CC 1/2	1.000 (0.555)
Completeness (%)	99.97 (99.54)
Redundancy	12.1 (6.3)
Refinement	
Resolution (Å)	52.95 - 1.34 (1.36 - 1.34)
No. reflections	75868 (2742)
R_{work} / R_{free}	0.159/0.185
No. atoms	
Protein	2920
Ligand/other	17/61
Water	198
B-factors (Å^2)	
All atoms	29.51
Ligand	26.57
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.79

Values in parentheses are for highest-resolution shell.

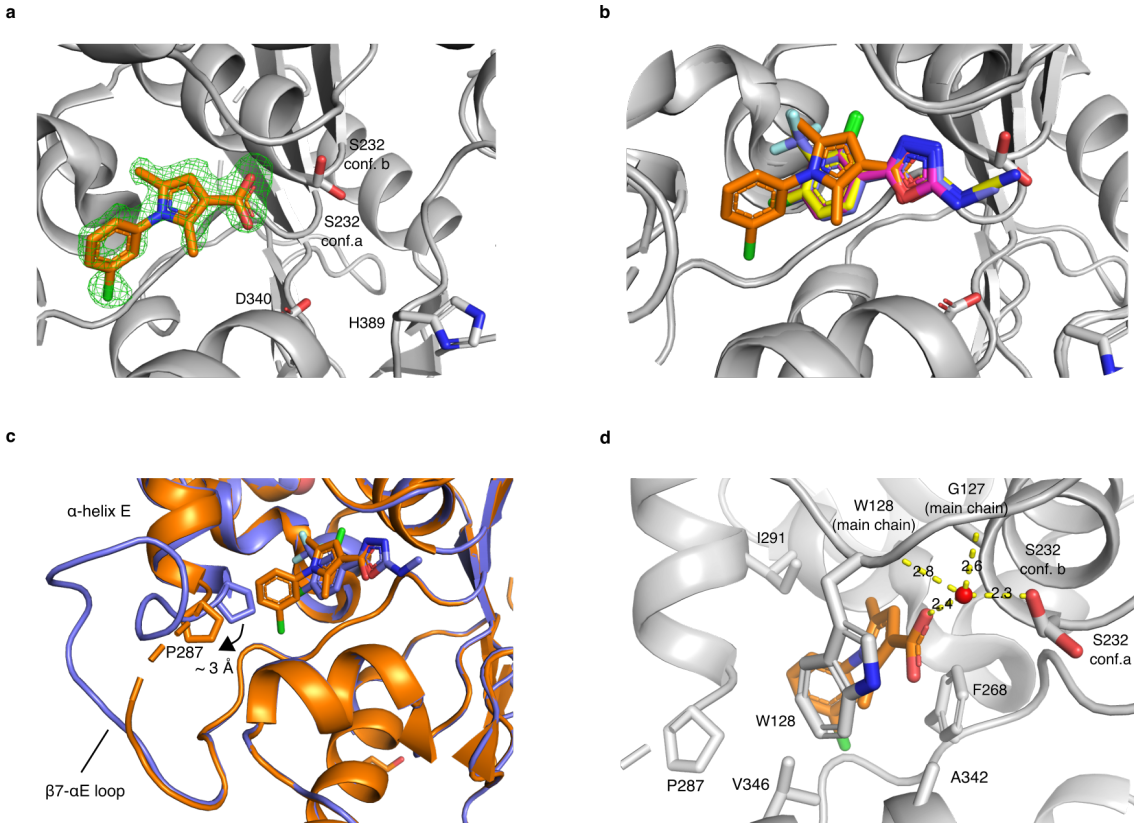


Figure 5.16. Structure of Notum in complex with ARUK3001246.
a) F_o-F_c omit map for ARUK3001246, contoured at ±3 σ. To help orientation, the residues forming the catalytic triad are shown as sticks. **b)** Superposition of ARUK3001246 (orange) with the three aminooxadiazoles, presented in the previous paragraph, ARUK3001213 (slate), ARUK3001233 (magenta) and ARUK3001244 (yellow). **c)** Shift of P287 (arrow) and disorder of the β7-αE loop caused by the positioning of ARUK3001246 (orange) deep into the Notum binding pocket, compared with ARUK3001213 (slate). **d)** Interactions of ARUK3001246 with the binding pocket. The polar interactions are shown by yellow dashed lines, labelled with the distance in Å.

For this reason, the pyrrole scaffold was abandoned; we preferred instead the pyrrolidine-3-carboxylic acids for further optimization. The process led to compound ARUK3003615, whose structure (solved by Dr Yuguang Zhao) is presented as a comparison to ARUK3001246 (**Figure 5.17**). The pyrrolidine-3-carboxylic acid ARUK3003615 does not sit as deep as the pyrrole-3-carboxylic acid ARUK3001246 in the cavity; this allows a better, full occupation of the cavity

and orients the compound in a very favourable geometry to engage the oxyanion hole, explaining its high potency (IC_{50} 110 nM). ARUK3003615 also showed good chemical properties (cLogP = 3.3, LLE = 5.4), metabolic stability, good CNS permeability, and high solubility.

In summary, this part of the work provided an additional set of compounds with high potency and good pharmacological properties to use as further drug-like molecules in Notum inhibition studies, although with an inhibitory effect not as high as the lead series discussed above.

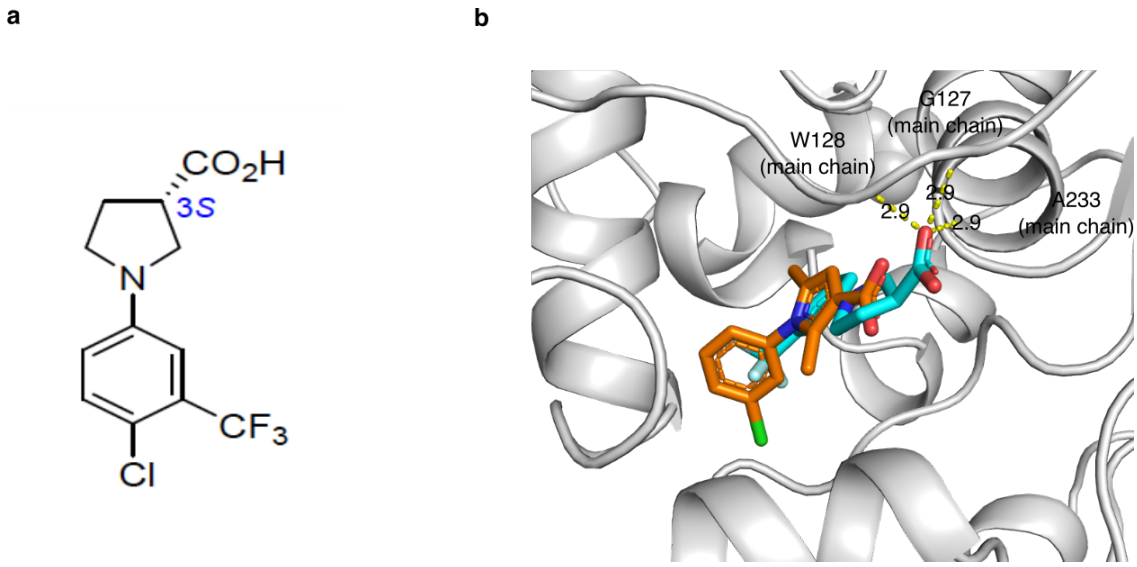


Figure 5.17. Comparison between the pyrrole-3 and pyrrolidine-3 carboxylic acids.
a) Chemical structure of the most potent pyrrolidine-3 carboxylic acid (ARUK3003615). **b)** Superposition of the structure in complex with Notum of ARUK3003615 (cyan, solved by Dr Yuguang Zhao, PDB ID: 6YSK) with the lead pyrrole-3 carboxylic acid ARUK3001246 (orange), showing how ARUK3003615 is better positioned to engage atoms forming the oxyanion hole during Notum enzymatic activity. The polar interactions are shown by yellow dashed lines, labelled with the distance in Å. Panel a is derived from ref. [168].

Discussion

This chapter has summarized the work performed to develop initial chemical fragments binding the Wnt extracellular inhibitor Notum into drug-like molecules with high potency. In particular, three main series occupying pretty different parts of chemical space and presenting diversified modes of target engagement have been optimized to reach potencies in the low nanomolar range.

One of these series is based on a heteroaryl nitrogen-based headgroup forming a set of polar interactions with the enzyme catalytic triad, linked to a substituted phenyl ring responsible for hydrophobic interactions with the Notum lipid-binding cavity. This series yielded a compound that also fulfils all the prerequisites for a drug candidate for Alzheimer's disease, in terms of metabolic stability, specificity and pharmacokinetics (ability to cross the blood-brain-barrier).

Furthermore, this lead candidate presents favourable properties in comparison to the two other Notum inhibitors currently available in the literature (LP-922056, developed by the Lexicon Pharmaceuticals US company as a potential drug candidate for osteoporosis[171], and ABC-99[172]). LP-922056 is in fact not designed to target the brain and, despite a good pharmacokinetic profile, with good bioavailability and low clearance, presents a carboxylic acid on its headgroup preventing the drug from penetration into the neural sites. ABC-99 is a potent Notum inhibitor that covalently binds to the catalytic Ser232. Although the advantages of covalent vs. non-covalent enzyme inhibition are a disputed topic, covalent inhibitors have generally been less favoured because of concerns related to their safety (either due to the formation of toxic, non-specific species

upon metabolic activation, or due to their longer-lasting drug effect because of their covalent target engagement, with the subsequent risk of a too sustained activity)[173,174]. For such a delicately balanced pathway like Wnt, which must be maintained in a tightly controlled homeostatic window, a more regulatable non-covalent inhibitor seems a preferable option for an *in vivo* setting. Moreover, because of their different pharmacological mechanism of action, covalent and non-covalent inhibitors actually do not represent alternatives, but complementary strategies that can be deployed to target Notum.

The fact that Notum could be inhibited with a range of compounds occupying essentially orthogonal chemical spaces, as shown by our work, also confirms its attractiveness as a potential pharmacological target.

The next step in the project, together with the *in vivo* testing of the lead compound in mouse models of Alzheimer's, will be the dissection of the specific mechanisms of Notum involvement in the disease pathology. To this goal, effort is ongoing in our Consortium to identify the pattern of Notum expression in wild-type and AD mouse models, and in the human brain. Preliminary results seem to point towards Notum expression being present even at physiological conditions and being mainly localized to the endothelial cells **(Figure 5.18 a**). Crucially – as assessed by comparison of Notum mRNA levels in age-matched healthy, prodromal and Alzheimer's brain – the expression is upregulated in Alzheimer's disease, both in the mouse APP/PS1 model and in humans with clinically manifest disease with dementia and amyloid plaques, confirming a potential direct role of Notum hyperactivation in the natural history of the disease (**Figure 5.18 b**). The localization of the expression at the endothelial cell level also suggests that

Notum might exert its effect mainly by dysregulating the maintenance of the blood-brain-barrier, in line with the literature presented in the introduction suggesting a particularly important function of the Wnt pathway in the BBB homeostasis.

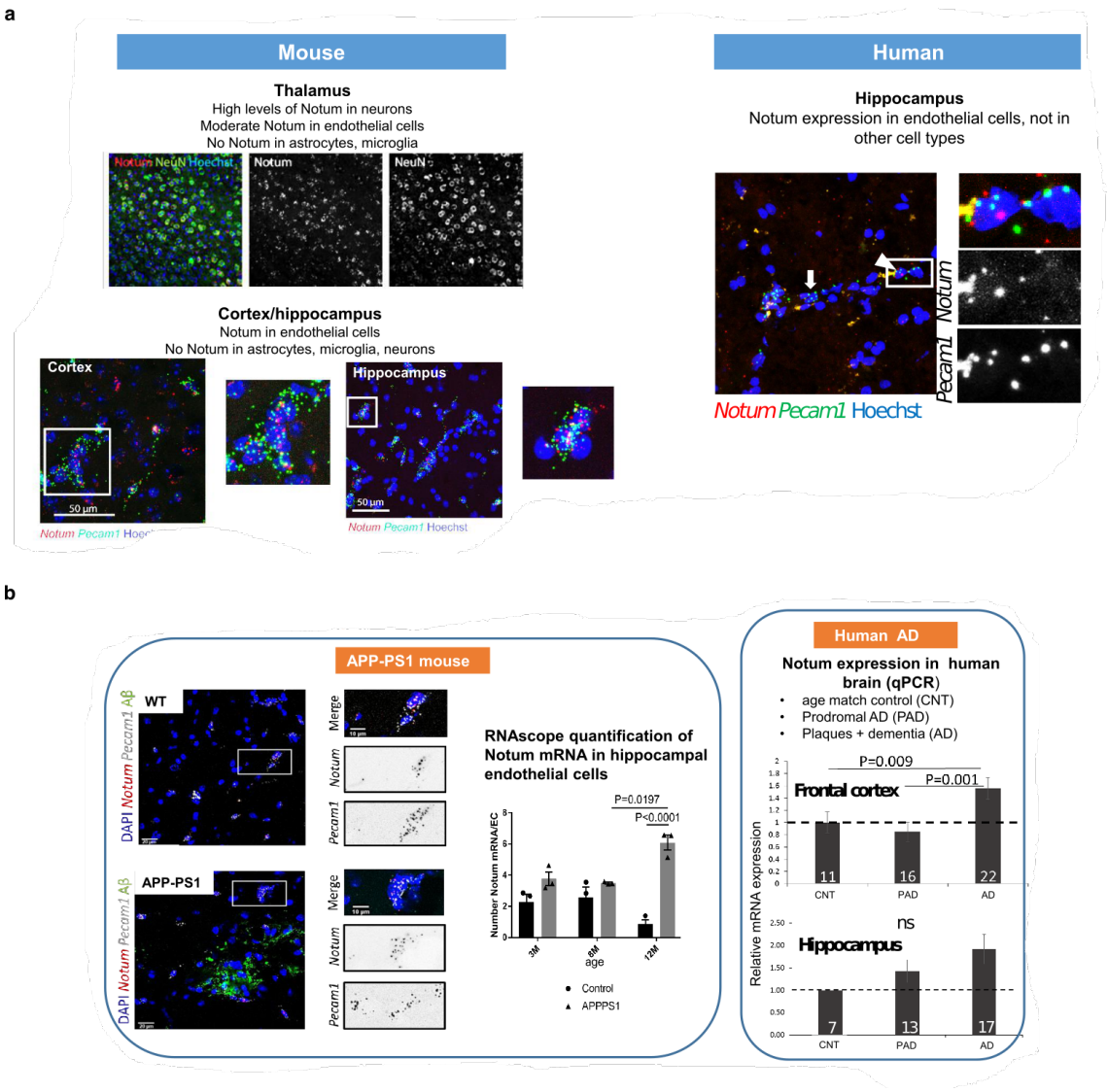


Figure 5.18. Notum expression in mouse and human brain in physiological and pathological conditions.
a) Normal mouse and human brain Notum expression. **b)** Notum expression in the APP-PS1 AD mouse model (see above) and in human AD samples. (Data and figure reproduced as a courtesy of Dr Sarah Jolly and colleagues, UCL London).

Further studies to validate this role are already being pursued in our Consortium, both by using *in vitro* cellular models of the BBB, and by generating brain Notum knockout mice to analyse the modifications in the synapse network and in the cerebral vasculature.

One obvious caveat about Notum targeting - as well as about targeting the Wnt pathway in general - relates to the potential side effects that might arise by altering such a ubiquitous pathway central to stem cell niche maintenance and to regenerative processes.

Important unwanted events, particularly at the level of the tissues with the highest homeostatic turnover (bone and intestine), have in fact been reported in almost all preclinical and clinical trials of Wnt drugs conducted so far. In some cases (e.g. tankyrase inhibitors - tankyrase interacts with and leads to Axin degradation[175]) they have resulted in dose-limiting toxicities[176], whereas in other cases they have led to temporary stopping of clinical trials (e.g. for the anti-Frizzled antibody vantictumab and the Wnt ligand decoy receptor ipafricept[177]). It is therefore likely that Notum inhibitors could face similar problems, with the additional caveat that boosting rather than inhibiting the Wnt pathway might result in even more undesirable hyperproliferative effects.

On the other hand, pre-clinical and clinical studies with inhibitors of Porcupine in multiple cancers have instead shown that therapeutic doses can be achieved without the appearance of significant adverse events, and the two aforementioned temporarily stopped phase 1 clinical trials with vantictumab and ipafricept were successfully restarted after implementation of bone loss mitigation strategies in the therapeutic schedule[177]. Moreover, conditional, liver-specific,

Notum knockout in adult mice has shown that Wnt signalling is increased but still within physiological ranges[178], raising hopes that a similar effect might be obtained in the brain with Notum inhibitors.

Even more interestingly, a recent study has uncovered a mechanistic link between aged intestinal epithelium and increased Notum expression[179]. In particular, it has shown that aged Paneth cells (surrounding stem cells in the intestinal crypts) constitutively overexpress Notum, and this contributes to a general decrease of Wnt signalling in the stem cell niche and a decline in the regenerative potential of the intestinal lining. Moreover, in the same study, treatment of mice showing mucositis caused by 5-fluorouracil (a first-line chemotherapy in human metastatic colorectal cancer) with the Notum covalent inhibitor ABC-99 led to recovery from the weight loss caused by the chemotherapy, without adverse effects caused by the Notum inhibition itself.

This has a double relevance in the context of the current discussion about the potential of Notum inhibition as an effective therapy. First, it reinforces the idea that Notum could be safely targeted; second, it also opens new opportunities for the use of Notum inhibitors in contexts additional to Alzheimer's disease. In relation to the mentioned study, they could for example be deployed to mitigate the side effects of conventional cancer chemotherapy, in situations where impaired Wnt signalling is implicated. Moreover, they might also find application in regenerative medicine contexts, for instance in hair follicle regeneration[180].

Of course, some properties of our lead series that are useful when targeting neurodegeneration - like the ability to easily cross the brain-blood-barrier - might become undesirable when the focus of the therapeutic action is not the brain.

However, the variety of highly potent chemotypes provided by our work, and the crystallographic information supporting and rationalizing their structure-activity relationships, create a powerful platform for further adaptation - e.g. by variation in the drug formulation, or by further tailored addition of chemical substituents based on the information gathered from the available Notum-inhibitor crystal structures.

Chapter 6 - General discussion, future perspectives and concluding remarks

The Wnt pathway represents a key embryonic morphogen signal, and a crucial regulator of tissue homeostasis and regeneration in the adult; dysregulation of its control leads to multiple degenerative and proliferative abnormalities, including several types of cancers. At the molecular level, Wnts are lipid-modified proteins, and this modification is crucial for receptor engagement at the target cells; however, coupled with a complex folding, it hampers protein solubility during Wnt intracellular and extracellular movement upstream of receptor activation.

What are the molecular mechanisms of lipid engagement that help Wnt overcome this hurdle, ensuring at the same time a regulated pathway activation? What are their structural determinants? And can we act on these mechanisms to open further therapeutic avenues?

My D.Phil. research tackled these questions by focusing on two main aspects: identifying the molecular basis and key interacting partners for lipid shielding during Wnt intracellular and extracellular trafficking; and targeting the extracellular regulator of its lipidation, i.e. the Wnt-dedicated deacylase Notum.

Towards the first goal, I have uncovered a novel molecular module of interaction that can provide a first clue to explain how this “Wnt solubility problem” is biologically solved. This solution relies on the lipid-binding ability of a subset of heparan sulphate-modified, GPI-anchored proteoglycans, called glypicans. As visible in the 2.2 Å resolution crystal structure that I solved of the *Drosophila* glypican Dlp in complex with a Wnt palmitoleoylated peptide, a conformational

rearrangement, mediated by a 6 Å tilting of an α-helix, opens a deep binding pocket for the lipid in the so-called glypican C-lobe. The structural findings were further confirmed by mutational analysis *in vitro* and *in vivo*, and expanded to human glypicans, to show how also Gpc4 and Gpc6, but not Gpc3 and Gpc5, possess this lipid binding capability, and the potential to act as morphogen reservoirs.

The work on Notum built on the positive hits from a previously performed x-ray fragment screen that provided the basis for a drug development campaign for small-molecule inhibitors of the enzyme. Three series in particular were optimized, yielding high potency candidates with diversified chemical structures. One of them, based on a triazole scaffold, also showed an interesting pharmacology in preliminary assays, and has become our lead drug for downstream *in vivo* testing.

In the following sections, I will discuss how these results fit in the broader picture of Wnt trafficking upstream of receptor activation and I will outline some far-reaching outstanding questions that remain to answered and can be addressed by future work building on the premises presented here.

Handover of Wnt

The work on glypicans has proven the role, among other models proposed in the literature, of this family of proteins as key regulators of Wnt extracellular trafficking. The lipid-shielding ability of Dlp, Gpc4 and Gpc6, in particular, could allow these molecules to spatially and temporarily control and delay the activation

of the Wnt signalling cascade until suitable conditions, in terms of receptor concentration and appropriate membrane localization, are formed at the target cells.

Uncovering this interaction, however, only provides a single, although important, snapshot of the complicated journey of the morphogen from the source to receptor engagement. In fact, since the glypican molecules themselves are devoid of signal transduction capability and act as carrier proteins, for the signalling to eventually happen Wnt still needs to be transferred from glypicans to Frizzled/non-canonical Wnt receptors at the target membrane. In other words, a handover process has to take place. But, if this involves a prolonged contact of the free Wnt lipid with the aqueous environment, it would be energetically disfavoured.

To overcome this hurdle, the proteins involved in the handover need therefore, first of all, to be spatially close. In this sense, possibly, cytonemes or similar cellular protrusions might play a role. As already pointed out in **Chapter 3**, in fact, the glypican model does not rule out an involvement of cytonemes, which, on the contrary, have been reported to be enriched in glypicans. Thus, these cellular substructures may allow the glypican-carried Wnt to be brought into close proximity with its target receptors even at a distance from the cell body. Coupled with kinetic considerations (higher affinity for Wnt of target receptors compared to glypicans - as in the case of Frizzled receptors) this might provide a rationale making the handover process more energetically favourable.

To better untangle the conundrum of Wnt handover to receptors, it would therefore be interesting to assess if (and which) glypicans can bind (which)

Frizzleds, and, given binding occurs, if the interaction involves the HS chains, the core protein of the glypicans, or both. As any potential physical contact should involve only the extracellular domain of the Frizzled receptor (because the glypicans do not have any transmembrane region apart from the GPI anchor), and several structures of Frizzled extracellular CRDs have already been solved[181], it should be relatively straightforward to first evaluate the binding between glypicans and Frizzled CRDs by biophysical methods, as a guide, eventually, for a more detailed structural analysis.

Possibly, the evaluation of potential glypican-Frizzled interactions might also shed light on the reason for the incorporation of a CRD-like domain in the glypican proteins. Although the evolutionarily conserved lipid-binding role of the domain[182] is hampered in glypicans by the groove-blocking helix α1 (see **Figure 3.7**), the CRD might still act as an interactive scaffold. Indeed, evidence is present in the literature that recombinantly expressed CRDs of Frizzled (e.g Fz7) have the propensity to dimerize in solution[181]; a potential hetero-dimerization between the CRD of glypicans and the CRD of Frizzled might thus contribute to bringing the carrier protein and the receptor into closer proximity for a subsequent Wnt handover, and/or it might help to increase the local concentration of Frizzled. Alternatively, transducing conformational rearrangements through the rigid glypican structure, the region might work as an allosteric modulator, regulating the open-closed state of the palmitoleate-binding pocket and prompting offload of the lipid.

Similarly, how Wnt gets to glypicans in the first place upon secretion also needs clarification. The answer to this question overlaps with the need to fully

understand the function of the intracellular carrier Wntless, in particular where the lipid-binding unit is located, at which point in the secretory route Wnt dissociate from this carrier, and what is the affinity between Wntless and Wnt.

Whereas, as discussed in **Chapter 4**, the recently solved structure of Wls in complex with Wnt has provided clarity on the location of the Wnt binding pocket in a deep hydrophobic cavity at the core of the Wls transmembrane region (a binding mode that suggests high affinity binding between the two proteins), the dynamics of Wnt release from Wls is still unclear.

Either in the case that Wnt remains bound to Wntless up to the extracellular space, or that Wnt is released first in the plasma membrane to then be passed onto glypicans, it is likely that the process faces kinetic and energetic barriers similar to the handover to Frizzled. How then does the transfer happen? Might the enlarged, not palmitoleate-occupied, bottom part of the binding tunnel I observed in Dlp have a role?

It could be in fact postulated that this unoccupied portion of the binding tunnel is oriented towards the plasma membrane, based on experimentally-supported modelling of the orientation of the glypican Gpc1 with respect to the cell-surface[78]. It is therefore fascinating to speculate that this region might somehow be involved in a “pocket kissing” mechanism similar to that reported for the NPC1 (Niemann-Pick C1) - NPC2 (Niemann Pick C2) interaction[183,184]. NPC1 and NPC2 are an integral lysosomal membrane protein and a lysosomal soluble protein, respectively, responsible for the transfer of free cholesterol from endocytosed lipoprotein particles to the cell. Soluble NPC2 and the luminal domain of NPC1 bind cholesterol in opposite orientations, and in turn their binding pockets interact

and orient themselves to form a continuous tunnel which allows cholesterol to move, shielded from the aqueous environment, from the lumen to the lysosomal membrane. Might a similar tunnel formation happen also in the case of Dlp and the lipid-binding tunnel of Wntless? Of course, a key difference is that in the case of NPCs the lipid is free to travel into the tunnel, while the palmitoleate in Wnt is bound to the whole Wnt protein, and this complicates any lipid offload happening by such a “kissing” mechanism. Even in this case, however, the results from the work presented in this thesis (identification of the role of glypicans as extracellular carriers, successful purification of Wls in lipid-based polymers) represent an important starting point for biophysical analysis, to assess possible direct physical interactions between the two Wnt-binding molecules and start addressing these questions.

In summary, the next efforts in the area should aim at bridging the detailed structural and functional information obtained about the Wnt-glypican interaction with the broader, dynamic picture of the Wnt journey to finally activate downstream signalling.

Glypicans: a more general hub for lipid-modified morphogens?

As already introduced in the discussion of **Chapter 3**, in addition to the Wnt pathway, glypicans have also been widely implicated in the regulation of Hedgehog (Hh) signalling. However, a more nuanced picture emerges there. First of all, the two glypican subclasses have opposing effects on the signalling activity, with the Dlp class presenting a stimulatory effect and the Dally class being instead

inhibitory[185]. Moreover, functional studies in *Drosophila* seem to suggest that, in this case, Dlp behaves more as a co-receptor for Hh, increasing the short-range signalling while repressing the long-range one, therefore with an opposite action compared to that it has in the Wnt pathway[186]. Additionally, as already mentioned in **Chapter 3**, in the work that led to the solution of the apo Dlp structure[76], co-immunoprecipitation experiments between the purified glypican and Hh were performed, but they failed to show any direct interaction between the two proteins (although the form of Hh used was unlipidated). However, in the same paper, mutagenesis studies assessed in cellular reporter assays identified some clusters of mutations in Dlp that negatively affected the signalling strength.

It is therefore interesting to compare and map these mutations onto the Dlp-Wnt peptide structure presented in this thesis. The mutation strategy of the authors of the apo structure followed two steps. Initially, they changed several, surface-exposed, charged or polar residues to alanines, and showed that 4 groups in particular (Group 4, 5, 6 and 7 shown in **Figure 6.1 a**), located in the C-lobe, reduced, although did not abolish, Hh pathway activation. This is in line with our data regarding Wnt: all 4 groups contain residues from the helices forming the binding pocket in our structure, and the mutations for which the strongest reduction was observed (Group 4 and Group 7) include residues from helix α10, which moves to open up the Wnt lipid-binding pocket, and its opposing helix α1. However, since all the mutated residues are surface-exposed, i.e. not in the binding pocket, they are unlikely to directly and completely disrupt a hydrophobic interaction with a lipid moiety. Subsequently, the authors re-clustered some of the mutations and found that a new cluster was able to impair Hh activity almost

completely. In this case, the group maps on the surface of Dlp opposite to the entrance of the Wnt peptide binding pocket, above the portion of the binding tunnel not occupied by the Wnt lipid (**Figure 6.1 b**).

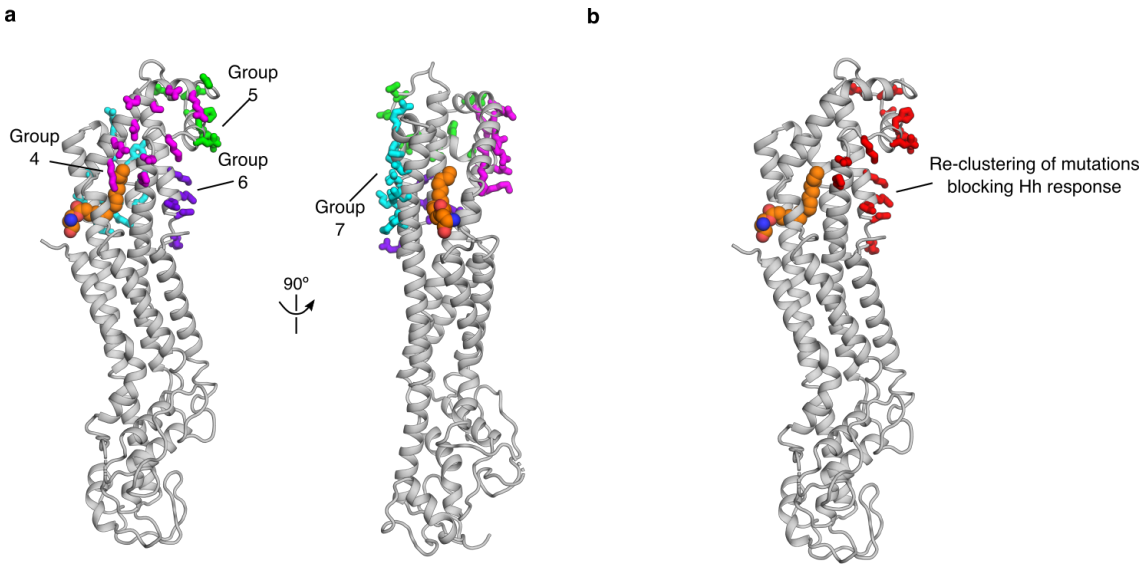


Figure 6.1. The Dlp-Wnt peptide structure and the Hedgehog pathway. **a)** Mapping of residues on the Wnt peptide-Dlp complex structure affecting Hh binding as identified by ref. [76]: clusters of residues resulting in a reduction of Hh response below 70% compared to wild-type when mutated to alanine (group 4, 5, 6 and 7 according to the authors' nomenclature). Group 4 includes residues belonging to helix α10, while group 7 includes residues belonging to helix α1. **b)** Subsequent reclustering of residues from the former groups resulting in almost complete impairment of Hedgehog response (R206, D208, L210, E211, K215, D235, E236, N237, R240, E244, H245, E248, K250, D254, K258).

It would however be too speculative to interpret these findings in light of the Wnt-Dlp structure without experimental support.

First of all, it cannot be excluded that some effects of the mutations might be caused by a too drastic change of the chemical properties of the local environment, including loss of polar patches. Second, direct comparison of Wnt structure-function data with Hh is complicated: while Wnt has a single

palmitoleate, Hh is dually lipidated with palmitate and cholesterol, and both modifications influence pathway activation. Moreover, even the molecular mechanism of receptor engagement is different, and less direct. In the Hh pathway, in fact, in absence of the morphogen a 12-pass integral membrane protein, called Patched 1, inhibits the signalling function of the key 7-pass G protein-coupled receptor Smoothened. This inhibition is thought to be generated by Patched functioning as sterol transporter, restricting and preventing the interaction of Smoothened with its activating free cholesterol substrate[187]. When Hh triggers signalling, it does so by binding Patched in a 1:2 architecture: a single morphogen interacts with high affinity with one of the two receptor molecules through its protein portion, while its palmitate and cholesterol protrusions are inserted into the extracellular domain of the second Patched molecule. In doing so, Hh closes the Patched sterol-binding pocket, and therefore blocks its cholesterol transport activity; the free sterol thus becomes available to bind and activate Smoothened and, in turn, downstream signalling[188]. However, many questions remain concerning the determinants of Hh activity, as unlipidated Hh, and a version of the morphogen carrying only the palmitate, have both been shown to still bind and inhibit Patched when used at high concentrations *in vitro*, while this does not seem to be the case for a peptide presenting the cholesterol modification alone[188].

Further studies are therefore clearly needed to first determine if the glypicans, or a subset of them, directly bind Hh, and in turn if they do so by interacting with the palmitate, the cholesterol, or both. To start addressing these questions, I have performed some preliminary biophysical experiments with a palmitate-modified

Hh peptide. The results suggest the existence of a direct interaction between the peptide and Dlp (**Figure 6.2 a**). In addition, I have designed and optimized a suitable construct and purification protocol to obtain high yields of the closest Dlp homologue, human Gpc6 (**Figure 6.2 b**). Gpc6 apo crystals have already been obtained and will soon undergo x-ray analysis, while crystallization screenings are ongoing to grow crystals also in the presence of palmitate and cholesterol-modified peptides. Finally, a panel of Gpc6 mutants (based on three-dimensional homology modelling) has also been generated (**Figure 6.2 c**), incorporating residue-to-cysteine mutations, to introduce an artificial disulphide bond locking the putative binding pocket in the closed conformation - following a strategy similar to the one successfully used for the Wnt peptide-Dlp interaction. Hence, methods for the recombinant expression of Dlp, Dally, Gpc6 and Gpc3 are now ready, and this panel can be used for comprehensive biophysical and structural studies to identify any potential Hh binding module in the glypicans.

Ultimately, if the glypicans will be shown to have a broader lipid sequestration and handover role in both pathways, they might prove to be key regulators of the balance between cellular proliferation and differentiation. In several contexts, for example the intestinal stem cell niche, the Wnt pathway and the Hedgehog pathway present opposing functional roles, with the former promoting self-renewal of stem cells at the bottom of the crypts, and the latter aiding the progressive cellular differentiation towards the villi and intestinal lumen. The ability of the glypicans to shield both morphogens, depending on the specific context of concentrations and conditions, might control the balance and determine the cellular fate.

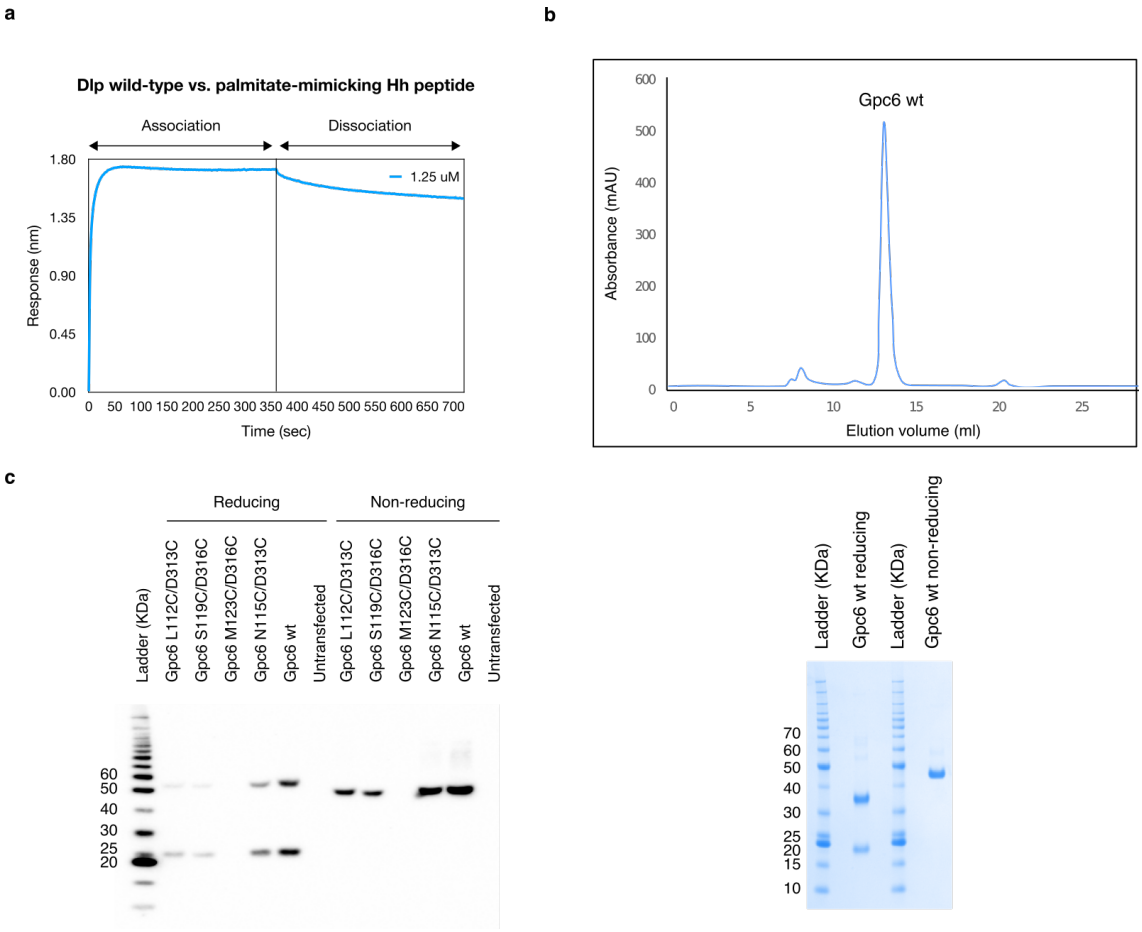


Figure 6.2. Preliminary results for the study of glypican role in the Hedgehog pathway.
a) Biolayer interferometry profile of Dlp tested against Hh peptide N-terminally modified with palmitic acid (kindly provided by Dr Christian Siebold, Division of Structural Biology, University of Oxford). **b)** Size-exclusion chromatography (Superdex 200 Increase 10_300 GL, GE Healthcare) elution profile and SDS-PAGE of the pooled peak fractions in reducing and non-reducing conditions for human Gpc6 (aa. 24 - 488). The SDS-PAGE confirms protein purity and identity: since the protein contains a furin cleavage site, it undergoes furin processing. The two resulting N-terminal and C-terminal fragments are held together by disulphide bonds and run therefore as two bands in reducing conditions. The sum of their molecular weights corresponds to the molecular weight of the Gpc6 construct (~ 55 Kda). In non-reducing conditions, a single band is present as expected. **c)** Western Blot on supernatant from small-scale transfection of Gpc6 mutants L112C/D313C, N115C/D313C, S119C/D316C and M123C/D316C in HEK293T cells. The proteins were probed with a His-reactive anti-PentaHistidine primary antibody (Qiagen). All the mutants except for M123C/D316C expressed. The positive control is represented by Gpc6 wild-type, while the negative control is represented by the supernatant from untransfected cells. The two bands in reducing conditions represent partially furin-cleaved Gpc6, with the top band

corresponding to uncleaved Gpc6 and the lower band to the tagged cleaved C-terminal fragment, as explained for panel b.

Therapeutic targeting of the Wnt pathway

The preliminary *in vivo* studies conducted with our lead Notum inhibitor candidate have shown how the compound has very promising pharmacokinetic properties and blood-brain-barrier permeability; the discussion of **Chapter 5** has analysed, with reference to the available literature, why a reasonable safety profile can be expected from drugs targeting Notum. The key question that remains to be addressed pertains therefore to the efficacy of the drug, a caveat in particular when devising therapeutic strategies for Alzheimer's disease.

AD drug development has been in fact termed the "valley of death"[189], to indicate the lack of success in translating promising treatments from the preclinical to the clinical stage. Whereas the reasons for this failure are multiple, and include the multifactorial nature of the disease, an incomplete understanding of the underlying biology, and the inappropriate selection of preclinical models, the need to improve the validity and translatability between animal and human findings has been recognized as particularly important[190]. With respect to this aim, two aspects are of significant relevance: selecting biological targets that share similar, pathological, changes between mice and humans; and finding biomarkers that represent a direct readout of target engagement at a pharmacodynamic level, and can be used to define the right dose and select the appropriate endpoints in proof-of-principle clinical trials[190]. In this sense, it is particularly encouraging that Notum appears to be involved in Alzheimer's pathogenesis both in mice and in

humans, as shown by the preliminary *in vivo* data of our Consortium (see **Figure 5.18**). Further studies should be directed at strengthening this evidence and fully elucidating the underlying mechanistic link.

Beyond Alzheimer's disease, however, the Notum inhibitors we are developing could add a further important bullet to our ammunitions aimed at the upstream Wnt pathway, and in particular at the Wnt protein itself. Why might hitting Wnt at the source be so important from a therapeutic perspective, for example in cancers?

As already introduced in this thesis, the Wnt cascade can be targeted at several points.

Focusing only on drugs that have reached the clinical trial stage, many candidate therapeutics interfere with the pathway at the transcriptional level: in particular, the aim has been to disrupt the interaction between β-catenin and its transcriptional co-activator CREB-binding protein (CBP)[175]. While this approach has the advantage of targetting the final effector of the cascade - and therefore being effective irrespective of the upstream or downstream localization of the mutation or molecular alteration responsible for the disease - it does not cover the β-catenin independent Wnt signalling. Moreover, it has been reported that even tumours that mostly arise because of downstream mutations (e.g. of APC), like colon cancers, are still dependent on upstream Wnt receptor engagement to sustain their growth[191], and might require an additional layer of therapeutic blockade.

Alternative strategies have therefore focused on the receptor level, in particular with monoclonal antibodies against Frizzleds. In this case, however, the problem

lies in our incomplete understanding of the specificity of particular Wnt-Fz combinations, and in the apparent redundancy that allows other Frizzled receptors to surrogate the functional role of the specific family member(s) that is bound by the monoclonal antibody[192].

That is why going even more upstream, and targeting a biological feature that is shared between all Wnt members, and probably is crucial to the interaction with many receptors - the lipidation - is so attractive. As already mentioned, in this sense the drugs currently in clinical trials inhibit Porcupine, the enzyme which, intracellularly, chemically attaches the palmitoleate to Wnt[175,177]. Compared to Porcupine, Notum is easier to effectively be reached pharmacokinetically thanks to its extracellular localization. Furthermore, the drugs against Notum have the advantage of a structure-based drug design not possible for Porcupine, as the structure of the latter has not yet been solved. However, it should be kept in mind that strategies targeting Porcupine vs. Notum are not alternative, but complementary. Inhibitors of Porcupine block the Wnt lipidation, thereby reducing Wnt signalling; on the other hand, Notum inhibitors prevent the enzyme from de-palmitoleoylating Wnt, which can therefore maintain the engagement of its target receptors.

Having drugs serving both to broadly inhibit and to broadly boost the Wnt pathway with relative safety might therefore represent a big step forward in our therapeutic options in several tumour contexts. In fact, as already noted in the **Introduction**, the “increased Wnt pathway - increased proliferation” paradigm is overly simplified, and an upregulated Wnt pathway can also play a protective role depending on the cancer location, histology and cellular composition, and even

on the progression stage of the tumour natural history[193]. Therefore, the most effective therapy will likely use combination treatments, carefully selecting and alternating the drug regimens based on patient-specific information derived by clinical samples and biopsies.

Finally, the glypican studies reported here also represent an important building block to further expand the number of drugs that can effectively affect Wnt lipidation. Indeed, they provide structural information to guide further strategies to modulate the upstream Wnt distribution. This could be achieved, for instance, by antibodies and/or small molecules against the Wnt binding region of glypicans, or by recombinant soluble mimetics of the binding pocket, to be used as decoy traps for Wnt.

Appendix

Table 7.1. Compounds in the Notum drug development project on which I have worked in collaboration with Dr Yuguang Zhao, Dr James Hillier and Mr Reinis Reinholds Ruza (the compounds highlighted are those for which I carried out the whole pipeline).

Compound class	Compound code	Structure with inhibitor
Carboxylic acids	ARUK3000454	Yes
	ARUK3000458	Yes
	ARUK3000471	Yes
	ARUK3000510	Yes
	ARUK3000511	Yes
	ARUK3000523	Yes
	ARUK3000534	
	ARUK3000559	
	ARUK3000569	
	ARUK3000572	Yes
	ARUK3000586	
	ARUK3000587	Yes
	ARUK3000589	
	ARUK3000598	Yes
	ARUK3000605	

	ARUK3000634	Yes
	ARUK3000644	Yes
	ARUK3000647	
	ARUK3000655	Yes
	ARUK3000686	Yes
	ARUK3001246	Yes
Heteroaryl compounds	ARUK3000223	Yes
	ARUK3000263	Yes
	ARUK3000276	Yes
	ARUK3000395	
	ARUK3000438	Yes
	ARUK3000694	
	ARUK3000697	
	ARUK3000701	Yes
	ARUK3000705	
	ARUK3000754	Yes
	ARUK3000759	Yes
	ARUK3000769	Yes
	ARUK3000776	
	ARUK3000781	Yes

	ARUK3000784	Yes
	ARUK3000833	Yes
	ARUK3000874	
	ARUK3000875	Yes
	ARUK3000900	Yes
	ARUK3000902	
	ARUK3001043	Yes
	ARUK3001045	Yes
	ARUK3001185	Yes
	ARUK3001195	Yes
	ARUK3001213	Yes
	ARUK3001221	Yes
	ARUK3001233	Yes
	ARUK3001241	Yes
	ARUK3001244	Yes
	ARUK3001261	Yes
Phenoxyamides	ARUK3000339	Yes
	ARUK3000367	
	ARUK3000385	
	ARUK3000400	Yes

	ARUK3000442	Yes
	ARUK3000744	
	ARUK3000745	
	ARUK3000752	
	ARUK3000761	
	ARUK3000763	Yes
	ARUK3000774	
	ARUK3000849	Yes
	ARUK3000893	
	ARUK3000894	
	ARUK3000896	
	ARUK3000897	

List of publications derived from the work presented in this thesis

- Mahy W, Willis NJ, Zhao Y, Woodward HL, Svensson F, Sipthorp J, **Vecchia L**, Ruza RR, Hillier J, Kjaer S, Frew S, Monaghan A, Bictash M, Salinas P, Whiting, P, Vincent JP, Jones EY, Fish, PV. "**5-Phenyl-1,3,4-oxadiazol-2(3H)-ones are potent inhibitors of Notum carboxylesterase activity identified by the optimization of a crystallographic fragment screening hit".** *J Med Chem* 2020; 63(21):12942-12956.

- Mahy W, Patel M, Steadman D, Woodward HL, Atkinson BN, Svensson F, Willis NJ, Flint A, Papatheodorou D, Zhao Y, **Vecchia L**, Ruza RR, Hillier J, Frew S, Monaghan A, Costa A, Bictash M, Walter MW, Jones EY, Fish PV. "**Screening of a custom-designed acid fragment library identifies 1-phenylpyrroles and 1-phenylpyrrolidines as Inhibitors of Notum carboxylesterase activity**". *J Med Chem* 2020; 63(17):9464-9483.

- McGough IJ*, **Vecchia L***, Bishop B, Malinauskas T, Beckett K, Joshi D, O'Reilly N, Siebold C, Jones EY, Vincent JP. "**Glypicans shield the Wnt lipid moiety to enable signalling at a distance**". *Nature* 2020; 585(7823):85-90.
  * Equivalent contribution

- Atkinson BN, Steadman D, Zhao Y, Sipthorp J, **Vecchia L**, Ruza RR, Jeganathan F, Lines G, Frew S, Monaghan A, Kjær S, Bictash M, Jones EY, Fish PV. "**Discovery of 2-phenoxyacetamides as inhibitors of the Wnt-depalmitoleating enzyme NOTUM from an X-ray fragment screen**". *Medchemcomm*. 2019 Apr 29; 10(8):1361-1369 (awarded journal cover page).

- Fish PV, Mahy W, Willis NJ, Woodward HL, Atkinson BN, Bayle ED, Sipthorp J, Jones EY, Zhao Y, **Vecchia L**, Ruza RR. International patent application n. PCT/EP2019/073177, titled **"Compounds" (Wnt pathway activators)**, filing date 30.08.2019.

Copyright information

All the figures except the ones listed below are of my creation and/or derive from published papers of which I am an author. The licenses to reuse published figures not of my creation are listed below.

Figure 1.1. Copyright © 2017 Elsevier Inc. All rights reserved. License Number 5011980241862.

Figure 1.2. Reprinted with permission from AAAS. No license is needed for reproduction in a dissertation/thesis in accordance with the publisher's policy.

Figure 1.3. Copyright © 2016, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved. License Number 5011990818915.

Figure 1.4. Panel a, Copyright © 2012, Springer Nature. All Rights Reserved. License Number 5012020744439.

Figure 1.5 and **Figure 1.6**. Copyright © 2019, The Company of Biologists Ltd. All Rights Reserved. License 1098420-1.

Figure 1.7. Panel a, CC BY 4.0 license http://creativecommons.org/licenses/by/4.0/ and Copyright © 2014 Annual Reviews, Inc.. All rights reserved. License Number 1100118-1. Panel b, Copyright © 2008, BioMed Central Ltd. All rights reserved. License Number 5016040375665.

Figure 1.8. Panel a-c, Copyright © 2015, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved. License Number 5016070440928. Panel d, Copyright © 2016, Nature Publishing Group, a division

of Macmillan Publishers Limited. All Rights Reserved. License Number 5011990818915.

Figure 4.9. Panel a-c, Copyright © 2020 Elsevier Inc. All Rights Reserved. License Number 5023730099602.

Figure 5.1. Copyright © 2018, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved. License Number 5014250319135.

References

1 Kusserow, A. *et al.* Unexpected complexity of the Wnt gene family in a sea anemone. *Nature* **433**, 156-160, doi:10.1038/nature03158 (2005).

2 Petersen, C. P. & Reddien, P. W. Wnt signaling and the polarity of the primary body axis. *Cell* **139**, 1056-1068, doi:10.1016/j.cell.2009.11.035 (2009).

3 Nusse, R. & Clevers, H. Wnt/beta-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* **169**, 985-999, doi:10.1016/j.cell.2017.05.016 (2017).

4 Wiese, K. E., Nusse, R. & van Amerongen, R. Wnt signalling: conquering complexity. *Development* **145**, doi:10.1242/dev.165902 (2018).

5 Nusslein-Volhard, C. & Wieschaus, E. Mutations affecting segment number and polarity in Drosophila. *Nature* **287**, 795-801, doi:10.1038/287795a0 (1980).

6 Steinhart, Z. & Angers, S. Wnt signaling in development and tissue homeostasis. *Development* **145**, doi:10.1242/dev.146589 (2018).

7 Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262-265, doi:10.1038/nature07935 (2009).

8 Kinzler, K. W. *et al.* Identification of a gene located at chromosome 5q21 that is mutated in colorectal cancers. *Science* **251**, 1366-1370, doi:10.1126/science.1848370 (1991).

9 Nishisho, I. *et al.* Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* **253**, 665-669, doi:10.1126/science.1651563 (1991).

10 Kinzler, K. W. & Vogelstein, B. Lessons from hereditary colorectal cancer. *Cell* **87**, 159-170, doi:10.1016/s0092-8674(00)81333-1 (1996).

11 Cancer Genome Atlas, N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330-337, doi:10.1038/nature11252 (2012).

12 Schatoff, E. M., Leach, B. I. & Dow, L. E. Wnt Signaling and Colorectal Cancer. *Curr Colorectal Cancer Rep* **13**, 101-110, doi:10.1007/s11888-017-0354-9 (2017).

13 Anastas, J. N. & Moon, R. T. WNT signalling pathways as therapeutic targets in cancer. *Nat Rev Cancer* **13**, 11-26, doi:10.1038/nrc3419 (2013).

14 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).

15 Zhong, Z., Yu, J., Virshup, D. M. & Madan, B. Wnts and the hallmarks of cancer. *Cancer Metastasis Rev*, doi:10.1007/s10555-020-09887-6 (2020).

16 Biehs, B. *et al.* A cell identity switch allows residual BCC to survive Hedgehog pathway inhibition. *Nature* **562**, 429-433, doi:10.1038/s41586-018-0596-y (2018).

17 Zeng, X. *et al.* A dual-kinase mechanism for Wnt co-receptor phosphorylation and activation. *Nature* **438**, 873-877, doi:10.1038/nature04185 (2005).

18 Mao, J. *et al.* Low-density lipoprotein receptor-related protein-5 binds to Axin and regulates the canonical Wnt signaling pathway. *Mol Cell* **7**, 801-809, doi:10.1016/s1097-2765(01)00224-6 (2001).
19 Li, V. S. *et al.* Wnt signaling through inhibition of beta-catenin degradation in an intact Axin1 complex. *Cell* **149**, 1245-1256, doi:10.1016/j.cell.2012.05.002 (2012).
20 Hernandez, A. R., Klein, A. M. & Kirschner, M. W. Kinetic responses of beta-catenin specify the sites of Wnt control. *Science* **338**, 1337-1340, doi:10.1126/science.1228734 (2012).
21 Gammons, M. V., Renko, M., Johnson, C. M., Rutherford, T. J. & Bienz, M. Wnt Signalosome Assembly by DEP Domain Swapping of Dishevelled. *Mol Cell* **64**, 92-104, doi:10.1016/j.molcel.2016.08.026 (2016).
22 van Amerongen, R. Alternative Wnt pathways and receptors. *Cold Spring Harb Perspect Biol* **4**, doi:10.1101/cshperspect.a007914 (2012).
23 Zebisch, M. *et al.* Structural and molecular basis of ZNRF3/RNF43 transmembrane ubiquitin ligase inhibition by the Wnt agonist R-spondin. *Nat Commun* **4**, 2787, doi:10.1038/ncomms3787 (2013).
24 Zebisch, M. & Jones, E. Y. Crystal structure of R-spondin 2 in complex with the ectodomains of its receptors LGR5 and ZNRF3. *J Struct Biol* **191**, 149-155, doi:10.1016/j.jsb.2015.05.008 (2015).
25 Hao, H. X. *et al.* ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner. *Nature* **485**, 195-200, doi:10.1038/nature11019 (2012).
26 Willert, K. *et al.* Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature* **423**, 448-452, doi:10.1038/nature01611 (2003).
27 Janda, C. Y., Waghray, D., Levin, A. M., Thomas, C. & Garcia, K. C. Structural basis of Wnt recognition by Frizzled. *Science* **337**, 59-64, doi:10.1126/science.1222879 (2012).
28 Hirai, H., Matoba, K., Mihara, E., Arimori, T. & Takagi, J. Crystal structure of a mammalian Wnt-frizzled complex. *Nat Struct Mol Biol* **26**, 372-379, doi:10.1038/s41594-019-0216-z (2019).
29 Chu, M. L. *et al.* structural Studies of Wnts and identification of an LRP6 binding site. *Structure* **21**, 1235-1242, doi:10.1016/j.str.2013.05.006 (2013).
30 Nile, A. H., Mukund, S., Stanger, K., Wang, W. & Hannoush, R. N. Unsaturated fatty acyl recognition by Frizzled receptors mediates dimerization upon Wnt ligand binding. *Proc Natl Acad Sci U S A* **114**, 4147-4152, doi:10.1073/pnas.1618293114 (2017).
31 Nile, A. H. & Hannoush, R. N. Fatty acylation of Wnt proteins. *Nat Chem Biol* **12**, 60-69, doi:10.1038/nchembio.2005 (2016).
32 Takada, R. *et al.* Monounsaturated fatty acid modification of Wnt protein: its role in Wnt secretion. *Dev Cell* **11**, 791-801, doi:10.1016/j.devcel.2006.10.003 (2006).
33 Herr, P. & Basler, K. Porcupine-mediated lipidation is required for Wnt recognition by Wls. *Dev Biol* **361**, 392-402, doi:10.1016/j.ydbio.2011.11.003 (2012).
34 Langton, P. F., Kakugawa, S. & Vincent, J. P. Making, Exporting, and Modulating Wnts. *Trends Cell Biol* **26**, 756-765, doi:10.1016/j.tcb.2016.05.011 (2016).

35 Miranda, M. *et al.* Identification of the WNT1 residues required for palmitoylation by Porcupine. *FEBS Lett* **588**, 4815-4824, doi:10.1016/j.febslet.2014.11.016 (2014).
36 Tuladhar, R. *et al.* Stereoselective fatty acylation is essential for the release of lipidated WNT proteins from the acyltransferase Porcupine (PORCN). *J Biol Chem* **294**, 6273-6282, doi:10.1074/jbc.RA118.007268 (2019).
37 Banziger, C. *et al.* Wntless, a conserved membrane protein dedicated to the secretion of Wnt proteins from signaling cells. *Cell* **125**, 509-522, doi:10.1016/j.cell.2006.02.049 (2006).
38 Bartscherer, K., Pelte, N., Ingelfinger, D. & Boutros, M. Secretion of Wnt ligands requires Evi, a conserved transmembrane protein. *Cell* **125**, 523-533, doi:10.1016/j.cell.2006.04.009 (2006).
39 Korkut, C. *et al.* Trans-synaptic transmission of vesicular Wnt signals through Evi/Wntless. *Cell* **139**, 393-404, doi:10.1016/j.cell.2009.07.051 (2009).
40 Gasnereau, I., Herr, P., Chia, P. Z., Basler, K. & Gleeson, P. A. Identification of an endocytosis motif in an intracellular loop of Wntless protein, essential for its recycling and the control of Wnt protein signaling. *J Biol Chem* **286**, 43324-43333, doi:10.1074/jbc.M111.307231 (2011).
41 Yu, J. *et al.* WLS retrograde transport to the endoplasmic reticulum during Wnt secretion. *Dev Cell* **29**, 277-291, doi:10.1016/j.devcel.2014.03.016 (2014).
42 Glaeser, K. *et al.* ERAD-dependent control of the Wnt secretory factor Evi. *EMBO J* **37**, doi:10.15252/embj.201797311 (2018).
43 Jin, J., Morse, M., Frey, C., Petko, J. & Levenson, R. Expression of GPR177 (Wntless/Evi/Sprinter), a highly conserved Wnt-transport protein, in rat tissues, zebrafish embryos, and cultured human cells. *Dev Dyn* **239**, 2426-2434, doi:10.1002/dvdy.22369 (2010).
44 Das, S., Yu, S., Sakamori, R., Stypulkowski, E. & Gao, N. Wntless in Wnt secretion: molecular, cellular and genetic aspects. *Front Biol (Beijing)* **7**, 587-593, doi:10.1007/s11515-012-1200-8 (2012).
45 Fu, J., Jiang, M., Mirando, A. J., Yu, H. M. & Hsu, W. Reciprocal regulation of Wnt and Gpr177/mouse Wntless is required for embryonic axis formation. *Proc Natl Acad Sci U S A* **106**, 18598-18603, doi:10.1073/pnas.0904894106 (2009).
46 Coombs, G. S. *et al.* WLS-dependent secretion of WNT3A requires Ser209 acylation and vacuolar acidification. *J Cell Sci* **123**, 3357-3367, doi:10.1242/jcs.072132 (2010).
47 Zecca, M., Basler, K. & Struhl, G. Direct and long-range action of a wingless morphogen gradient. *Cell* **87**, 833-844, doi:10.1016/s0092-8674(00)81991-1 (1996).
48 Kiecker, C. & Niehrs, C. A morphogen gradient of Wnt/beta-catenin signalling regulates anteroposterior neural patterning in Xenopus. *Development* **128**, 4189-4201 (2001).
49 Farin, H. F. *et al.* Visualization of a short-range Wnt gradient in the intestinal stem-cell niche. *Nature* **530**, 340-343, doi:10.1038/nature16937 (2016).

50 Alexandre, C., Baena-Lopez, A. & Vincent, J. P. Patterning and growth control by membrane-tethered Wingless. *Nature* **505**, 180-185, doi:10.1038/nature12879 (2014).

51 Tian, A., Duwadi, D., Benchabane, H. & Ahmed, Y. Essential long-range action of Wingless/Wnt in adult intestinal compartmentalization. *PLoS Genet* **15**, e1008111, doi:10.1371/journal.pgen.1008111 (2019).

52 Bartscherer, K. & Boutros, M. Regulation of Wnt protein secretion and its role in gradient formation. *EMBO Rep* **9**, 977-982, doi:10.1038/embor.2008.167 (2008).

53 Yamamoto, H. *et al.* The apical and basolateral secretion of Wnt11 and Wnt3a in polarized epithelial cells is regulated by different mechanisms. *J Cell Sci* **126**, 2931-2943, doi:10.1242/jcs.126052 (2013).

54 Yamazaki, Y. *et al.* Godzilla-dependent transcytosis promotes Wingless signalling in Drosophila wing imaginal discs. *Nat Cell Biol* **18**, 451-457, doi:10.1038/ncb3325 (2016).

55 Routledge, D. & Scholpp, S. Mechanisms of intercellular Wnt transport. *Development* **146**, doi:10.1242/dev.176073 (2019).

56 Panakova, D., Sprong, H., Marois, E., Thiele, C. & Eaton, S. Lipoprotein particles are required for Hedgehog and Wingless signalling. *Nature* **435**, 58-65, doi:10.1038/nature03504 (2005).

57 Neumann, S. *et al.* Mammalian Wnt3a is released on lipoprotein particles. *Traffic* **10**, 334-343, doi:10.1111/j.1600-0854.2008.00872.x (2009).

58 Gross, J. C., Chaudhary, V., Bartscherer, K. & Boutros, M. Active Wnt proteins are secreted on exosomes. *Nat Cell Biol* **14**, 1036-1045, doi:10.1038/ncb2574 (2012).

59 Beckett, K. *et al.* Drosophila S2 cells secrete wingless on exosome-like vesicles but the wingless gradient forms independently of exosomes. *Traffic* **14**, 82-96, doi:10.1111/tra.12016 (2013).

60 Mulligan, K. A. *et al.* Secreted Wingless-interacting molecule (Swim) promotes long-range signaling by maintaining Wingless solubility. *Proc Natl Acad Sci U S A* **109**, 370-377, doi:10.1073/pnas.1119197109 (2012).

61 Naschberger, A. *et al.* Structural Evidence for a Role of the Multi-functional Human Glycoprotein Afamin in Wnt Transport. *Structure* **25**, 1907-1915 e1905, doi:10.1016/j.str.2017.10.006 (2017).

62 Stanganello, E. *et al.* Filopodia-based Wnt transport during vertebrate tissue patterning. *Nat Commun* **6**, 5846, doi:10.1038/ncomms6846 (2015).

63 Moti, N., Yu, J., Boncompain, G., Perez, F. & Virshup, D. M. Wnt traffic from endoplasmic reticulum to filopodia. *PLoS One* **14**, e0212711, doi:10.1371/journal.pone.0212711 (2019).

64 Mattes, B. *et al.* Wnt/PCP controls spreading of Wnt/beta-catenin signals by cytonemes in vertebrates. *Elife* **7**, doi:10.7554/eLife.36953 (2018).

65 Yan, D. & Lin, X. Shaping morphogen gradients by proteoglycans. *Cold Spring Harb Perspect Biol* **1**, a002493, doi:10.1101/cshperspect.a002493 (2009).

66 Binari, R. C. *et al.* Genetic evidence that heparin-like glycosaminoglycans are involved in wingless signaling. *Development* **124**, 2623-2632 (1997).

67 Hacker, U., Lin, X. & Perrimon, N. The Drosophila sugarless gene modulates Wingless signaling and encodes an enzyme involved in polysaccharide biosynthesis. *Development* **124**, 3565-3573 (1997).
68 Haerry, T. E., Heslip, T. R., Marsh, J. L. & O'Connor, M. B. Defects in glucuronate biosynthesis disrupt Wingless signaling in Drosophila. *Development* **124**, 3055-3064 (1997).
69 Lin, X. & Perrimon, N. Dally cooperates with Drosophila Frizzled 2 to transduce Wingless signalling. *Nature* **400**, 281-284, doi:10.1038/22343 (1999).
70 Han, C., Yan, D., Belenkaya, T. Y. & Lin, X. Drosophila glypicans Dally and Dally-like shape the extracellular Wingless morphogen gradient in the wing disc. *Development* **132**, 667-679, doi:10.1242/dev.01636 (2005).
71 Franch-Marro, X. *et al.* Glypicans shunt the Wingless signal between local signalling and further transport. *Development* **132**, 659-666, doi:10.1242/dev.01639 (2005).
72 Topczewski, J. *et al.* The zebrafish glypican knypek controls cell polarity during gastrulation movements of convergent extension. *Dev Cell* **1**, 251-264, doi:10.1016/s1534-5807(01)00005-3 (2001).
73 Tao, Q. *et al.* Maternal wnt11 activates the canonical wnt signaling pathway required for axis formation in Xenopus embryos. *Cell* **120**, 857-871, doi:10.1016/j.cell.2005.01.013 (2005).
74 Xu, D. & Esko, J. D. Demystifying heparan sulfate-protein interactions. *Annu Rev Biochem* **83**, 129-157, doi:10.1146/annurev-biochem-060713-035314 (2014).
75 Yan, D., Wu, Y., Feng, Y., Lin, S. C. & Lin, X. The core protein of glypican Dally-like determines its biphasic activity in wingless morphogen signaling. *Dev Cell* **17**, 470-481, doi:10.1016/j.devcel.2009.09.001 (2009).
76 Kim, M. S., Saunders, A. M., Hamaoka, B. Y., Beachy, P. A. & Leahy, D. J. Structure of the protein core of the glypican Dally-like and localization of a region important for hedgehog signaling. *Proc Natl Acad Sci U S A* **108**, 13112-13117, doi:10.1073/pnas.1109877108 (2011).
77 Svensson, G., Awad, W., Hakansson, M., Mani, K. & Logan, D. T. Crystal structure of N-glycosylated human glypican-1 core protein: structure of two loops evolutionarily conserved in vertebrate glypican-1. *J Biol Chem* **287**, 14040-14051, doi:10.1074/jbc.M111.322487 (2012).
78 Awad, W. *et al.* Structural Aspects of N-Glycosylations and the C-terminal Region in Human Glypican-1. *J Biol Chem* **290**, 22991-23008, doi:10.1074/jbc.M115.660878 (2015).
79 Filmus, J., Capurro, M. & Rast, J. Glypicans. *Genome Biol* **9**, 224, doi:10.1186/gb-2008-9-5-224 (2008).
80 Cruciat, C. M. & Niehrs, C. Secreted and transmembrane wnt inhibitors and activators. *Cold Spring Harb Perspect Biol* **5**, a015081, doi:10.1101/cshperspect.a015081 (2013).
81 Malinauskas, T. & Jones, E. Y. Extracellular modulators of Wnt signalling. *Curr Opin Struct Biol* **29**, 77-84, doi:10.1016/j.sbi.2014.10.003 (2014).
82 Zhang, X. *et al.* Tiki1 is required for head formation via Wnt cleavage-oxidation and inactivation. *Cell* **149**, 1565-1577, doi:10.1016/j.cell.2012.04.039 (2012).

83 Dann, C. E. *et al.* Insights into Wnt binding and signalling from the structures of two Frizzled cysteine-rich domains. *Nature* **412**, 86-90, doi:10.1038/35083601 (2001).

84 Malinauskas, T., Aricescu, A. R., Lu, W., Siebold, C. & Jones, E. Y. Modular mechanism of Wnt signaling inhibition by Wnt inhibitory factor 1. *Nat Struct Mol Biol* **18**, 886-893, doi:10.1038/nsmb.2081 (2011).

85 Kakugawa, S. *et al.* Notum deacylates Wnt proteins to suppress signalling activity. *Nature* **519**, 187-192, doi:10.1038/nature14259 (2015).

86 Traister, A., Shi, W. & Filmus, J. Mammalian Notum induces the release of glypicans and other GPI-anchored proteins from the cell surface. *Biochem J* **410**, 503-511, doi:10.1042/BJ20070511 (2008).

87 Petersen, C. P. & Reddien, P. W. Polarized notum activation at wounds inhibits Wnt function to promote planarian head regeneration. *Science* **332**, 852-855, doi:10.1126/science.1202143 (2011).

88 Flowers, G. P., Topczewska, J. M. & Topczewski, J. A zebrafish Notum homolog specifically blocks the Wnt/beta-catenin signaling pathway. *Development* **139**, 2416-2425, doi:10.1242/dev.063206 (2012).

89 Torisu, Y. *et al.* Human homolog of NOTUM, overexpressed in hepatocellular carcinoma, is regulated transcriptionally by beta-catenin/TCF. *Cancer Sci* **99**, 1139-1146, doi:10.1111/j.1349-7006.2008.00814.x (2008).

90 Mullis, K. *et al.* Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* **51 Pt 1**, 263-273, doi:10.1101/sqb.1986.051.01.032 (1986).

91 Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51-59, doi:10.1016/0378-1119(89)90358-2 (1989).

92 Aricescu, A. R., Lu, W. & Jones, E. Y. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr D Biol Crystallogr* **62**, 1243-1250, doi:10.1107/S0907444906029799 (2006).

93 Elegheert, J. *et al.* Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins. *Nat Protoc* **13**, 2991-3017, doi:10.1038/s41596-018-0075-9 (2018).

94 Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-5467, doi:10.1073/pnas.74.12.5463 (1977).

95 McGough, I. J. *et al.* Glypicans shield the Wnt lipid moiety to enable signalling at a distance. *Nature*, doi:10.1038/s41586-020-2498-z (2020).

96 Reeves, P. J., Callewaert, N., Contreras, R. & Khorana, H. G. 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, doi:10.1073/pnas.212519299 (2002).

97 Chang, V. T. *et al.* Glycoprotein structural genomics: solving the glycosylation problem. *Structure* **15**, 267-273, doi:10.1016/j.str.2007.01.011 (2007).

98 Walter, T. S. *et al.* A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization. *Acta Crystallogr D Biol Crystallogr* **61**, 651-657, doi:10.1107/S0907444905007808 (2005).
99 Newman, J. Novel buffer systems for macromolecular crystallization. *Acta Crystallogr D Biol Crystallogr* **60**, 610-612, doi:10.1107/S0907444903029640 (2004).
100 Winter, G., Lobley, C. M. & Prince, S. M. Decision making in xia2. *Acta Crystallogr D Biol Crystallogr* **69**, 1260-1273, doi:10.1107/S0907444913015308 (2013).
101 Winter, G. *et al.* DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr D Struct Biol* **74**, 85-97, doi:10.1107/S2059798317017235 (2018).
102 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**, 1204-1214, doi:10.1107/S0907444913000061 (2013).
103 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).
104 McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674, doi:10.1107/S0021889807021206 (2007).
105 Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352-367, doi:10.1107/S0907444912001308 (2012).
106 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
107 Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293-315, doi:10.1002/pro.3330 (2018).
108 Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797, doi:10.1016/j.jmb.2007.05.022 (2007).
109 Tian, W., Chen, C., Lei, X., Zhao, J. & Liang, J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res* **46**, W363-W367, doi:10.1093/nar/gky473 (2018).
110 Zimmermann, L. *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J Mol Biol* **430**, 2237-2243, doi:10.1016/j.jmb.2017.12.007 (2018).
111 Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539, doi:10.1038/msb.2011.75 (2011).
112 Almagro Armenteros, J. J. *et al.* SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**, 420-423, doi:10.1038/s41587-019-0036-z (2019).
113 Kall, L., Krogh, A. & Sonnhammer, E. L. Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Res* **35**, W429-432, doi:10.1093/nar/gkm256 (2007).

114 Buchan, D. W. A. & Jones, D. T. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res* **47**, W402-W407, doi:10.1093/nar/gkz297 (2019).
115 Huang, Y. J., Acton, T. B. & Montelione, G. T. DisMeta: a meta server for construct design and optimization. *Methods Mol Biol* **1091**, 3-16, doi:10.1007/978-1-62703-691-7_1 (2014).
116 Kim, D. E., Chivian, D. & Baker, D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* **32**, W526-531, doi:10.1093/nar/gkh468 (2004).
117 Zhang, Y. *et al.* Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein. *Nature* **546**, 248-253, doi:10.1038/nature22394 (2017).
118 Yin, W. *et al.* Crystal structure of the human 5-HT1B serotonin receptor bound to an inverse agonist. *Cell Discov* **4**, 12, doi:10.1038/s41421-018-0009-2 (2018).
119 Lee, S. C. *et al.* A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nat Protoc* **11**, 1149-1162, doi:10.1038/nprot.2016.070 (2016).
120 Laskowski, R. A. & Swindells, M. B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778-2786, doi:10.1021/ci200227u (2011).
121 Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-W303, doi:10.1093/nar/gky427 (2018).
122 Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858, doi:10.1038/nprot.2015.053 (2015).
123 Beira, J. V. & Paro, R. The legacy of Drosophila imaginal discs. *Chromosoma* **125**, 573-592, doi:10.1007/s00412-016-0595-4 (2016).
124 Harmansa, S., Hamaratoglu, F., Affolter, M. & Caussinus, E. Dpp spreading is required for medial but not for lateral wing disc growth. *Nature* **527**, 317-322, doi:10.1038/nature15712 (2015).
125 Hessvik, N. P. & Llorente, A. Current knowledge on exosome biogenesis and release. *Cell Mol Life Sci* **75**, 193-208, doi:10.1007/s00018-017-2595-9 (2018).
126 Alonso, Y. A. M., Migliano, S. M. & Teis, D. ESCRT-III and Vps4: a dynamic multipurpose tool for membrane budding and scission. *FEBS J* **283**, 3288-3302, doi:10.1111/febs.13688 (2016).
127 Duckert, P., Brunak, S. & Blom, N. Prediction of proprotein convertase cleavage sites. *Protein Eng Des Sel* **17**, 107-112, doi:10.1093/protein/gzh013 (2004).
128 Dhamdhere, G. R. *et al.* Drugging a stem cell compartment using Wnt3a protein as a therapeutic. *PLoS One* **9**, e83650, doi:10.1371/journal.pone.0083650 (2014).
129 Pei, J. & Grishin, N. V. Cysteine-rich domains related to Frizzled receptors and Hedgehog-interacting proteins. *Protein Sci* **21**, 1172-1184, doi:10.1002/pro.2105 (2012).

130 Liebschner, D. *et al.* Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr D Struct Biol* **73**, 148-157, doi:10.1107/S2059798316018210 (2017).
131 Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191, doi:10.1093/bioinformatics/btp033 (2009).
132 Tu, R., Duan, B., Song, X. & Xie, T. Dlp-mediated Hh and Wnt signaling interdependence is critical in the niche for germline stem cell progeny differentiation. *Sci Adv* **6**, eaaz0480, doi:10.1126/sciadv.aaz0480 (2020).
133 Gonzalez-Mendez, L., Gradilla, A. C. & Guerrero, I. The cytoneme connection: direct long-distance signal transfer during development. *Development* **146**, doi:10.1242/dev.174607 (2019).
134 Serralbo, O. & Marcelle, C. Migrating cells mediate long-range WNT signaling. *Development* **141**, 2057-2063, doi:10.1242/dev.107656 (2014).
135 Hayashi, Y., Kobayashi, S. & Nakato, H. Drosophila glypicans regulate the germline stem cell niche. *J Cell Biol* **187**, 473-480, doi:10.1083/jcb.200904118 (2009).
136 Wang, X. & Page-McCaw, A. A matrix metalloproteinase mediates long-distance attenuation of stem cell proliferation. *J Cell Biol* **206**, 923-936, doi:10.1083/jcb.201403084 (2014).
137 Savitsky, P. *et al.* High-throughput production of human proteins for crystallization: the SGC experience. *J Struct Biol* **172**, 3-13, doi:10.1016/j.jsb.2010.06.008 (2010).
138 Tektas, S. S. *Structural and functional analysis of Wntless in Wnt signaling*, University of Delaware, (2015).
139 Choy, B. C., Cater, R. J., Mancia, F. & Pryor, E. E., Jr. A 10-year meta-analysis of membrane protein structural biology: Detergents, membrane mimetics, and structure determination techniques. *Biochim Biophys Acta Biomembr* **1863**, 183533, doi:10.1016/j.bbamem.2020.183533 (2021).
140 Dorr, J. M. *et al.* The styrene-maleic acid copolymer: a versatile tool in membrane research. *Eur Biophys J* **45**, 3-21, doi:10.1007/s00249-015-1093-y (2016).
141 Broecker, J., Eger, B. T. & Ernst, O. P. Crystallogenesis of Membrane Proteins Mediated by Polymer-Bounded Lipid Nanodiscs. *Structure* **25**, 384-392, doi:10.1016/j.str.2016.12.004 (2017).
142 Nygaard, R. *et al.* Structural Basis of WLS/Evi-Mediated Wnt Transport and Secretion. *Cell* **184**, 194-206 e114, doi:10.1016/j.cell.2020.11.038 (2021).
143 Yang, S. *et al.* Crystal structure of the Frizzled 4 receptor in a ligand-free state. *Nature* **560**, 666-670, doi:10.1038/s41586-018-0447-x (2018).
144 Oliva, C. A., Vargas, J. Y. & Inestrosa, N. C. Wnt signaling: role in LTP, neural networks and memory. *Ageing Res Rev* **12**, 786-800, doi:10.1016/j.arr.2013.03.006 (2013).
145 Shimogori, T., VanSant, J., Paik, E. & Grove, E. A. Members of the Wnt, Fz, and Frp gene families expressed in postnatal mouse cerebral cortex. *J Comp Neurol* **473**, 496-510, doi:10.1002/cne.20135 (2004).

146 Ahmad-Annuar, A. *et al.* Signaling across the synapse: a role for Wnt and Dishevelled in presynaptic assembly and neurotransmitter release. *J Cell Biol* **174**, 127-139, doi:10.1083/jcb.200511054 (2006).
147 Ciani, L. *et al.* Wnt7a signaling promotes dendritic spine growth and synaptic strength through Ca(2)(+)/Calmodulin-dependent protein kinase II. *Proc Natl Acad Sci U S A* **108**, 10732-10737, doi:10.1073/pnas.1018132108 (2011).
148 McLeod, F. & Salinas, P. C. Wnt proteins as modulators of synaptic plasticity. *Curr Opin Neurobiol* **53**, 90-95, doi:10.1016/j.conb.2018.06.003 (2018).
149 Lucas, F. R. & Salinas, P. C. WNT-7a induces axonal remodeling and increases synapsin I levels in cerebellar neurons. *Dev Biol* **192**, 31-44, doi:10.1006/dbio.1997.8734 (1997).
150 McLeod, F. *et al.* Wnt Signaling Mediates LTP-Dependent Spine Plasticity and AMPAR Localization through Frizzled-7 Receptors. *Cell Rep* **23**, 1060-1071, doi:10.1016/j.celrep.2018.03.119 (2018).
151 Palomer, E., Buechler, J. & Salinas, P. C. Wnt Signaling Deregulation in the Aging and Alzheimer's Brain. *Front Cell Neurosci* **13**, 227, doi:10.3389/fncel.2019.00227 (2019).
152 Seib, D. R. *et al.* Loss of Dickkopf-1 restores neurogenesis in old age and counteracts cognitive decline. *Cell Stem Cell* **12**, 204-214, doi:10.1016/j.stem.2012.11.010 (2013).
153 Folke, J., Pakkenberg, B. & Brudek, T. Impaired Wnt Signaling in the Prefrontal Cortex of Alzheimer's Disease. *Mol Neurobiol* **56**, 873-891, doi:10.1007/s12035-018-1103-z (2019).
154 Liu, C. C. *et al.* Deficiency in LRP6-mediated Wnt signaling contributes to synaptic abnormalities and amyloid pathology in Alzheimer's disease. *Neuron* **84**, 63-77, doi:10.1016/j.neuron.2014.08.048 (2014).
155 Hooper, C., Killick, R. & Lovestone, S. The GSK3 hypothesis of Alzheimer's disease. *J Neurochem* **104**, 1433-1439, doi:10.1111/j.1471-4159.2007.05194.x (2008).
156 Lu, T. *et al.* REST and stress resistance in ageing and Alzheimer's disease. *Nature* **507**, 448-454, doi:10.1038/nature13163 (2014).
157 Sweeney, M. D., Sagare, A. P. & Zlokovic, B. V. Blood-brain barrier breakdown in Alzheimer disease and other neurodegenerative disorders. *Nat Rev Neurol* **14**, 133-150, doi:10.1038/nrneurol.2017.188 (2018).
158 Hardy, J. & Allsop, D. Amyloid deposition as the central event in the aetiology of Alzheimer's disease. *Trends Pharmacol Sci* **12**, 383-388, doi:10.1016/0165-6147(91)90609-v (1991).
159 Cho, C., Smallwood, P. M. & Nathans, J. Reck and Gpr124 Are Essential Receptor Cofactors for Wnt7a/Wnt7b-Specific Signaling in Mammalian CNS Angiogenesis and Blood-Brain Barrier Regulation. *Neuron* **95**, 1056-1073 e1055, doi:10.1016/j.neuron.2017.07.031 (2017).
160 Tran, K. A. *et al.* Endothelial beta-Catenin Signaling Is Required for Maintaining Adult Blood-Brain Barrier Integrity and Central Nervous System Homeostasis. *Circulation* **133**, 177-186, doi:10.1161/CIRCULATIONAHA.115.015982 (2016).

161 Jia, L., Pina-Crespo, J. & Li, Y. Restoring Wnt/beta-catenin signaling is a promising therapeutic strategy for Alzheimer's disease. *Mol Brain* **12**, 104, doi:10.1186/s13041-019-0525-5 (2019).
162 Zhang, X. *et al.* Notum is required for neural and head induction via Wnt deacylation, oxidation, and inactivation. *Dev Cell* **32**, 719-730, doi:10.1016/j.devcel.2015.02.014 (2015).
163 Perez-Palma, E. *et al.* Early Transcriptional Changes Induced by Wnt/beta-Catenin Signaling in Hippocampal Neurons. *Neural Plast* **2016**, 4672841, doi:10.1155/2016/4672841 (2016).
164 Cox, O. B. *et al.* A poised fragment library enables rapid synthetic expansion yielding the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chem Sci* **7**, 2322-2330, doi:10.1039/c5sc03115j (2016).
165 Tarver, J. E., Jr. *et al.* Stimulation of cortical bone formation with thienopyrimidine based inhibitors of Notum Pectinacetylesterase. *Bioorg Med Chem Lett* **26**, 1525-1528, doi:10.1016/j.bmcl.2016.02.021 (2016).
166 Atkinson, B. N. *et al.* Discovery of 2-phenoxyacetamides as inhibitors of the Wnt-depalmitoleating enzyme NOTUM from an X-ray fragment screen. *Medchemcomm* **10**, 1361-1369, doi:10.1039/c9md00096h (2019).
167 Rankovic, Z. CNS drug design: balancing physicochemical properties for optimal brain exposure. *J Med Chem* **58**, 2584-2608, doi:10.1021/jm501535r (2015).
168 Mahy, W. *et al.* Screening of a Custom-Designed Acid Fragment Library Identifies 1-Phenylpyrroles and 1-Phenylpyrrolidines as Inhibitors of Notum Carboxylesterase Activity. *J Med Chem* **63**, 9464-9483, doi:10.1021/acs.jmedchem.0c00660 (2020).
169 Nardini, M. & Dijkstra, B. W. Alpha/beta hydrolase fold enzymes: the family keeps growing. *Curr Opin Struct Biol* **9**, 732-737, doi:10.1016/s0959-440x(99)00037-8 (1999).
170 Zhao, Y. *et al.* Structural characterization of melatonin as an inhibitor of the Wnt deacylase Notum. *J Pineal Res* **68**, e12630, doi:10.1111/jpi.12630 (2020).
171 Brommage, R. *et al.* NOTUM inhibition increases endocortical bone formation and bone strength. *Bone Res* **7**, 2, doi:10.1038/s41413-018-0038-3 (2019).
172 Suciu, R. M., Cognetta, A. B., 3rd, Potter, Z. E. & Cravatt, B. F. Selective Irreversible Inhibitors of the Wnt-Deacylating Enzyme NOTUM Developed by Activity-Based Protein Profiling. *ACS Med Chem Lett* **9**, 563-568, doi:10.1021/acsmedchemlett.8b00191 (2018).
173 Singh, J., Petter, R. C., Baillie, T. A. & Whitty, A. The resurgence of covalent drugs. *Nat Rev Drug Discov* **10**, 307-317, doi:10.1038/nrd3410 (2011).
174 Ghosh, A. K., Samanta, I., Mondal, A. & Liu, W. R. Covalent Inhibition in Drug Discovery. *ChemMedChem* **14**, 889-906, doi:10.1002/cmdc.201900107 (2019).
175 Jung, Y. S. & Park, J. I. Wnt signaling in cancer: therapeutic targeting of Wnt signaling beyond beta-catenin and the destruction complex. *Exp Mol Med* **52**, 183-191, doi:10.1038/s12276-020-0380-6 (2020).

176 Zhong, Y. *et al.* Tankyrase Inhibition Causes Reversible Intestinal Toxicity in Mice with a Therapeutic Index < 1. *Toxicol Pathol* **44**, 267-278, doi:10.1177/0192623315621192 (2016).
177 Madan, B. & Virshup, D. M. Targeting Wnts at the source--new mechanisms, new biomarkers, new drugs. *Mol Cancer Ther* **14**, 1087-1094, doi:10.1158/1535-7163.MCT-14-1038 (2015).
178 Canal, F. *et al.* Generation of Mice with Hepatocyte-Specific Conditional Deletion of Notum. *PLoS One* **11**, e0150997, doi:10.1371/journal.pone.0150997 (2016).
179 Pentinmikko, N. *et al.* Notum produced by Paneth cells attenuates regeneration of aged intestinal epithelium. *Nature* **571**, 398-402, doi:10.1038/s41586-019-1383-0 (2019).
180 Lien, W. H. *et al.* In vivo transcriptional governance of hair follicle stem cells by canonical Wnt regulators. *Nat Cell Biol* **16**, 179-190, doi:10.1038/ncb2903 (2014).
181 Nile, A. H. & Hannoush, R. N. Fatty acid recognition in the Frizzled receptor family. *J Biol Chem* **294**, 726-736, doi:10.1074/jbc.REV118.005205 (2019).
182 Bazan, J. F. & de Sauvage, F. J. Structural ties between cholesterol transport and morphogen signaling. *Cell* **138**, 1055-1056, doi:10.1016/j.cell.2009.09.006 (2009).
183 Kwon, H. J. *et al.* Structure of N-terminal domain of NPC1 reveals distinct subdomains for binding and transfer of cholesterol. *Cell* **137**, 1213-1224, doi:10.1016/j.cell.2009.03.049 (2009).
184 Li, X., Saha, P., Li, J., Blobel, G. & Pfeffer, S. R. Clues to the mechanism of cholesterol transfer from the structure of NPC1 middle lumenal domain bound to NPC2. *Proc Natl Acad Sci U S A* **113**, 10079-10084, doi:10.1073/pnas.1611956113 (2016).
185 Williams, E. H. *et al.* Dally-like core protein and its mammalian homologues mediate stimulatory and inhibitory effects on Hedgehog signal response. *Proc Natl Acad Sci U S A* **107**, 5869-5874, doi:10.1073/pnas.1001777107 (2010).
186 Yan, D. *et al.* The cell-surface proteins Dally-like and Ihog differentially regulate Hedgehog signaling strength and range during development. *Development* **137**, 2033-2044, doi:10.1242/dev.045740 (2010).
187 Kowatsch, C., Woolley, R. E., Kinnebrew, M., Rohatgi, R. & Siebold, C. Structures of vertebrate Patched and Smoothened reveal intimate links between cholesterol and Hedgehog signalling. *Curr Opin Struct Biol* **57**, 204-214, doi:10.1016/j.sbi.2019.05.015 (2019).
188 Rudolf, A. F. *et al.* The morphogen Sonic hedgehog inhibits its receptor Patched by a pincer grasp mechanism. *Nat Chem Biol* **15**, 975-982, doi:10.1038/s41589-019-0370-y (2019).
189 Beach, T. G. Alzheimer's disease and the "Valley Of Death": not enough guidance from human brain tissue? *J Alzheimers Dis* **33 Suppl 1**, S219-233, doi:10.3233/JAD-2012-129020 (2013).
190 Scearce-Levie, K., Sanchez, P. E. & Lewcock, J. W. Leveraging preclinical models for the development of Alzheimer disease therapeutics. *Nat Rev Drug Discov* **19**, 447-462, doi:10.1038/s41573-020-0065-9 (2020).

191 Voloshanenko, O. *et al.* Wnt secretion is required to maintain high levels of Wnt activity in colon cancer cells. *Nat Commun* **4**, 2610, doi:10.1038/ncomms3610 (2013).
192 Zimmerli, D., Hausmann, G., Cantu, C. & Basler, K. Pharmacological interventions in the Wnt pathway: inhibition of Wnt secretion versus disrupting the protein-protein interfaces of nuclear factors. *Br J Pharmacol* **174**, 4600-4610, doi:10.1111/bph.13864 (2017).
193 Kahn, M. Can we safely target the WNT pathway? *Nat Rev Drug Discov* **13**, 513-532, doi:10.1038/nrd4233 (2014).