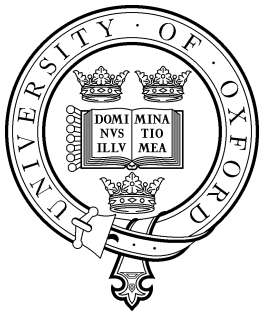
Deciphering the Molecular Mechanism of Hedgehog Signalling: A Structural Approach



Christiane Kowatsch

Exeter College

Nuffield Department of Medicine

Division of Structural Biology

University of Oxford

A thesis submitted in partial fulfillment of the requirements for the degree of *Doctor of Philosophy*

Trinity 2018

Why water boils at 100°C and methane at -161°C; why blood is red and grass is green; how sunlight makes plants grow and how living organisms have been able to evolve into ever complex forms - the answers to all these problems have come from structural analysis.

Max Perutz

Difficulties are just things to overcome, after all.

Ernest Shackleton

Abstract

Deciphering the Molecular Mechanism of Hedgehog Signalling: A Structural Approach

Christiane Kowatsch

A thesis submitted in partial fulfillment of the requirements for the degree *Doctor of Philosophy*

Exeter College, Nuffield Department of Medicine and
the Division of Structural Biology
University of Oxford, Trinity 2018

The Hedgehog (Hh) signalling pathway is key to proper embryonic development, regulation of human stem cells and tissue homoeostasis. Aberrant Hh signalling at the embryonic stage leads to developmental abnormalities, such as holoprosencephaly or, in severe cases, cyclopia. Constitutive pathway activation in adults is associated with the formation of various types of cancer. The pathway is activated by the Hh morphogen, which recognises the cell-surface receptor Patched-1 (Ptc1). The Hh-Ptc1 interaction subsequently leads to the activation of the GPCR Smoothened (Smo), which in turn activates downstream signalling and the transcription of Hh target genes. The mechanism by which Ptc1 conveys the Hh signal to Smo still remains a "black box". The goal of this thesis was to contribute to our current understanding of the underlying molecular mechanism of Hh signalling by combining structural and functional studies on two essential pathway members, Sonic Hh (Shh) and Ptc1. Firstly, Shh-specific nanobodies were identified that target Shh at two distinct binding sites. Both types of nanobody were able to inhibit Hh signalling, suggesting that Ptc1 either functions as a homodimer or by recruiting a co-receptor. Based on these results, bispecific nanobodies with enhanced inhibitory potential were designed. This study may pave the way for the development of a novel nanobody-based inhibitor for the treatment of Hh ligand-dependent cancers. Secondly, a thermostable full-length Ptc1 construct was designed that yielded a high quality sample at suitable milligram quantities for structural studies. An expression and purification protocol for successful protein extraction from mammalian cells was established. Thirdly, the first high resolution crystal structure of the Ptc1 extracellular domain 1 (ECD1) in complex with a Ptc1-specific nanobody was determined by single-wavelength anomalous dispersion of sulphur atoms (S-SAD) phasing to a resolution of 1.9 Å. The structure revealed a novel lipid-binding domain, which could either be involved in lipid transport or form an anchor point for lipidated Shh. The work presented in this thesis has advanced our understanding of the interplay and function of the morphogen Shh and its receptor Ptc1. This study may lay the basis for future drug development efforts targeting Hh in disease.

This thesis is dedicated to my family.

Acknowledgements

First and foremost, I would like to thank my supervisor, Christian Siebold, for his guidance throughout my DPhil studies, insights into the field of X-ray crystallography and ever-present enthusiasm for research. Working with him has been an incredibly valuable experience. I would also like to thank the Siebold group members: Amalie, Benjamin, Eamon, Joy, Lea, Rachel, Rebekka, Sam, and Tomas, for a memorable time in the lab, which has been both scientifically stimulating and enjoyable.

Being part of the Strubi community has been a privilege; working together with highly talented scientists, sharing their ideas and expertise. Thanks to Christina, Holly, James, Jonathan, Paul, Suzanne, and everyone else in Strubi for joining me on this journey. Many thanks to Margaret, Weixan, and Oliver for their continuous and invaluable lab support. Thanks to everyone who has attended the joint Aricescu/Jones/Siebold group meetings and who has been part of lively scientific discussions and giving valuable feedback. I would also like to acknowledge the support of Cancer Research UK who has funded this work.

I have been incredibly fortunate to have come to one of the most intercultural places imaginable: Oxford. Over the past four years, I have met wonderful people that I can now call my friends. Very special thanks to Andrea, who showed me the beauty of Japan; Clare, who took me on an adventure to New Zealand; Michael, who organised all our hiking trips across the UK, sometimes with spectacular (cloudy) views. Joy, who I ran my first 10k with; and our "colleague" Steinar, who would not let us have scallops in Iceland. Last, but not least, I would like to thank the fantastic Takopa crew! Thank you for all these special moments and your true friendship.

My scientific journey has also brought me to many other places in the world, and many marvellous people have crossed my path. I am grateful to Alexandra, Atsushi, Howard, Kyoko, Naser, Pranay, Rika, Tony, and everyone else who has made my times abroad an exceptional experience. Thank you.

None of this would have been possible without the continuous support and endless love of my family. My parents Heribert and Evi, my grandparents Mutter Lydia and Vatter Fritz, Oma Inge and Opa Gerhard, my brother Bjoern, my aunty Marianne and uncle Juergen, as well as my second family, Edith, Beate, Juergen, and Christian. Without you, I would not be where I am right now. Thank you with all my heart.

Technical acknowledgements

In alphabetical order.

- Dr B. Bishop, Dr E. Byrne, M. Jones and Prof C. Siebold for proof-reading and comments.
- Dr K. El Omari and A. Wagner (Beamline I23, Diamond Light Source, UK) for their help with S-SAD data collection and structure determination.
- Dr D. Karia and Dr S. Masiulis (Strubi, University of Oxford, UK) for their help with electron microscopy.
- Prof R. Rohatgi and L. Milenkovic (School of Medicine, Stanford University, USA) for fluorescence microscopy experiments.
- Dr A. Rudolf (Strubi, University of Oxford, UK) for her continuous advice, help with cloning and the Ptc1 ECD1-ECD2 work, as well as crystallising the Ptc1 ECD1-Nb64 complex.
- Dr R. Siebold-Schwab (Strubi, University of Oxford, UK) for testing the nanobodies in cellular Hh signalling assays.
- Prof J. Steyaert, Dr E. Pardon, and N. Buys (Center for Structural Biology, Vrije Universiteit Brussel, Belgium) for llama immunisation, cDNA library generation and phage display screening as part of the INSTRUCT Integrating Biology platform.
- Dr D. Staunton (Department of Biochemistry, University of Oxford, UK) for MALS measurements and data analysis.

Abbreviations

β-ME	β-mercaptoethanol
βOG	n-Octyl-β-D-Glucopyranoside
1D4-beads	Rho1D4 antibody-coupled sepharose beads
1D4-peptide	Rho1D4 epitope peptide (TETSQVAPA)
1D4-tag	Rho1D4 antibody epitope tag
A_{280}	Absorbance at 280 nm
BCC	Basal cell carcinoma
BCNS	Basal cell nevus syndrome
BLI	Bio-layer interferometry
Boc	Brother of Cdo
bRIL	Thermostabilised apocytochrome b_{562}RIL
CCD	Charge-coupled device
Cdo	Cell adhesion molecule-related/down regulated by oncogenes
CDR	Complementarity-determining region
CHS	Cholesteryl hemisuccinate
CMC	Critical micelle concentration
Cryo-EM	Cryo-electron microscopy
DDM	n-Dodecyl-β-D-Maltopyranoside
Dhh	Desert Hedgehog
Disp	Dispatched
DLS	Diamond Light Source
DMEM	Dulbecco's Modified Eagle's Medium
DMNG	Decyl Maltose Neopentyl Glycol
DNA	Deoxyribonucleic acid
ECD	Extracellular domain
EndoF1	Endoglycosidase F1
FCS	Fetal calf serum
FSEC	Fluorescent size-exclusion chromatography
Gas1	Growth arrest specific 1
Gli	Glioma-associated oncogene
GPCR	G-protein coupled receptor

HEK293 cells	Human embryonic kidney cells 293
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hh	Hedgehog
Hhat	Hedgehog acyltransferase
Hhip	Hedgehog-interacting protein
HPE	Holoprosencephaly
ICL	Intracellular loop
Ihh	Indian Hedgehog
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
K_D	Dissociation constant
kDa	Kilodalton
Kif7	Kinesin-like protein 7
LB	Lysogeny broth
LBD	Ligand binding domain
LCP	Lipidic cubic phase
MES	2-(N-Morpholino)ethanesulfonic acid hemisodium salt
MOPS	(3-(N-morpholino)propanesulfonic acid)
MP	Membrane protein
MR	Molecular replacement
mV	monoVenus
Nb	Nanobody
NEAA	Non-essential amino acids
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NPC1	Niemann-Pick C1
OD_{600}	Optical density at 600 nm
OGNG	Octyl Glucose Neopentyl Glycol
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEI	Polyethylenimine
pelB	Pectate lyase B

PEM	Protein Expression Medium
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase
Ptc1	Patched-1
Ptc2	Patched-2
RMSD	Root mean square deviation
RTK	Receptor tyrosine kinase
Sant-1	Smoothened antagonist
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
Shh	Sonic Hedgehog
ShhN	Sonic Hedgehog N-terminal domain
Smo	Smoothened
SOC	Super Optimal Broth
SPR	Surface plasmon resonance
S-SAD	Single wavelength anomalous dispersion of sulphur atoms
SSD	Sterol-sensing domain
SuFu	Suppressor of Fused
TAE	Tris base, acetic acid and EDTA
TB	Terrific Broth
TGF-β	Transforming growth factor-β
TM	Transmembrane
TMD	Transmembrane domain
UV	Ultra-violet
VD	Vapour diffusion

N.B. Throughout this thesis, one- or three-letter codes are used to refer to amino acid residues.

Contents

1 Introduction **1**

1.1 Developmental Biology 2

1.1.1 The importance of morphogens 2

1.1.2 The role of morphogens in *Drosophila* embryonic development 3

1.2 The Hedgehog signalling pathway 6

1.2.1 The discovery of Hh signalling 6

1.2.2 The Hh morphogen in vertebrate Hh signalling 9

1.2.3 Hh signalling: Signal transduction in vertebrates 13

1.2.4 The Hh receptor Ptc-1 16

1.3 Hh signalling and disease 20

1.3.1 Towards the treatment of Hh-related malignancies 20

1.3.2 Therapeutic nanobodies: A new strategy to target signalling pathways in disease 23

1.4 Thesis content 26

2 Materials and Methods **28**

2.1 Molecular Biology 29

2.1.1 Expression vectors 29

2.1.2 Construct design 30

2.1.3 Cloning 32

2.1.3.1 PCR and overlapping PCR 32
2.1.3.2 Restriction digest and ligation of DNA 34
2.1.3.3 Transformation of DNA 35
2.1.3.4 Purification of DNA 36
2.2 Protein Biochemistry 37
2.2.1 SDS-PAGE and Western blot 37
2.3 Protein expression in mammalian cells 38
2.3.1 Expression of secreted protein constructs 38
2.3.2 Ptc1 small-scale expression trials 38
2.3.3 Large-scale Ptc1 expression 39
2.4 Protein expression in bacterial cells 40
2.4.1 Expression of Shh 40
2.4.2 Nanobody expression 40
2.4.3 Expression of nanodisc forming proteins 41
2.5 Protein purification from mammalian cells 42
2.5.1 Purification of secreted protein constructs 42
2.5.2 Purification of Ptc1 from small-scale expressions 44
2.5.3 Large-scale Ptc1 purification 44
2.6 Protein purification from bacterial cells 46
2.6.1 Shh purification 46
2.6.2 Purification of nanobodies 46
2.6.3 Purification of nanodisc forming proteins 47
2.7 Biophysical characterisation 48
2.7.1 Protein thermostability assay 48
2.7.2 Multi-angle light scattering (MALS) 48
2.7.3 Bio-layer interferometry 49

2.7.3.1 Sample preparation 50
2.7.3.2 Experimental setup 51
2.8 Gli transcription factor reporter assay 52
2.9 Protein X-ray crystallography 54
2.9.1 The discovery and principle of X-ray crystallography 54
2.9.2 Methods for crystal structure determination 55
2.9.3 Crystallisation techniques 56
2.9.4 Sample preparation for crystallisation in vapour diffusion . . . 57
2.9.4.1 Shh-Nb complexes 57
2.9.4.2 Ptc1 ECD & Ptc1 ECD-Nb complexes 57
2.9.4.3 Ptc1 membrane protein constructs 58
2.9.5 Sample preparation for crystallisation in LCP 58
2.9.6 Crystallisation conditions and data collection 59
2.9.6.1 Shh-Nb complexes 59
2.9.6.2 Ptc1 ECD1-Nb64 complexes 60
2.9.7 Structure determination, refinement and model quality 61
2.9.7.1 Shh-Nb complexes 62
2.9.7.2 Ptc1 ECD1-Nb64 complexes 63
2.10 Electron microscopy of biomolecules 64
2.10.1 Principles of EM 64
2.10.2 Specimen preparation and data collection 65
2.10.2.1 Negative stain EM 65
2.10.2.2 Cryo-EM 67
2.10.3 Data processing of cryo-EM data 68

3 Nanobodies target Shh and its receptor Ptc1 70

3.1 Abstract 71

3.2 Introduction 72
3.2.1 Nanobodies, a smaller type of antibodies 72
3.2.2 Current applications of nanobodies 74
3.3 Results 77
3.3.1 Llama immunisation 77
3.3.2 Production of nanobodies 79
3.3.3 Binding studies of nanobodies using Bio-layer interferometry . 80
3.3.3.1 Shh-specific nanobodies 81
3.3.3.2 Ptc1-specific nanobodies 86
3.3.3.3 Nanobodies that target the Shh-Ptc1 complex 90
3.4 Discussion 93

4 Nanobodies are inhibitors of Hh signalling 98

4.1 Abstract 99
4.2 Introduction 100
4.2.1 Nanobodies: From structures to inhibitors 100
4.2.2 Targeting disease with bispecific antibodies 102
4.3 Results 105
4.3.1 Preparation of Shh & Shh-Nb complexes for structural studies 105
4.3.2 Shh crystallises readily in complex with nanobodies 107
4.3.3 Nanobodies target Shh at two distinct binding sites 108
4.3.4 Shh-specific nanobodies inhibit Hh signalling in cellular studies 115
4.3.5 The design of bispecific nanobodies to target disease 117
4.3.6 Bispecific nanobodies increase the inhibition potential 120
4.4 Discussion 123

5 Characterisation of Ptc1 for structural studies **129**

5.1 Abstract 130

5.2 Introduction 131

5.2.1 (Non)native environments for membrane proteins 131

5.2.2 Structural studies of membrane proteins 134

5.3 Results 137

5.3.1 Truncation of termini improves Ptc1 expression 137

5.3.2 Removal of the PPXY motif stabilises Ptc1 146

5.3.3 Nb64 binds to Ptc1 with high affinity 153

5.3.4 BRIL, a fusion partner to improve crystallisability 163

5.3.5 Reconstitution of Ptc1 in nanodiscs for cryo-EM studies . . . 172

5.4 Discussion 180

6 Structural studies of the Ptc1 extracellular domain **184**

6.1 Abstract 185

6.2 Introduction 186

6.2.1 Lipids and their implication in Hh signalling 186

6.3 Results 189

6.3.1 Ptc1 ECD1 & ECD2 can be expressed as a fusion protein . . . 189

6.3.2 Ptc1 ECD1 crystallises in complex with Nb64 195

6.3.3 Cholesterol binds to Ptc1 ECD1 202

6.4 Discussion 205

7 Conclusion **210**

7.1 The design of a novel Hh signalling inhibitor 211

7.2 Towards a structure of full-length Ptc1 215

7.3 Ptc1 ECD1 might be involved in cholesterol transport 217

7.4 Concluding remarks 221

A Plasmid maps **223**

B Constructs **229**

C Chemical structures of detergents and lipids **232**

Bibliography **234**

List of Figures

1.1 The French flag model 3

1.2 Gene type relationships in *D. melanogaster* development 5

1.3 Consequences of *Hh* gene mutations in *Drosophila* embryos 7

1.4 Segment polarity gene expression in *Drosophila* parasegment formation 8

1.5 Hh biosynthesis in vertebrates 11

1.6 Secretion of the Shh N-terminal domain via Disp and Scube2 12

1.7 Hh signalling in vertebrates 14

1.8 Schematic representation of the Ptc1 domain architecture 17

2.1 Overview of construct design 31

2.2 Introducing deletions and bRIL insertions using overlapping PCR . . 34

3.1 Antibodies vs. Nanobodies 73

3.2 Purification of nanobodies against Shh and Ptc1 79

3.3 Identification of Shh-specific nanobodies 82

3.4 Binding studies of Shh-specific nanobodies 84

3.5 Sequence alignment of Shh-specific nanobodies 85

3.6 Identification of Ptc1-specific nanobodies 87

3.7 Binding studies of Ptc1-specific nanobodies 88

3.8 Sequence alignment of Ptc1-specific nanobodies 89

3.9 Identification and characterisation of nanobodies against the Shh-Ptc1 complex 91
3.10 Shh-Nb interactions: A binding model 94
3.11 Recognition of the Shh-Ptc1 complex by nanobodies: A binding model 96

4.1 Purification of Shh and Shh-specific nanobodies 107
4.2 Crystal structures of Shh-Nb complexes 112
4.3 Comparison of Shh-Nb70 and Shh-Nb77 complexes 113
4.4 Detailed overview of Shh-Nb70 and Shh-Nb77 interface residues . . . 114
4.5 Assessment of Gli1 mRNA levels in response to nanobody administration 116
4.6 Design and production of bispecific nanobodies 119
4.7 Influence of bispecific nanobodies on Gli1 mRNA response levels in Hh responsive cells 121
4.8 Analysis of Shh-bispecific nanobody complexes by SEC 122
4.9 Interface comparison of Shh binding partners 126

5.1 Initial screening of Ptc1: construct overview 139
5.2 Initial screening of Ptc1 constructs by FSEC: Group 1 139
5.3 Initial screening of Ptc1 constructs by FSEC: Group 2 and 3 140
5.4 Initial screening of Ptc1 constructs by FSEC: Group 4 and 5 141
5.5 Comparison of FSEC traces of Ptc1-6 and Ptc1-20 extracted in different detergents 142
5.6 Purification and analysis of Ptc1-20 in DMNG 144
5.7 Crystallisation of Ptc1-20 in vapour diffusion 145
5.8 Screening of Ptc1 ICL3 deletion mutants 147
5.9 Ptc1 mutants and their localisation in the primary cilium 149
5.10 Purification and analysis of Ptc1-20ΔICL3-5 in DMNG 152

5.11 Purification and analysis of Ptc1-20ΔICL3-5 in DDM 153
5.12 Screening of nanobodies against Ptc1-20ΔICL3-5 156
5.13 Kinetic analysis of nanobodies targeting Ptc1-20ΔICL3-5 157
5.14 Purification and analysis of Ptc1-20ΔICL3-5 in complex with Nb64 . 159
5.15 Crystallisation trials of the Ptc1-Nb64 complex in DDM at 6.5°C . . 160
5.16 Crystallisation trials of the Ptc1-Nb64 complex in DDM at 20°C . . . 161
5.17 Purification and analysis of Amphipol-exchanged Ptc1-20ΔICL3-5 in complex with Nb64 162
5.18 Negative stain of Amphipol-exchanged Ptc1-20ΔICL3-5 in complex with Nb64 163
5.19 FSEC screening of Ptc1 mutants with bRIL insertion 165
5.20 MALS of Ptc1-20ΔICL3-5-bRIL in Amphipols 167
5.21 Purification and analysis of Ptc1-20ΔICL3-5-bRIL in complex with Nb64 in DDM 168
5.22 Crystallisation of Ptc1-20ΔICL3-5-bRIL in complex with Nb64 in DDM168
5.23 Reconstitution of Ptc1-20ΔICL3-5-bRIL in complex with Nb64 using lipidic bicelles and cubic phase 170
5.24 Thermostability of Ptc1-20ΔICL3-5-bRIL in complex with Nb64 . . . 172
5.25 Analysis of MSP nanodisc formation using analytical SEC 174
5.26 Purification and analysis of Ptc1-20ΔICL3-5-mV in DDM 175
5.27 SEC analysis of MSP1D1-based nanodiscs 177
5.28 SEC analysis of MSP2N2-based nanodiscs 177
5.29 Purification and analysis of Ptc1-20ΔICL3-5-mV in nanodiscs 179

6.1 The link between Hh signalling and cholesterol-derived disease 188
6.2 Purification and analysis of Ptc1 ECD1-ECD2 190
6.3 Purification and analysis of the Shh-Ptc1 ECD1-ECD2 complex . . . 191

6.4 Purification and analysis of Ptc1 ECD1-ECD2 in complex with Nb64 193
6.5 Purification and analysis of Ptc1 ECD1-ECD2 in complex with Nb65 194
6.6 Purification and analysis of Ptc1 ECD1-ECD2 in complex with Shh and Nb66 194
6.7 Purification and analysis of the Ptc1 ECD1-Nb64 complex 196
6.8 S-SAD phased crystal structure of the Ptc1 ECD1-Nb64 complex . . 197
6.9 Comparison of the Ptc1 ECD1 and NPC1 crystal structures 200
6.10 Detailed overview of Ptc1 ECD1-Nb64 interface residues 201
6.11 Crystal structure of the Ptc1 ECD1-Nb64:CHS complex 203
6.12 Ptc1 ECD1 interactions involving CHS 204

7.1 A model to explain the inhibitory function of Shh-specific nanobodies 214
7.2 Nb64 collides with residues located on Ptc1 ECD2 220

A.1 Plasmid maps for Shh and nanobody constructs 224
A.2 Plasmid maps for Ptc1 ECD constructs 225
A.3 Plasmid map for Ptc1 constructs 226
A.4 Plasmid maps for BirA constructs 227
A.5 Plasmid maps for MSP constructs 228

B.1 Overview of constructs used in this thesis (1/2) 230
B.2 Overview of constructs used in this thesis (2/2) 231

C.1 Overview of detergents and lipids used for Ptc1 work 233

Chapter 1

Introduction

1.1 Developmental Biology

1.1.1 The importance of morphogens

The development of complex, multicellular organisms, originating from a single, isolated cell, is an astonishing biological phenomenon. But what are the underlying mechanisms that determine the fate of a cell, lead to pattern formation, and body segmentation? Back in 1952, A. Turing proposed a mathematical theory on how the activation of specific genes influence the development of fertilised eggs and how spatial concentrations of morphogens ("diffusible chemical substances") may induce pattern formation in early embryonic development [273]. 30 years earlier, H. Spemann and H. Mangold conducted the first biological experiment on amphibian oocytes that validated the importance of specific morphogen distributions ("cytoplasmic determinants") inside a cell. They transplanted the embryo's origin of ontogeny, the Spemann organiser and confirmed that gastrulation can now set in at two different locations [254]. In the late 1960s, L. Wolpert introduced the first tangible description of a morphogen concentration gradient system. The simplified representation, called the "French flag model" (Figure 1.1) suggests that different morphogen thresholds result in the development of distinct cellular phenotypes; different downstream factors are up- or downregulated in response to varying morphogen concentrations [288].

Axis and pattern formation is key in early embryonic development. In the 1980s and 1990s, several morphogens have been identified that act as signal inducers and control the fate of cells. These molecules have the ability to establish opposing concentration gradients at opposite poles, which leads to egg polarisation and segmentation of the embryo. Besides their function as cytoplasmic determinants, most morphogens are secreted proteins that interact with signalling receptors of neighbouring cells and subsequently activate cellular signalling pathways. Morphogens can

also guide gene expression directly and function as transcription factors that either activate or inhibit the transcription of target genes. Morphogens hold a versatile functional repertoire and are indispensable for proper embryonic development.

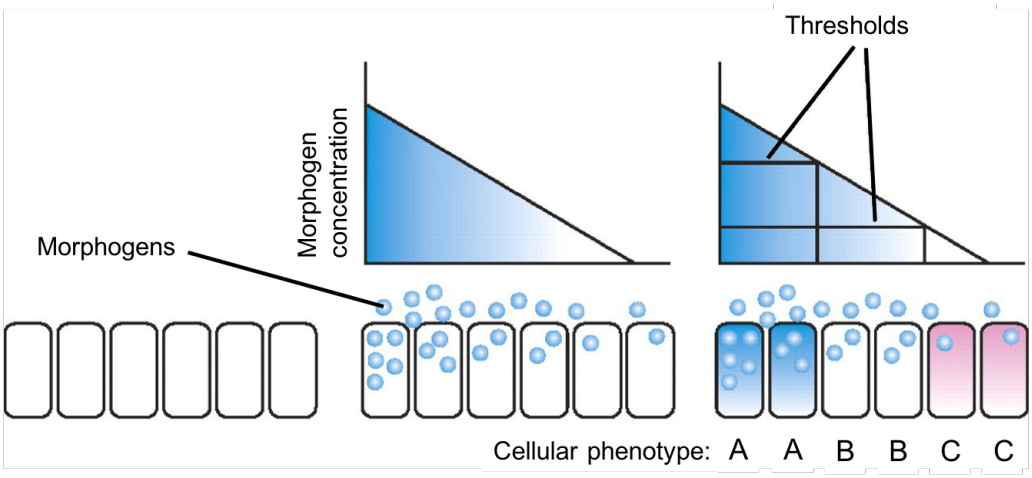


Figure 1.1: **The French flag model.** Cell fate is determined by establishing a morphogen concentration gradient across the embryo. The simplified model shows that different morphogen thresholds result in the development of distinct cellular phenotypes (A, B, or C) as different downstream factors are up- or downregulated as a response to the surrounding morphogen concentration. Figure adapted from [276].

1.1.2 The role of morphogens in *Drosophila* embryonic development

The correct distribution of morphogens in the oocyte and the embryo plays an important role for cell fate and differentiation. Three fundamental types of genes have been identified that encode for morphogens: maternal-effect genes, segmentation genes and homeotic selector genes [8]. These genes act in hierarchical order and are switched on or off at specific times during embryonic development. Incorrect gene expression at any point of ontogenesis can lead to mutations, which gives rise to deformed or nonviable embryos.

Maternal-effect genes translate into the first set of morphogens and are responsible for the polarisation of the egg before fertilisation. By directing the mRNA to the appropriate position and fine-tuning its concentration across the ovum, the anterior-posterior (longitudinal) axis is established. In 1988, W. Driever and C. Nuesslein-Volhart reported bicoid (bcd) as the first morphogen in *Drosophila melanogaster* (*D. melanogaster*). It was demonstrated that bcd mRNA is localised alongside microtubuli at the anterior pole of the egg [80]. The concentration of bcd mRNA decreases gradually towards the posterior end and is non-detectable in the last 30% of the egg cytoplasm [79]. Nanos (nos) mRNA builds the opposing concentration gradient to bcd and is localised at the posterior end of the embryo [162]. Together with the mRNA of other maternal-effect genes, bcd and nos form the anterior and posterior pole plasm. After fertilisation, the mRNAs are translated into proteins which maintain similar concentration gradients [256].

The spatial protein pattern formation is extended by a group of segmentation genes (gap-genes, pair-rule genes, segment polarity genes) and homeotic selector genes. As for maternal-effect genes, the concentration, time and duration of their expression is important for cell differentation and full segmentation of the *Drosophila* embryo (Figure 1.2). Similar rules for the establishment of a morphogen concentration gradient are applied for the formation of the dorsal-ventral axis of the embryo. After hatching, three thoracic and eight abdominal segments form the full body of a normal larvae [205].

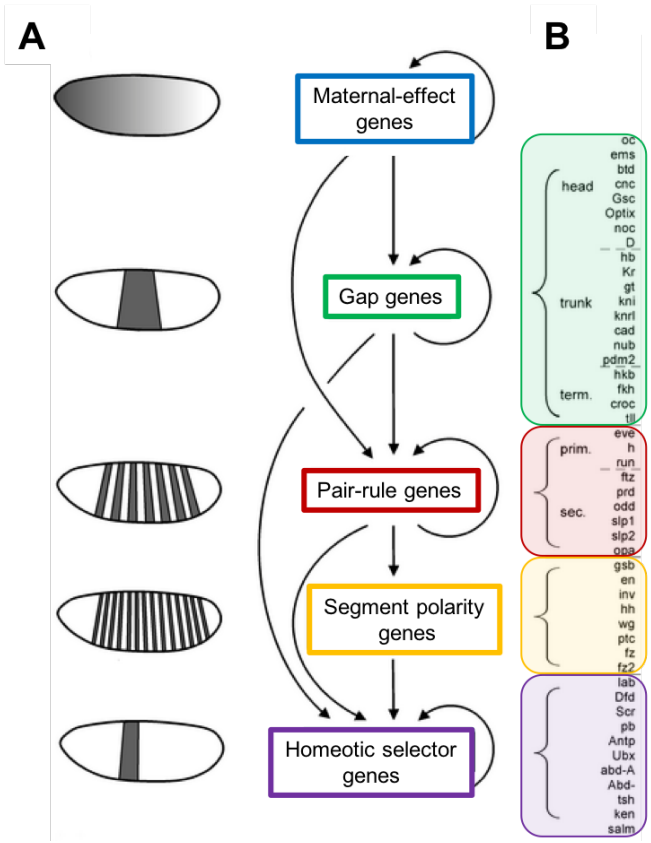


Figure 1.2: **Gene type relationships in *D. melanogaster* development. A**, Maternal-effect genes form concentration gradients along the anterior-posterior axis. Gap genes interpret the gradients and are expressed in broad, irregular stretches across the embryo. Tissue patterning is specified by the expression of pair-rule genes and segment polarity genes in seven and fourteen uniform streaks, respectively. Segment identity is defined by homeotic selector genes [247]. All gene types control each other in hierarchical order (as indicated by arrows). **B**, Proteins expressed at different stages of *Drosophila* embryonic development. Figure adapted from [247].

1.2 The Hedgehog signalling pathway

1.2.1 The discovery of Hh signalling

The Hedgehog gene (*Hh*) was first described in the fruit fly *D. melanogaster* by C. Nusslein-Volhard and E. Wieschaus in 1980 [205]. *Hh* was identified as a member of the segment polarity gene family, and is expressed in fourteen uniform stripes called parasegments [205]. Mutations in the *Hh* gene and therefore abnormal Hh morphogen expression causes deletions in the segmented pattern [192, 205]. Compared to the wildtype, the mutant phenotype of homozygous larvas manifests as a shorter larval body with spiked cuticles (Figure 1.3) [205, 279]. In fact, the spiky appearance of the larval body resembles a hedgehog-like creature and thus owes the *Hh* gene its name.

Three genes play a crucial role in *D. melanogaster* parasegment formation: Engrailed (*En*), Wingless (*Wg*) and *Hh*. The *En* gene encodes for a homeodomain-containing transcription factor and is expressed in the posterior compartment of each segment [94]. It regulates the expression of Hh, a protein which is first post-translationally modified and then secreted to fulfill its function on adjacent cells [159]. The secreted Hh morphogen interacts with the cell-surface receptor Patched (Ptc), which is presented on cells anterior to Hh-expressing cells [148]. Ptc subsequently activates downstream signalling and promotes, among others, the expression of the Wg morphogen. Wg is secreted and communicates with Frizzled (Fzd), a cell-surface receptor of Hh-expressing cells, which in turn maintains the expression of En (Figure 1.4). This elegant feedback loop of cell-cell communication results in the formation of fourteen parasegments during *D. melanogaster* embryonic development. In later studies, Hh signalling was also observed to take part in adult appendices formation, such as the patterning of the posterior part of imaginal discs [17, 262].

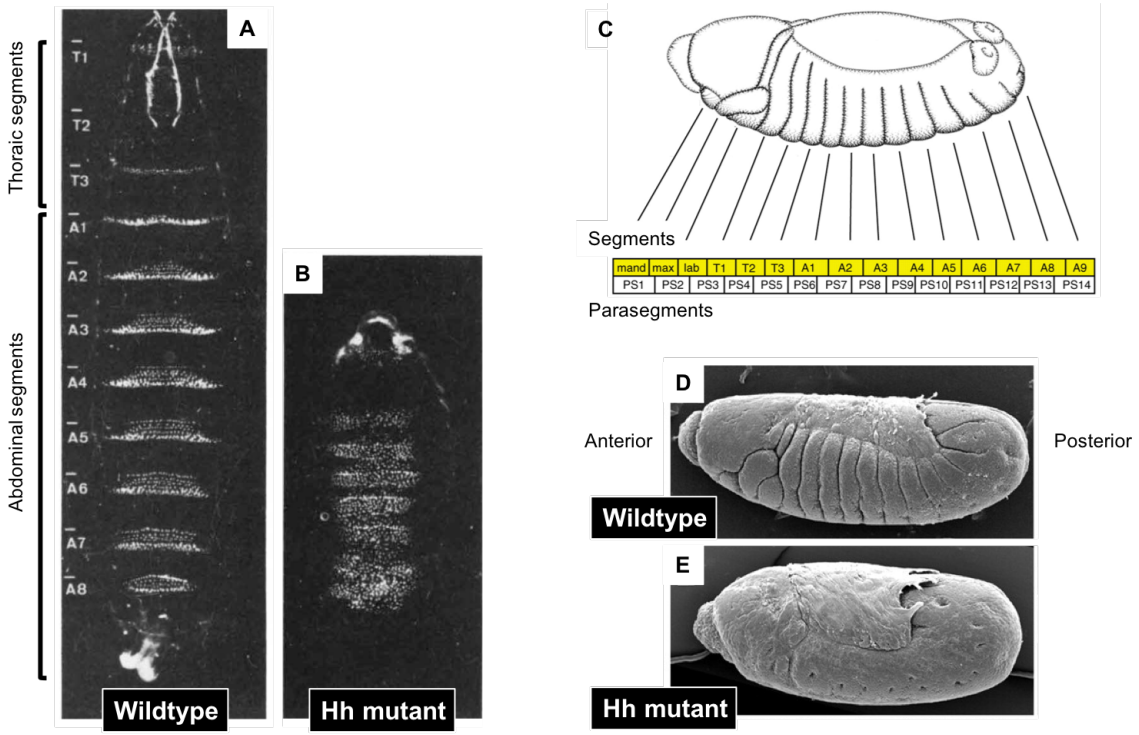


Figure 1.3: **Consequences of *Hh* gene mutations in *Drosophila* embryos. (A-B)** Analysis of cuticles presented on *Drosophila* late-stage embryos. Ventral view of **A**, wildtype (WT) embryo and **B**, embryo with homozygous *Hh* mutation. Denticles and naked cuticles alternate on WT embryos; Denticles are duplicated and naked cuticles are deleted in *Hh*-mutant embryos. **C**, Schematic representation of a *Drosophila* embryo. The embryo is subdivided into fourteen segments and parasegments. Parasegments are formed by the posterior half of one and the anterior half of the next segment [177]. Fourteen parasegments develop into three head (Ma, Mx, Lb), three thoraic (T1-T3) and eight abdominal (A1-A8) parts [177]. **(D-E)** Presentation of segmental grooves. **D**, Wildtype (WT) embryo and **E**, embryo with *Hh* mutation. *Hh* mutants fail to form segmental grooves, indicating that Hh is essential for the formation of segmental grooves [192]. Figures adapted from [177, 192, 205].

As for the dorsal-ventral patterning in *Drosophila*, the development of the neural tube is the best studied example for Hh-mediated pattern formation in vertebrates. Sonic Hedgehog (Shh), one of the three vertebrate Hh isoforms, is produced in the notochord and the floor plate of the neural tube, thus introducing ventral polarity. Neural progenitor cells that are positioned along the dorsal-ventral axis develop into distinct neuronal subtypes by translating the extracellular Shh gradient into intracellular signals such as the expression of homeodomain (HD) and basic helix-loop-helix

(bHLH) proteins [76]. Besides playing an important role for the formation of the vertebrate central nervous system, Hh signalling has been identified in the organisation of the anterior-posterior limb axis. The zone of polarising activity (ZPA) forms the posterior pole of the limb bud and also defines the region of Shh expression and secretion [83, 150, 231].

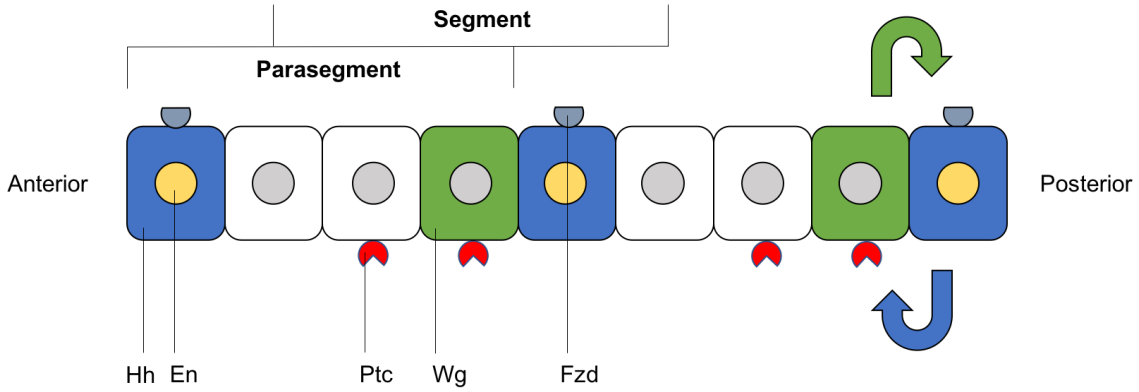


Figure 1.4: **Segment polarity gene expression in *Drosophila* parasegment formation.** Engrailed (En) initiates expression of Hedgehog (Hh) in the posterior part of each segment. The Hh receptor Patched (Ptc) is presented on cells anterior to Hh-expressing cells. Hh is secreted and activates signal transduction via Ptc by binding to its extracellular domain. As a response Wg is expressed and secreted. Wg initiates cell signalling via the Frizzled (Fzd) receptor and expression of En. En, Hh, and Fzd are expressed in cells that are located in the posterior part of each segment. Wg and Ptc are active in cells anterior to Hh-expressing cells, and form the posterior part of each parasegment. Cells are represented as rounded rectangles, nuclei as circles. Expression of Hh, En and Wg is indicated in blue, yellow and green, respectively. The cell-surface receptor Ptc is depicted as a red circular sector, Fzd as a light blue semicircle.

Hh signalling has been studied extensively during the last four decades and has been found to fulfill crucial functions across species. It is active in early embryonic development, controls cell fate and stem cell renewal and maintains the majority of adult tissues.

1.2.2 The Hh morphogen in vertebrate Hh signalling

Hh morphogens are essential, secreted signalling molecules in vertebrates and invertebrates. Remarkably, three isoforms of the Hh morphogen have been identified in mammals: Sonic Hedgehog (Shh), Desert Hedgehog (Dhh), and Indian Hedgehog (Ihh) [83]. Among them, Shh is the best studied, and is involved in crucial developmental processes such as the formation of the central nervous system and limb maturation. Sequence analysis of the three proteins has shown that mouse Shh and Ihh exhibit a higher sequence identity to each other (63%) than to Dhh (61% and 58%, respectively) [83]. Mouse Dhh in turn shows a slightly higher sequence identity to *Drosophila* Hh [83].

The Hh/Shh protein is synthesized as a 45 kDa precursor protein and translocated into the endoplasmatic reticulum (ER) for further processing [52]. Independent enzymatic reactions and post-translational modifications convert the precursor protein into the active signalling molecule. In humans, the amino (N)-terminal signal peptide is comprised of the first 23 amino acids and is removed by signal peptidases once the protein enters the ER [34]. Next, the precursor protein is autoprocessed to yield a 20 kDa N-terminal fragment Shh-N. This autocleavage step is carried out by the carboxy (C)-terminal enzymatic function of the Shh precursor protein and is accompanied by the attachment of a cholesterol moiety to the C-terminus of Shh-N. The thiol side-chain of Cys^{198} reacts with the carbonyl carbon of the preceding Gly^{197} residue to form a thioester intermediate via an intein-mediated splicing reaction [225, 237]. The nucleophilic attack of the activated 3β-hydroxyl group of cholesterol results in the rearrangement of the thioester linkage [225]. A conventional ester bond is formed by covalent coupling of cholesterol to Shh-N and the simultaneous release of Shh-C. This reaction is catalysed by the amino acid Asp^{243}, which acts as a base

and thus facilitates deprotonation of the 3β-hydroxyl group of cholesterol [113]. Positioning of cholesterol near the active site is most likely facilitated by a hydrophobic interface located on Shh-C [113]. Shortly after autoproteolysis, Shh-C is targeted for degradation via the proteasome [52]. After attachment of cholesterol, Shh-N undergoes another lipidation step. The N-terminal residue Cys^{24} is highly conserved among species, and is readily accessible for palmitic acid modification [217]. This chemical reaction is catalysed by the membrane-bound Hedgehog O-acyltransferase (Hhat) [34, 46]. Interestingly, Cys^{24} is palmitoylated on its α-amino terminal group and not the thiol side-chain [217]. The mechanistic details are undefined, but two possibilities have been proposed: (1) the α-amino group of Cys^{24} directly attacks Palmitoyl-CoA in a reaction similar to N-myristoylation or (2) a thioester is formed between the Cys^{24} thiol side-chain and Palmitoyl-CoA, followed by a rearrangement via a cyclic intermediate to yield the more stable amide [35, 217].

Dual lipidation of Shh-N is an important feature for proper protein function. Cell-based assays have shown that modified Shh-N is about thirty times more active than its non-palmiloylated equivalent [217]. However, its potency is partially reinstated when introducing a hydrophobic amino acid residue, and fully restored when substituting the palmitic acid moiety with a myristic acid residue [265]. This suggests that hydrophobicity of the N-terminus is more important than the chemical structure of the compound attached [265]. Palmitoylation together with cholesterol modification is essential for membrane localisation and subsequent distribution of Shh-N [50, 217, 220].

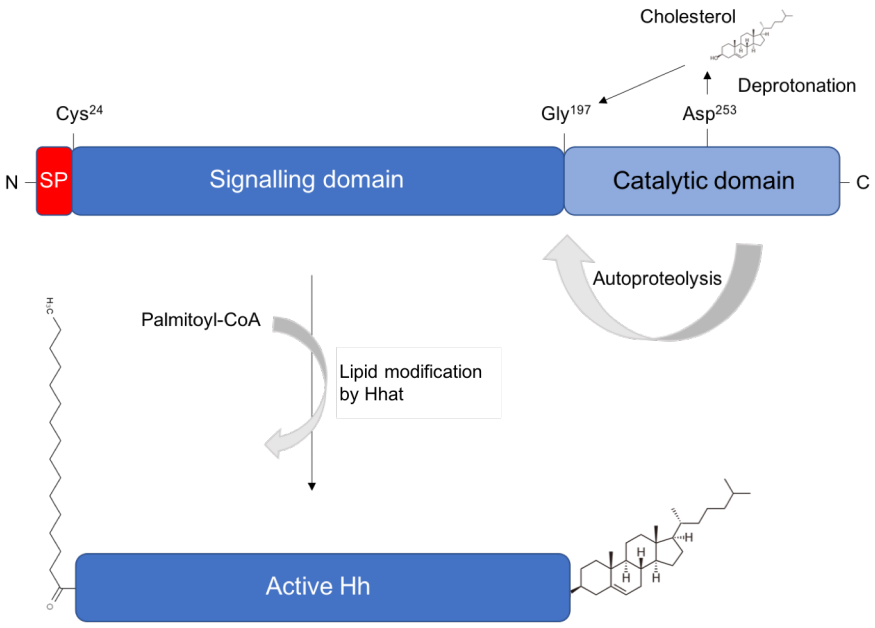


Figure 1.5: **Hh biosynthesis in vertebrates.** Hh is synthesized as a 45 kDa precursor protein. Autoproteolysis of the protein is facilitated by its C-terminal catalytic domain in an intein-like mechanism. Firstly, the thiol side-chain of Cys^{198} reacts with the carbonyl carbon of Gly^{197} to form an activated thioester intermediate. Next, Asp^{253} deprotonates the 3β-hydroxyl group of cholesterol, which then acts as a nucleophile on the thioester. This results in a conventional ester bond between the cholesterol residue and the N-terminal signalling domain of Hh. The C-terminal catalytic domain is released and degraded via the proteasome. The N-terminus of the Hh signalling domain is further modified by palmitic acid. The Hedgehog acyltransferase (Hhat) attaches the palmitoyl moiety on the α-amino group of Cys^{24} after cleavage of the signal peptide (SP).

In vertebrates, the release of membrane anchored Shh-N is facilitated by two proteins, Dispatched (Disp) and Scube2 [64]. Disp is a multipass transmembrane-spanning protein which belongs to the resistance-nodulation-division (RND) superfamily of transporters, whereas Scube2 is a secreted glycoprotein. Both proteins interact with the cholesterol moiety of Shh-N at two distinct sites, allowing the two proteins to cooperate in the release of Shh-N [271]. This ensures that the hydrophobic cholesterol residue of Shh-N is shielded from the aqueous environment at all times. Initially, monomeric Shh-N binds to Disp, lowering the energy barrier for its extraction from the plasma membrane [271]. Via an unknown mechanism, Shh-N is then transferred to Scube2. The solubility of Shh-N in complex with Scube2 is greatly

enhanced, and thus enables its release from the lipid bilayer [271]. Unsurprisingly, cells lacking Disp have shown impaired secretion of Shh-N [38, 176].

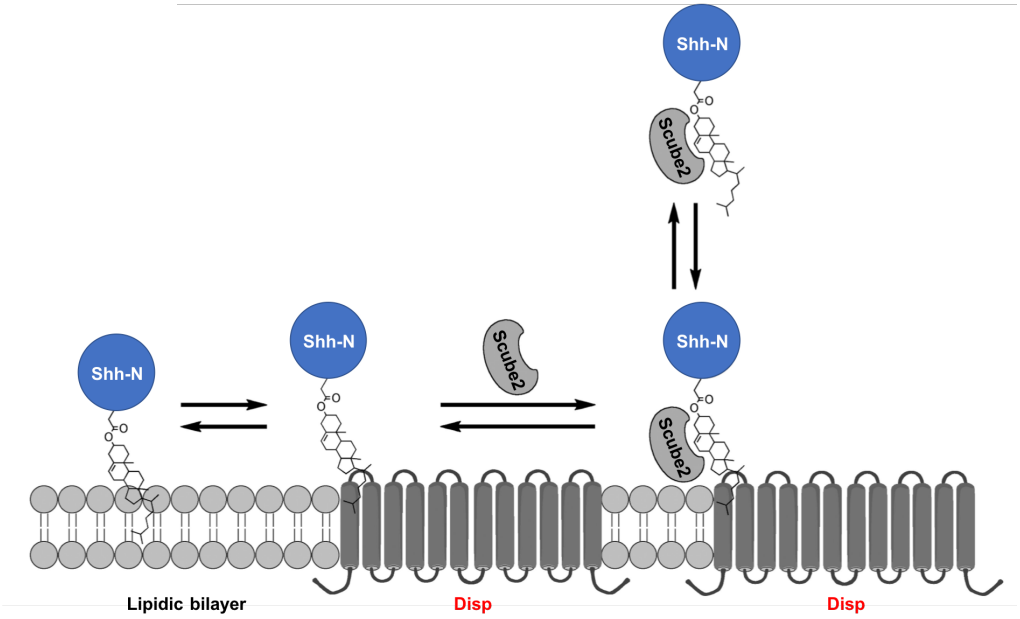


Figure 1.6: **Secretion of the Shh N-terminal domain via Dispatched and Scube2.** Shh-N is positioned at the outside of the cell membrane via its cholesterol and palmitoyl anchor. Release from the plasma membrane is initiated by the membrane protein Dispatched (Disp), and subsequent binding of Scube2. Once released, Scube2 is responsible for shielding the hydrophobic cholesterol residue of Shh-N from the aqueous extracellular environment. Figure modified after [271].

In addition to the aforementioned mechanism of Shh-N extraction from the cell membrane, three other secretion pathways have been proposed [32, 81]. M. Chen *et al.* have shown that lipidated Shh-N is able to form multimers along the plasma membrane that can be released as soluble multimeric complexes [50, 91]. Another secretion mechanism that has been suggested involves the formation of lipoprotein particles [207]. Due to their hydrophilic-lipophilic nature, lipoproteins readily integrate lipid modified Shh-N for secretion. Lastly, Shh-N multimers can be secreted from the membrane in the form of extracellular vesicles following the endocytic pathway [281]. All four mechanisms underline the importance of Shh-N lipidation and

demonstrate the versatility of Hh signalling. It is widely understood that cells utilise these mechanisms to facilitate long-range signalling of Shh-N [164, 296]. Without N- and C-terminal modification, long-range signalling of Shh-N secreting cells is greatly impaired. At the same time, membrane anchoring is important for short-range signalling of Shh-N as it limits its movements in and out of the plasma membrane.

1.2.3 Hh signalling: Signal transduction in vertebrates

Hh signalling was originally discovered in the fruit fly *D. melanogaster*, and later identified in vertebrates [205]. The centre of vertebrate Hh signalling is the primary cilium [100], a membrane protrusion formed on most mammalian cells by directed microtubule assembly. In the absence of the Hh ligand, Hh signalling is silenced. The cell-surface receptor Patched-1 (Ptc1) is delivered into the primary cilium via Caveolin-1 type vesicles (caveolae) [140], and located in the plasma membrane until the Hh signal is received [238]. Meanwhile, Ptc1 constitutively suppresses Hh signalling by inhibiting the seven transmembrane-spanning receptor Smoothened (Smo) via an unknown mechanism.

Hh signalling is initiated by binding of the Hh ligand to the extracellular domains of Ptc1 [179, 260], which in turn releases its inhibition of Smo, a member of the G protein-coupled receptor (GPCR) family. The exact mechanism of this interaction remains to be determined. However, a small molecule such as a sterol might be involved in Smo activation, and a change in Smo subcellular localisation might play an important role [82, 264]. Next, Smo is phosphorylated at multiple sites by Casein kinase 1 α (CK1α) and G protein-coupled receptor kinase 2 (GRK2), which both interact directly with the C-terminal domain of Smo [54]. Phosphorylation of Smo leads to its relocation into the primary cilium by lateral transport (transport to

the cilium via the cell membrane) [54, 73, 188, 238]. Once Smo has accumulated in cilia, Kinesin-like protein 7 (Kif7), Supressor of Fused (SuFu) and Gli proteins are placed at the tip of the primary cilium [118, 129]. In the SuFu-Gli complex, SuFu might be responsible for keeping Gli proteins silent by preventing them from entering the nucleus [160]. Upon Hh activation, Kif7 promotes dissociation of the SuFu-Gli complex and initiates cleavage of Gli proteins to yield their active forms (GliA) [160]. GliA is then translocated into the nucleus, where it is responsible for Hh target gene regulation. On the contrary, Gli repressor forms (GliR) are generated in the absence of the Hh ligand. When entering the nucleus, GliR has the ability to inhibit Hh target gene transcription.

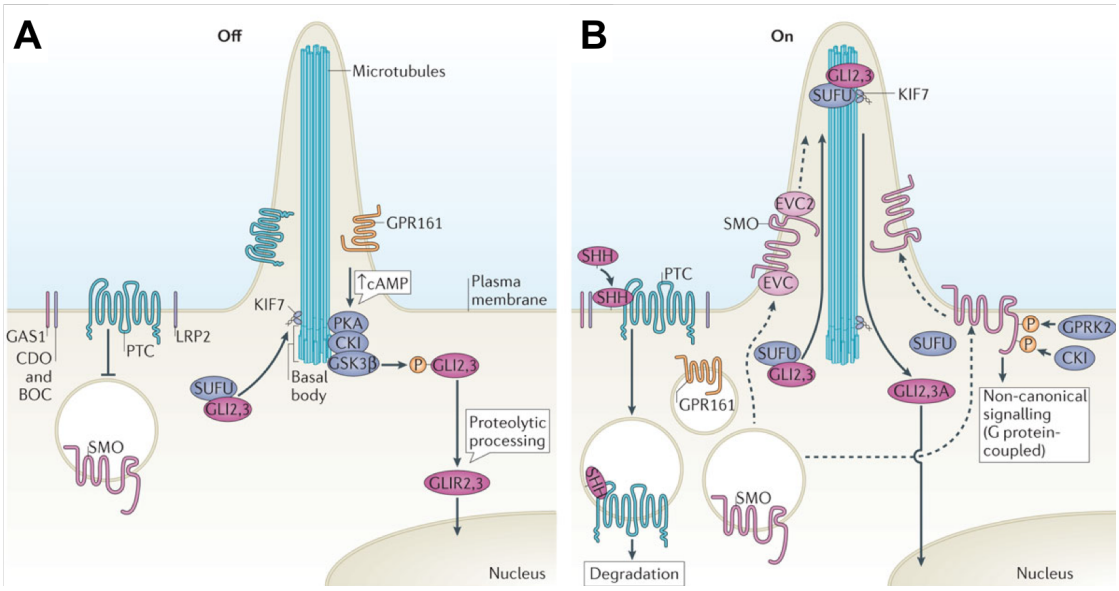


Figure 1.7: **Hh signalling in vertebrates. A**, In the absence of the Hh ligand, Hh signalling is inactivated. **B**, Binding of Hh to the cell-surface receptor Ptc releases the inhibition of the GPCR Smo. Hh signalling is activated. The signal is transmitted to the nucleus via downstream factors such as Gli. As a result, transcription of Hh target genes is stimulated. See main text for details. Figure modified after [32].

In vertebrates, three different Gli proteins are known: Gli1, Gli2 and Gli3. Gli proteins belong to the family of zinc-finger transcription factors that harbour five tandem Cys2-His2 (C2H2) domains to bind to DNA. Full-length Gli proteins act as Hh signalling activators as they possess a C-terminal activator domain [32]. Partial proteasomal degradation that generates an N-terminal fragment including the zinc-finger domains, results in Hh signalling repressor forms of the proteins [245]. Gli1 does not undergo proteasomal degradation and thus acts as an activator only [65, 180]. Gli2 mainly functions as an activator, however, it has been shown that upon phosphorylation by Protein kinase A (PKA), the repressor molecule can be formed [160, 208]. The Gli2 activator form is responsible to downregulate *Ptc* and *Gli3* gene transcription [128]. Gli3 acts predominantly as a repressor of Hh signalling, but might also function as an activator in initial response to signalling [128, 160].

Noteworthy, Ptc1 cooperates with several co-receptors to recruit Hh and to transduce the Hh signal from the cell surface to the nucleus. The co-receptors CAM-related/downregulated by oncogenes (Cdo) and Brother of Cdo (Boc) are single-pass transmembrane (TM) proteins with several Ig- and fibronectin-type III (FNIII) repeats [21]. For Cdo it has been shown that Hh makes contact via the FNIII repeats and that this interaction is calcium dependent [186]. Growth arrest specific 1 (Gas-1), another co-receptor, is anchored to the membrane via a glycosylphosphatidylinositol (GPI) linker and is formed of two glial-derived neurotrophic factor family receptor (GFR) domains [21]. Cdo, Boc, and Gas-1 all work together with Ptc1 to form a signalling complex. Interestingly, the co-receptors can compensate for the loss of one group member, but in the event of complete co-receptors deficiency, Hh signalling is abolished [10, 133]. While Cdo, Boc and Gas-1 have the ability to promote Hh signalling, Hh-interacting protein (Hhip) has been classified a negative regulator of the Hh signalling pathway [58]. Hhip consists of a cysteine-rich domain (CRD), a β-

propeller, two epidermal growth factor (EGF) domains and a C-terminal hydrophobic region [21]. Hh binding is facilitated by the third blade of the β-propeller region [27, 30].

1.2.4 The Hh receptor Ptc-1

The canonical Hh receptor Ptc1 is the primary element of the Hh signalling pathway. The *Ptc1* gene was first discovered in *D. melanogaster* in 1989 [120, 201], and it was not until a decade later, when two vertebrate *Ptc* genes were identified, *Ptc1* and *Ptc2* [41]. Sequence analysis predicts Ptc1 to be a 12 transmembrane (TM)-spanning protein (TM-I to TM-XII), with two large extracellular loops (ECD1 and ECD2) and both, N- and C- termini facing the cytoplasm (Figure 1.8). ECD1 and ECD2 are located in between TM-I and TM-II, as well as TM-VII and TM-VIII, respectively. Despite no structural information being available to date,[1] loss-of-function and gain-of-function mutations have demonstrated that the Hh ligand interacts with the Ptc1 ECDs, however, the exact binding interface remains unknown [31, 193, 216]. Further indications of how Hh might interact with Ptc1 resulted from structural studies of Shh with Hh signalling inhibitors Hhip and the Shh-specific antibody 5E1 [27, 30, 182]. The structure of Hhip has revealed three loops that interact with Shh, with an Asp^{383} residue on the second loop (L2) reaching far into the Shh zinc-binding site [27, 30]. Interestingly, Hhip L2 shares a common motif with Ptc1 ECD2, suggesting a similar binding interface [30]. The binding site of 5E1 on Shh largely overlaps with the binding site of Hhip, however, 5E1 is not making direct contact with the Shh Zn^{2+} ion [182]. As 5E1 competes with Ptc1 and Hhip for Shh binding, 5E1 might have therapeutic potential as a Hh signalling inhibitor.

[1] In the process of writing this thesis, cryo-EM structures of the Shh-Ptc1 complex have been published [101, 227, 228].

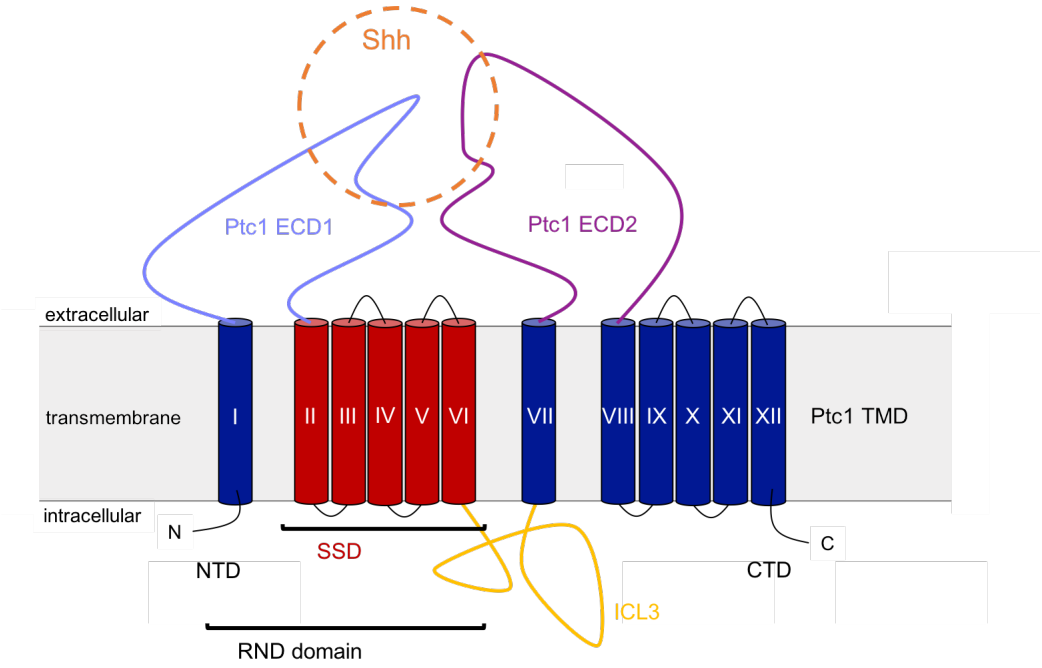


Figure 1.8: **Schematic representation of the Ptc1 domain architecture.** The receptor consists of 12 transmembrane (TM) helices (TM-I to TM-XII) and two extracellular domains (Ptc1 ECD1 and Ptc1 ECD2). N- and C-termini face the cytoplasm. Ptc1, Patched-1; Shh, Sonic Hedgehog; ECD1, extracellular domain 1; ECD2, extracellular domain 2; NTD, N-terminal domain; CTD, C-terminal domain; ICL3, intracellular loop 3; TMD, transmembrane domain; SSD, sterol-sensing domain; RND domain, resistance-nodulation-division domain.

The TM region TM-II to TM-VI can be classified as a sterol-sensing domain (SSD), a region which has been shown to interact with cholesterol [154]. It has been proposed that Ptc1 might be involved in the transport of cholesterol or its derivatives, but it is unclear whether the SSD is responsible for this function [25, 82]. The SSD was first identified in 3-Hydroxy-3-methylglutaryl coenzyme A reductase (HMGCR) and the sterol regulatory element-binding protein (SREBP)-cleavage activating protein (SCAP), which both exhibit important functions in cholesterol biosynthesis and homeostasis, respectively [98, 121]. Even though the role of the Ptc1 SSD remains to be determined, the SSD seems to be crucial for Ptc1 activity [154]. Interestingly, a mutation in the SCAP SSD leads to cholesterol insensitivity of the protein. Introducing the mutation at the equivalent position in the Ptc1 SSD (Asp^{583} to Asn)

causes embryonic lethality as Ptc1 loses its ability to inhibit Smo [154, 181]. This underlines the hypothesis that Ptc1 might regulate Smo via a small molecule such as cholesterol and/or influences Smo intracellular trafficking [261]. Moreover, E. Byrne *et al.* have determined the crystal structure of Smo with a cholesterol molecule bound to its extracellular cysteine-rich domain (CRD) [39]. Is it possible that Ptc1 provides the cholesterol for Smo? There is evidence that Ptc1 acts on Smo in a catalytic rather than a stoichiometric manner and most likely a small molecule is involved in the process [264].

TM-I to TM-VI, as well as TM-VII to TM-XII constitute a resistance-nodulation-division (RND) domain, which have classified Ptc1 as a member of the RND superfamily of transporters [117, 270]. Ptc1 and RND-type proteins have 12 TM-spanning regions and a conserved GxxxD motif, which is located in their fourth TM (TM-IV) domain [26, 107]. Interestingly, this motif is linked to transport activity of prokaryotic RND family members [26, 107]. RND transporters catalyse the efflux of drugs such as antibiotics and chemotherapeutics, as well as small molecules like cholesterol and lipids [102, 203]. Recent studies suggest that Ptc1 is involved in doxorubicin transport, a chemotherapeutic drug used to treat various types of cancer [26, 263]. Further indication that Ptc1 might have transporter function comes from its sequence homology to Niemann-Pick C1 (NPC1) protein and the bacterial RND transporter AcrB. It has been shown that NPC1 is involved in the transport of cholesterol derived from low-density lipoproteins (LDL) and can function as a lipid exporter [67, 127, 278]. Structural studies of NPC1 confirmed that cholesterol binds to the N-terminal domain (NTD), as well as the TM domain (TMD), suggesting that cholesterol first interacts with the NTD and is then handed over to the TMD to be transported across the membrane [155, 167]. Although the structure of NPC1 suggests the protein to be monomeric, structures of bacterial RND permeases, such as

the multidrug efflux transporter AcrB, have shown a trimeric assembly [102, 195, 194]. Considering previous biochemical studies and the homology to AcrB (36% sequence similarity) and RND multi-drug transporters in general, Ptc1 is believed to form a trimeric complex [92, 172].

The third intracellular loop (ICL3), located between TM-VI and TM-VII, and the C-terminal domain (CTD) have important functions in Ptc1 regulation and localisation. The CTD is targeted by E3 protein ligase Itch at Lys^{1413}, whereas the ICL3 exhibits a PPXY motif which is recognised by HECT-domain ubiquitin E3 ligases Smurf1 and Smurf2 [53, 294]. E3 ligases ubiquitinate and target Ptc1 for proteasomal degradation [294]. Once labelled, Hh-bound Ptc1 is removed from the plasma membrane by endocytosis and guided to degradation in lysosomes [238, 294]. Interestingly, Smurf proteins are part of Caveolin-1 type vesicles that transport Ptc1 into the primary cilium in the absence of the Hh ligand [140, 294]. This negative feedback loop ensures the precise and finely tuned cell signalling response to a Hh ligand gradient [294].

1.3 Hh signalling and disease

1.3.1 Towards the treatment of Hh-related malignancies

Members of the Hh signalling pathway are key players in vertebrate embryonic development, regulation of human stem-cells and maintenance of adult tissue. Aberrant signalling is often associated with severe developmental abnormalities and diseases such as cancer and neurological disorders.

Very early in embryonic development, misregulation of Hh signalling leads to developmental malfunction, such as holoprosencephaly, a cephalic disorder in which the two hemispheres of the forebrain fail to develop correctly [236]. Cyclopia is a rare and lethal form of holoprosencephaly. Defects in symmetry formation manifest in the development of a single cyclopic eye and the failure to form a nose. Instead newborn patients show a proboscis, a nose-like formation above the eye [244]. The most common cause for holoprosencephaly are point or frame-shift mutations in the *Shh* gene and thus loss of Shh morphogen activity and Hh signalling. Interestingly, cyclopamine, a chemical naturally occurring in plants of the genus *Veratrum* (for instance *Veratrum californicum* or corn lily) can cause cyclopia in lambs if ingested by the mother sheep during gestation [48]. Cyclopamine binds to the transmembrane domain of Smo which leads to Hh pathway inhibition [39, 48]. Although silencing of Hh signalling is fatal in early embryonic development, it is desirable in cancer therapy. Therefore, cyclopamine has been proposed to be suitable molecule to target cancers that originate from Smo (or upstream factor) hyperactivation [161]. However, the pharmaceutical properties of cyclopamine are unfavourable, which led to the development of other drug candidates such as cyclopamine derivatives, as well as vismodegib and sonidegib [269].

Common cancer types associated with over-activation of Hh signalling are basal cell carcinoma (BCC), a universal type of skin cancer, and medulloblastoma, a malignant primary brain tumour in children [112, 292]. Both types of cancer result from mutations in either of the two Hh receptors, Ptc1 or Smo. Ptc1 loss-of-function and Smo gain-of-function mutations lead to constitutive Hh signalling. While medulloblastoma cases are rare, with 500 affected children in the United States (US) each year, BCC accounts for more than one million cases in the same country and timespan [242]. Patients with Gorlin syndrome, also known as nevoid basal cell carcinoma syndrome (NBCCS) develop multiple BCCs during their lifetime and have a higher risk to develop medulloblastoma and rhabdomyosarcoma [242]. The inherited autosomal dominant disorder originates from mutations in the *Ptc1* gene and is found in approximately 1 in 19,000 individuals [135]. The cancers mostly appear in the face, chest, and back of the patients and usually start to occur in the third decade of life (age 20 to 29) [135].

To date, most efforts have been made to target the oncoprotein and GPCR Smo. The small molecule inhibitors vismodegib (Erivedge®) and sonidegib (Odomzo®) have both progressed to the clinic for the treatment of BCC [269]. Unfortunately, patients rapidly develop resistance towards these drugs and the cancer reappears after a few weeks of treatment [243]. Molecules have also been screened and developed to inhibit the downstream factors Gli. Arsenic trioxide (ATO) is an US Food and Drug Administration (FDA)-approved drug for the treatment of acute promyelocytic leukemia and is also currently being investigated in clinical phases I to IV with the vision to treat various types of other cancers [22, 233]. GANTs (Gli ANTagonist), such as GANT-58 and GANT-61 are also Gli1 and Gli2 inhibitors, however, these molecules are not involved in clinical trials yet [158].

The canonical Hh receptor Ptc1, which functions as a tumour suppressor and dependence receptor, is a potential drug target itself [103]. Studies suggest that Ptc1 is a member of multi-drug efflux pumps of the resistance-nodulation-division (RND)-type. Ptc1 transporter function could explain why chemotherapeutic treatments of cancer often fail [26]. To date, no Ptc1 inhibitors have been identified, however, there is great potential for these inhibitors to circumvent therapeutic resistance. Instead, anti-Ptc1 antibodies and peptides have been developed that inhibit proliferation of pancreatic cancer cells [199, 200].

In adults, ligand-dependent activation of Hh signalling, which can be caused by autocrine or paracrine Hh ligand secretion, can stimulate tumour growth in many cancer types, such as lung, stomach, pancreas, breast and prostate cancer [11, 242, 293]. This fact suggests that the Hh ligand is another drug target itself, albeit Hh inhibitors have not entered clinical trials yet. Research has mainly focused on Shh, the most characterised and potent ligand of all three vertebrate Hh proteins [215]. The antibody 5E1 and the small molecule inhibitor Robotnikinin interact with Shh directly [182, 257]. 5E1 has demonstrated reduced growth of medulloblastoma and pancreatic cancer in mouse models, whereas Robotnikinin blocked Hh signalling in human epidermal keratinocytes [233, 257]. The small molecule inhibitor RU-SKI 43 impedes Hh signalling by inhibiting the function of Hhat, the Hedgehog *O*-acyltransferase that is responsible for palmitoylation of Shh [34, 221]. RU-SKI 43 treatment restricts growth of breast and pancreatic cancer cells in *in-vitro* and animal models, respectively [233]. However, RU-SKI 43 has shown off-target cytotoxicity [235].

Finally, it is noteworthy that abnormal Shh signalling has been implicated in the development of endodermally derived human cancers which account for approximately 25% of cancer patient death [175]. This highlights the importance of targeting

the Hh signalling pathway in cancer therapy. However, treatment of cancer is not a straightforward, but a quite complex matter. Most effective treatment of Hh-derived cancers could be offered to patients by personalised medicine and the appropriate choice of anti-cancer drug combinations.

1.3.2 Therapeutic nanobodies: A new strategy to target signalling pathways in disease

Almost one third of marketed drugs are of biological origin; this number includes more than 75 (as of 2017) monoclonal antibodies (mAb) that have been approved by either the European Medicines Agency (EMA) or the US Food and Drug Administration (FDA) [4, 258]. Currently available on the market are mAbs for the treatment of Alzheimer's, autoimmune disease and cancer. The general function of mABs is to block signalling pathways or to induce apoptosis of targeted cells. A major disadvantage is their strenuous and costly manufacturing process. Moreover, their large size (∼150 kDa) impedes tissue penetration and therefore delivery to target cells. Another common problem is activity neutralisation due to rapid response of the human immune system to foreign substances [258]. In 1993, C. Hamers-Casterman and coworkers discovered a new type of antibody in *Camelus dromedarius* [114]. These antibodies consist of heavy-chains only and contain an immunogenic region (V_{HH}) which recognises the antigen. When isolated, V_{HH} regions are better known as nanobodies and are excellent biological (immunostaining, structural studies) and medical (therapeutics) tools. Nanobodies show very low immunogenicity and toxicity, and are stable over a wide pH range, which is significant for therapeutic administration [9]. Several nanobodies that have entered clinical trials are developed for intravenous, oral, nebular or intratumeral application and target autoimmune disease, cancer, hematological disease, inflammation, and respiratory disease [258]. The first therapeutic nanobody

is expected to be approved in 2018; Caplacizumab® (Ablynx) is currently in clinical phase 3 and evaluated for the treatment of thrombotic thrombocytopenic purpura (TPP) and thrombosis.[2] Caplacizumab® binds to the von Willebrand factor (vWF) and prevents its interaction with thrombocytes and therefore blood coagulation [222].

In most cases nanobodies are receptor antagonists that either prevent ligand binding or conformational change and thus intercept signal activation [145]. For instance, nanobodies currently investigated for cancer therapy target the chemokine receptor CXCR7, receptor tyrosine kinases EGFR (epidermal growth factor receptor), HGFR (hepatocyte growth factor receptor), HEK2 (human embryonal kinase 2) as well as VEGFR (vascular endothelial growth factor receptor) and soluble targets such as chemokines and HGF [145, 280]. The intravenous retention time of nanobodies for medical treatment and in particular cancer therapy is of critical importance as it correlates directly with the accumulation of the drug in the tumour tissue [145]. To increase the therapeutic potential of nanobodies, three strategies have been developed: A, Monomeric/multimeric nanobodies; B, nanobodies linked to effector domains; C, nanobodies integrated into liposomes or nanoparticles. Strategy A utilises the small size of nanobodies, which simplifies their delivery to target cells and difficult to reach antigens. Multimers such as bivalent nanobodies show increased avidity and longer retention times on target molecules [15, 145]. Therapeutic efficiency of nanobodies is enhanced by fusing the nanobodies to effector domain such as toxins or radionuclides (Strategy B) [145]. The nanobody recognises the antigen on the target cell and the toxin induces cell death. Nanobody-decorated nanoparticles that have been engineered as drug delivery systems are an example for strategy C; the drug molecules are encapsulated into the nanoparticles, which present the nanobodies on their sur-

[2]During the submission process, Caplacizumab® has been approved by the EMA and is now being sold by Novartis under the name Cablivi™ [1].

face. This elegant targeted-delivery system is highly selective and ensures a high drug dosage at the tumour site [145]. For clinical applications, nanobodies often undergo further modifications. "Humanisation" minimises autoimmune response, and the removal of enzymatic cleavage sites increases the half-life of nanobodies. Furthermore, high affinity of nanobodies to the target site is essential to prevent early renal or hepatic clearance [145].

1.4 Thesis content

The aim of this work was to decipher the molecular mechanism of Hh signalling by studying the Hh receptor Ptc1 and its ligand Shh. Both proteins fulfill important functions in the Hh signalling pathway; Shh activates signalling, whereas Ptc1 silences the pathway. As of today, the mechanism by which Ptc1 regulates Hh signalling and, in particular, its downstream target Smo remains unknown. The herein presented study could help to identify and understand the underlying principles of signal transduction. Furthermore, structural information could guide the development of therapeutics that prevent Shh from binding to Ptc1 and therefore target Hh ligand-dependent malignancies.

The underlying work of the first part of this thesis (Chapter 3, Chapter 4) guided the identification of Shh-specific nanobodies that show inhibitory effects on Hh signalling. In order to characterise a variety of Shh-Nb complexes, a combination of biophysical methods, cellular assays and X-ray crystallography was used. The comprehensive study led to the identification of two inhibitory nanobodies that target Shh at two distinct binding sites. Consequently, this work may lay the ground for the design of a novel nanobody-based Hh signalling inhibitor with therapeutic potential. Due to the sparse repertoire of known Shh-specific inhibitors, new developments targeting Hh ligand-dependent cancers are much anticipated. Also presented in Chapter 3 is the identification of Ptc1-specific nanobodies which were used in co-crystallisation experiments of various Ptc1 constructs (see below).

The second part of this thesis (Chapter 5, Chapter 6) aimed at the characterisation of the Hh receptor Ptc1. Chapter 5 discusses various protein engineering strategies that were pursued in order to design a thermostable, high yielding full-length Ptc1 construct. Preliminary data suggest that this engineered construct may

have potential for future structural studies using X-ray crystallography and cryo-EM. Chapter 6 focuses on the structural characterisation of the Ptc1 extracellular domains (ECDs). This chapter presents the crystal structure of Ptc1 ECD1 in complex with Nb64 at 1.9 Å resolution. The high resolution led to the identification of a novel hydrophobic lipid-binding pocket that potentially accommodates a molecule of a lipidic nature. These findings contribute to the hypothesis that Ptc1 functions as a cholesterol transporter and/or suggests that the Ptc1 ECD1 harbours an anchor point for lipid-modified Shh.

Chapter 2

Materials and Methods

2.1 Molecular Biology

2.1.1 Expression vectors

All plasmid maps can be found in Appendix A.

For mammalian protein expression, constructs (human Patched-1, Ptc1; human Sonic Hedgehog, Shh; bispecific nanobodies) were cloned into the pHLsec vector (Addgene) containing an N-terminal Kozak consensus sequence, followed by a signal peptide (from chicken receptor type protein tyrosine phosphatase, RPTP) [12]. Full-length Ptc1 constructs were lacking the secretion signal sequence. A C-terminal monoVenus (mV)-tag was included in constructs intended to be used for initial membrane protein expression screenings. Constructs used for bio-layer interferometry (BLI) experiments contained an AviTag^{TM} sequence for protein biotinylation, recognised by BirA biotin ligase [87]. All constructs contained either a C-terminal His_6-tag or Rho1D4 antibody epitope (1D4)-tag sequence for purification. Membrane protein sequences were cloned into pHLsec using the restriction enzymes EcoRI and KpnI (New England Biolabs), secreted proteins were cloned with AgeI and KpnI (New England Biolabs). For protein fusions, a pHLsec vector was designed with a new multiple cloning site containing Gly/Ser (G/S)-rich protein linker sequences of different length. Protein sequences were then cloned in a 2-step cloning process using a combination of AgeI/BamHI for inserting a sequence upstream of the linker, and NheI/KpnI for inserting a sequence downstream of the linker.

For bacterial protein expression, human Sonic Hedgehog (Shh) constructs were cloned into the pET22b vector containing an N-terminal lac operon sequence for inducible expression, preceding a T7 promotor sequence for transcription initiation. All constructs contained a C-terminal His_6-tag for purification. Nanobody constructs

were received in the bacterial expression vector pMESy4 from the Steyaert laboratory at Vrije Universiteit Brussels (VUB), Belgium. Constructs included a N-terminal T7 promotor sequence, followed by a pectate lyase B (pelB) sequence for periplasmic protein expression and a C-terminal His_6-tag sequence for protein purification. Plasmids (pET28a) encoding membrane scaffold proteins (MSP) with an N-terminal His_7-tag and TEV-protease cleavage site were acquired from Addgene.

2.1.2 Construct design

Unless otherwise noted, DNA sequences used in this study encode proteins of human origin. The nanobodies (Nbs) originated from camels. A table with details on constructs (e.g. domain boundaries) can be found in Appendix B.

An optimised gene (synthetic gene) of human *Ptc1* (UniProt ID: Q13635) was obtained from GeneArt® (Thermo Fisher Scientific) using their GeneOptimizer® service. For initial expression screening, 30 membrane protein (MP) constructs, with truncations at both N- and C-termini were designed (Ptc1-1 to Ptc1-30). Of these constructs, Ptc1-20 was used for structural studies (Figure 2.1, B). To improve expression and stability of the protein, an intracellular loop deletion was introduced between Ptc1 transmembrane (TM) domain VI and VII (Ptc1-20ΔICL3-5, Figure 2.1, B). The mutant was designed based on the structure of the Ptc1 homologue Niemann-Pick C1 (NPC1) reported in 2016 by X. Li *et al.* and a paper published by S. Yue *et al.* in 2014 [167, 294]. Yue and co-workers reported a PPXY motif located in Ptc1 intracellular loop 3 (ICL3) that targets Ptc1 for ubiquitin-mediated protein degradation. In addition, a crystallisation scaffold thermostabilised apocytochrome b_{562} (bRIL) was introduced to substitute the Ptc1 ICL3 sequence (Ptc1-20ΔICL3-5-bRIL, Figure 2.1, B). All constructs contained a C-terminal 1D4-tag for protein purification.

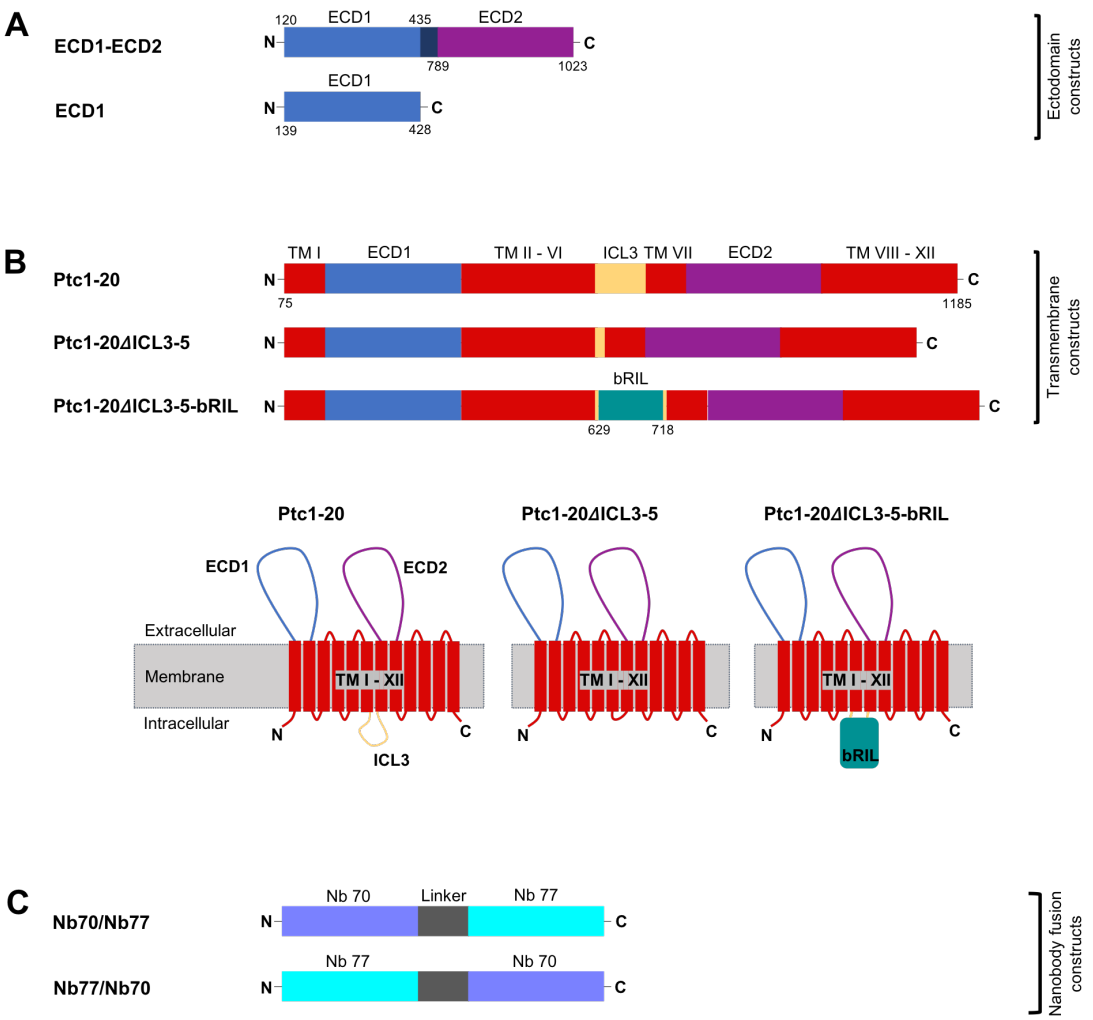


Figure 2.1: **Overview of construct design. A-C**, Schematic linear domain organisation. Domain boundaries of the proteins are indicated by numbers. ECD, extracellular domain; TM, transmembrane; ICL, intracellular linker; bRIL, thermostabilised apocytochrome b_{562}RIL; Nb, nanobody. **A**, Ptc1 extracellular domain constructs. ECD1 is highlighted in blue, ECD2 in magenta and the linker domain (GGSNGSGG) in dark blue. ECD1-ECD2 contains a C-terminal 1D4-tag, ECD1 a C-terminal His_6-tag for purification. **B**, Ptc1 transmembrane constructs. TM domains are highlighted in red, ECD1 in blue, ECD2 in magenta, ICL3 in yellow, bRIL in pine. All constructs carry a C-terminal 1D4-tag. **C**, Nanobody fusion constructs. Nb70 is shown in slate, Nb77 in cyan and the flexible linker (G/S-rich linker, variable length) in grey. All constructs with C-terminal His_6-tag for purification.

Ptc1 ectodomain (ECD) constructs, Shh, and nanobody fusion constructs (bispecific nanobodies) were cloned into pHLsec including a protein secretion signal sequence for expression in mammalian cells. Ptc1 ECD1 and ECD2 were linked via a GGSNGSGG linker (Ptc1 ECD1-ECD2) and carried a 1D4-tag for purification (Figure 2.1, A). Single loop construct Ptc1 ECD1 (Figure 2.1, A) and Shh carried a His_6-tag sequence for purification. For bispecific nanobody constructs (Figure 2.1, C), proteins were connected via a Gly/Ser (G/S)-rich linker, with linker length of 30, 60 or 90 nucleotides (10, 20, or 30 amino acid residues). Constructs carried a His_6-tag for purification.

2.1.3 Cloning

2.1.3.1 PCR and overlapping PCR

Pyrobest™ DNA polymerase kit (Takara) was used for the amplification of DNA fragments from a plasmid DNA template. PCR reactions were prepared as follows: 10 ng of template DNA was mixed with 1 μL of each forward and reverse primer (10 μM, Eurofins Genomics). 5 μL of 10 x *Pyrobest*™ buffer II, 4μL of dNTP mixture (2.5 mM of each NTP) and 0.25 μL of *Pyrobest*™ DNA polymerase (5 units/μL) were added to the DNA and adjusted to a final reaction volume of 50 μL with nuclease-free water. PCRs were carried out in a T100™ thermal cycler (BioRad) using the following settings: 30 sec at 98°C, followed by 25 cycles of 10 sec at 98°C (denaturation), 30 sec at 63°C (primer annealing) and 1 min/1 kb at 72°C (elongation). PCR reactions were completed at 72°C for 5 min. Samples were mixed with 200 μL of buffer PB (Qiagen) and purified using the QIAquick Gel Extraction Kit (Qiagen). QIAquick spin handbook (03/2001) instructions were followed, starting from step 7 (applying the sample to the QIAquick spin column). DNA was eluted in 42 μL of nuclease-free water.

Overlapping PCR method (Figure 2.2, A) was used to introduce intracellular loop (ICL) deletions and bRIL insertions into Ptc1 TM constructs. For ICL deletions, a total number of four primers was designed. The first half of the construct required a forward primer from the start of the construct (Primer 1) and a reverse primer before the start of the deletion site (Primer 2). The second part of the construct required a forward primer at the end of the deletion site (Primer 3) and a reverse primer from the end of the construct (Primer 4). In addition, Primer 2 and Primer 3 contained complementary regions to create the base for overlapping PCR. At first, two separate PCR reactions were prepared with plasmid DNA template and Primer 1 and 2 (to generate PCR fragment A), or Primer 3 and 4 (PCR fragment B). PCRs were carried out according to the procedure described above. Next, PCR fragment A and PCR fragment B were mixed together with 5 μL of 10 x *Pyrobest*^{TM} buffer II, 4μL of dNTP mixture (2.5 mM of each NTP) and 0.25 μL of *Pyrobest*^{TM} DNA polymerase (5 units/μL) and adjusted to a final reaction volume of 48 μL with nuclease-free water. Ten PCR cycles were carried out without the primers present: 1 min at 98°C, followed by 5 cycles of 10 sec at 98°C, 30 s at 63°C and 1 min/1 kb at 72°C. The PCR reactions were then kept at 4°C until 1 μL of Primer 1 and Primer 4 (10 μM) was added. PCR cycles were continued for 1 min at 98°C, followed by 30 cycles of 10 sec at 98°C, 30 s at 63°C and 1 min/1 kb at 72°C. PCR reactions were completed at 72°C for 10 min to yield PCR fragment C. Samples were purified as described earlier.

For bRIL insertions, a similar strategy was used (Figure 2.2, B) and an additional set of primers (Primer 5 and Primer 6) was required to create a bRIL DNA fragment (PCR fragment D). Primer 2 and 3 contained complementary regions to the bRIL sequence, and Primer 5 and 6 exhibited complementary regions to the Ptc1 construct. The first overlapping PCR was carried out using Primer 1 and 6 and PCR fragment A and D, creating PCR fragment E. Primer 1 and 4 were then

combined with PCR fragment B and E to yield PCR fragment F. PCR settings as described above.

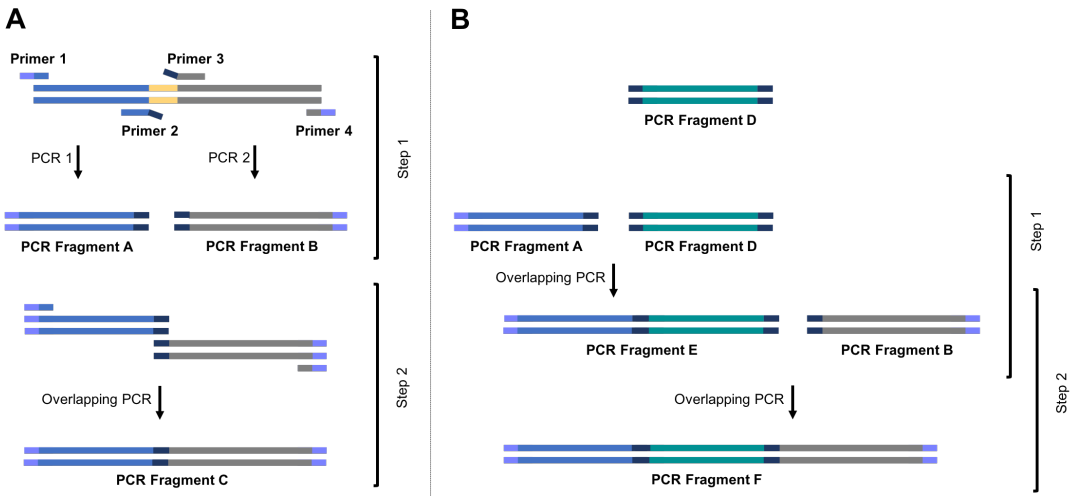


Figure 2.2: **Introducing deletions and bRIL insertions using overlapping PCR. A**, For linker deletions four primers are required. Two PCR fragments (A and B) are amplified from a DNA template, the central primers (Primer 2 and 3) introduce overlapping sequences (dark blue). The linker deletion is introduced by combining the two newly synthesised PCR fragments A and B and two terminal primers (Primer 1 and 4). **B**, The same method is applied for introducing bRIL insertions. The method described in A is repeated twice.

2.1.3.2 Restriction digest and ligation of DNA

Purified DNA fragments and 4 μg of plasmid DNA were digested separately with 1 μL of the required restriction enzymes (New England Biolabs) in 1 x CutSmart® NEBuffer. The reaction volume was adjusted to 50 μL with nuclease-free water. Restriction digests were carried out for 1 hour at 37°C. Samples were loaded on a 1% w/v agarose gel in TAE buffer and separated by applying a constant voltage. DNA fragments (inserts) were excised with a clean scalpel and purified using the QIAquick Gel Extraction Kit (Qiagen) according to the manufacturer's instructions.

Ligation reactions were performed at room temperature using NEB Quick Ligation™ Kit (New England Biolabs). Plasmid DNA and DNA inserts were mixed in a 1:3 molar ratio and nuclease-free water was added to a final volume of 10 μL. Quick ligase reaction buffer was added in a 1:1 volumetric ratio. Ligation of DNA was initiated by adding 1 μL of Quick Ligase and the reaction was allowed to proceed for 15 min at room temperature. 3.5 μL of the ligation mixture (ca. 4.5 ng of DNA) were mixed with 50 μL of *E.coli* DH5α^{TM} (Thermo Fisher) cells and incubated for 5 min on ice. Cells were heat-shocked for 30 sec at 42°C, placed on ice and 200 μL of SOC medium (Invitrogen) was added. Cells were incubated for 1 hour at 37°C, collected by centrifugation and spread on a LB-Agar plate containing the appropriate antibiotic. Plates were incubated at 37°C overnight. The next day, 14 mL of LB-medium supplemented with the required antibiotic was inoculated with one colony of *E.coli* DH5α^{TM} containing the plasmid of interest. Bacterial cultures were grown overnight at 37°C and 180 rpm. Plasmid DNA was extracted from cells using QIAprep spin miniprep kit (QIAGEN) according to the manufacturer's instructions. DNA sequences were verified by Sanger sequencing (GATC Biotech).

2.1.3.3 Transformation of DNA

Transformation of plasmid DNA for DNA purification (1) or protein expression (2) was carried out as follows. *E.coli* DH5α^{TM} (Thermo Fisher) cells (1) or *E.coli* Rosetta™2(DE3) Singles™ (Novagen) cells (2) were thawed on ice for 10 min and 0.5 μL of plasmid DNA was added. Cells were incubated for another 5 min on ice before being heat-shocked for 30 sec at 42°C. Cells were placed on ice immediately and spread onto LB-Agar plates containing the appropriate antibiotic. Plates were incubated overnight at 37°C.

2.1.3.4 Purification of DNA

QIAGEN Plasmid Giga Kit (Qiagen) was used to prepare high quantities of DNA for large scale transient transfections and protein expression in mammalian cells. The day prior to purification, a starter culture was prepared by inoculating 12 mL of LB medium (supplemented with 100 μg mL^{-1} Ampicillin) with one colony of *E.coli* DH5α^{TM} cells (Thermo Fisher) containing the plasmid of interest. After 6 hours, the culture was expanded into 3 litres of LB medium (supplemented with 100 μg mL^{-1} Ampicillin) and cells were grown overnight at 37°C and 180 rpm. Cells were collected by centrifugation at 3000 x g and DNA was purified from cells according to Qiagen's Plasmid Giga Kit instructions. This strategy typically resulted in 12- 15 mg of DNA, enough for transient transfections of 24- 30 litres of HEK293S cells in suspension or 6- 8 litres of HEK293S cells in roller bottles (Cellmaster®, VWR).

2.2 Protein Biochemistry

2.2.1 SDS-PAGE and Western blot

Protein samples were analysed by SDS-PAGE using 4- 12% Bis-Tris NuPAGE® (Life Technologies) gels. Samples were mixed with 6x SDS Loading dye prior to loading onto a gel. Soluble proteins such as nanobodies, Shh or Ptc1 ECD constructs were denatured at 95°C for 5 min prior to loading, whereas MP samples were loaded directly onto the gel to avoid protein aggregation. SDS-Page was carried out at 180 V for 40 min in either NuPage™ MES SDS or MOPS SDS running buffer (Thermo Fisher Scientific), depending on the molecular weight of the protein. BenchMark™ Protein Ladder (Thermo Fisher Scientific) was used for molecular weight estimation.

For western blotting, proteins were separated on a 4- 12% Bis-Tris NuPAGE® (Life Technologies) gel before being transferred onto a nitrocellulose membrane (Hybond ECL, GE Healthcare) at 15 V for 70 min in NuPAGE™ transfer buffer (Thermo Fisher Scientific). BenchMark™ pre-stained protein ladder was included for molecular weight estimation and to monitor transfer efficiency. After transfer, the membrane was blocked in 5% milk in PBS at 4°C overnight. Depending on the protein construct, either mouse anti-1D4 (1:5000) or mouse anti-His_6 (1:1000) antibody in 5% milk/PBS was added to the membrane and allowed to incubate for 1 hour at room temperature. The membrane was washed three times with PBS-Tween (0.1% Tween-20 in PBS), before incubation with peroxidase-coupled goat α-mouse IgG antibody (1:2000) in 5% milk/PBS for 40 min at room temperature. The membrane was washed two times with PBS-Tween (0.1% Tween-20 in PBS) and three times with PBS. Prior to digital imaging on a ChemiDoc™ instrument (BioRad), Clarity™ Western ECL blotting substrate (BioRad) was added onto the membrane and incubated for 5 min at room temperature.

2.3 Protein expression in mammalian cells

2.3.1 Expression of secreted protein constructs

Proteins expressed using this method: Ptc1 ECD1-ECD2, Ptc1 ECD1, Shh and bispecific nanobody constructs.

HEK293T cells were cultivated in T175 flasks (VWR) using DMEM (Gibco) supplemented with 10% FCS (Gibco), 1% NEAA (Gibco), 1% L-Glutamine (Gibco). Near confluency, cells were transferred into roller bottles (Cellmaster®, VWR) 3 days prior to transfection. For six roller bottles, 3 mg of DNA and 6 mL of PEI (1 mg mL^{-1}) were mixed in 300 mL of serum-free DMEM (1% NEAA, 1% L-Glutamine) and incubated for 20 min at room temperature. 50 mL of DNA:PEI mixture was added to each roller bottle containing 200 mL of DMEM (2% FCS, 1% NEAA, 1% L-Glutamine). Protein expression was carried out for 48- 72 hours post-transfection at 37°C and 5% CO_2 in the presence of kifunensine (5 μM).

2.3.2 Ptc1 small-scale expression trials

Proteins screened using this method: all Ptc1 membrane protein constructs.

HEK293T cells were cultivated in DMEM (Gibco) supplemented with 10% FCS (Gibco), 1% NEAA (Gibco) and 1% L-Glutamine (Gibco). Cells were seeded in 6-well plates 24 hours prior transfection. Transient transfections were carried out at 90-100% confluency using 4 μg of DNA and 8 μL Lipofectamine® 2000 (Invitrogen). Cells were incubated in DMEM (2% FCS, 1% NEAA, 1% L-Glutamine) at 37°C and 5% CO_2 for 48 hours. Cells were washed with PBS and collected by centrifugation.

2.3.3 Large-scale Ptc1 expression

Proteins expressed using this method: Ptc1-20, Ptc1-20ΔICL3-5 and Ptc1-20ΔICL3-5-bRIL.

9.6 litres of HEK293S-GnTI$^-$ (ATCC® CRL-3022™) cells in suspension were cultivated in PEM (Invitrogen) supplemented with 1% FCS (Gibco), 1% NEAA (Gibco) and 1% L-Glutamine (Gibco) at 37°C, 8% CO_2, 130 rpm. At a cell density of 2 x 10^{-6} cells mL^{-1}, cells were transiently transfected in 12 batches of 0.8 litres. Cells were collected at 1400 rpm for 7 min at room temperature and resuspended in 120 mL of FreeStyle 293 (Invitrogen) expression medium (1% NEAA, 1% L-Glutamine, 5 mM VPA [Sigma Aldrich]), containing 0.4 mg of plasmid DNA and 1.2 mL of PEI Max (Polysciences, 1 mg mL^{-1} in PBS). Cells were incubated at 37°C, 8% CO_2 and 160 rpm. After 4- 6 hours, 0.6 litres of PEM was added to each batch of cells and proteins were expressed for 60- 72 hours at 37°C, 8% CO_2 and 130 rpm. Cells were pelleted by centrifugation at 1100 x g for 4 min at room temperature. Cell pellets were snap-frozen in liquid nitrogen and stored at -80°C.

2.4 Protein expression in bacterial cells

2.4.1 Expression of Shh

For the sake of clarity, the Shh construct used throughout the work of this thesis refers to the non-lipidated, N-terminal domain of the protein (Cys^{24}- Gly^{197}). Cys^{24} was substituted by two isoleucine residues to imitate the naturally occurring N-terminal palmitoyl moiety.

The plasmid (pET22b) containing the sequence of Shh was transformed into *E.coli* Rosetta^{TM} 2(DE3) Singles^{TM} (Novagen). Cells were plated onto LB-Agar plates (100 μg mL^{-1} Ampicillin) and incubated at 37°C, overnight. The following day, a single colony was picked to start a 10 mL LB pre-culture (supplemented with 100 μg mL^{-1} Ampicillin) at 37°C, 200 rpm, overnight. The next day, 1 litre of LB medium was inoculated with the starter culture and the cells were grown at 37°C, 170 rpm until an optical density OD_{600} of 0.4 was reached. The temperature was reduced to 20°C and protein expression was induced at an OD_{600} of 0.7 with 0.2 mM IPTG. Cultures were grown overnight at 20°C, 170 rpm. Bacteria were collected by centrifugation at 8000 x g for 20 min at 4°C and cell pellets were stored at -80°C.

2.4.2 Nanobody expression

Two days prior to expression, plasmids with nanobody sequences were transformed into chemically competent *E.coli* WK6Su^{-} cells and spread onto LB-Agar plates supplemented with 2% glucose and 100 μg mL^{-1} Ampicillin. The next day, a single colony was picked and used to inoculate 14 mL of LB medium containing 2% glucose, 1 mM $MgCl_2$ and 100 μg mL^{-1} Ampicillin. Starter cultures were incubated for 16 hours at 37°C and 180 rpm. 2 litres of TB medium (supplemented with 0.1% glucose, 0.4%

glycerol, 1 mM $MgCl_2$ and 100 μg mL^{-1} Ampicillin) were then inoculated with the pre-culture and incubated at 37°C, 180 rpm until an optical density OD_{600} of 0.7-0.8 was reached. Protein expression was induced with 1 mM IPTG and cells were grown for 20 hours at 22°C and 180 rpm. Cells were collected by centrifugation at 5000 rpm for 15 min and cell pellets were snap-frozen in liquid nitrogen and stored at -80°C.

2.4.3 Expression of nanodisc forming proteins

Glycerol stocks of *E.coli* BL21-Gold(DE3) cells (kind gift from E. Carpenter, University of Oxford, UK) containing plasmids with sequences of MSP1D1 or MSP2N2 (Addgene) were used to start a 250 mL LB pre-culture supplemented with 25 μg mL^{-1} Kanamycin. Cells were grown overnight at 37°C, 200 rpm. The starter culture was used to prepare a 1 litre TB bacterial culture (supplemented with 0.1% glucose, 0.4% glycerol and 10 μg mL^{-1} Kanamycin) with a starting optical density OD_{600} of 0.2. Cells were incubated at 37°C, 220 rpm and protein expression was induced with 1 mM IPTG at OD_{600} of 2.5-3.0. In baffled Erlenmeyer flasks, bacteria were grown for 3- 4 hours at 28°C, 220 rpm. Cells were collected by centrifugation at 8000 x g for 15 min and cell pellets were stored at -80°C.

2.5 Protein purification from mammalian cells

2.5.1 Purification of secreted protein constructs

For purification of Ptc1 ECD1-ECD2, the medium was collected from roller bottles (Cellmaster®, VWR) three days post-transfection. Cell debris was removed by centrifugation at 3000 rpm for 15 min and by filtering the medium through a Millipore Stericup® filter unit (0.22 μm membrane, Merck). The filtered medium was then incubated for 90 min at 16°C and 80 rpm with 1 mL of Rho1D4 antibody-coupled sepharose beads (University of British Columbia, GE Healthcare) per litre of culture. The beads were collected and washed with 0.5 litre of 20 mM HEPES pH 7.5, 500 mM NaCl (wash buffer). The protein was eluted in wash buffer containing 0.5 mM Rho1D4 peptide (TETSQVAPA, Genscript) for 1 hour at 4°C and gentle shaking. The eluate was collected and loaded onto a Superdex 200 16/60 column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl. The peak fractions were pooled and concentrated to 3- 6 mg mL^{-1} using a 30 kDa MWCO filter (Millipore). Proteins were either immediately subjected to crystallisation trials or stored at -80°C for future experiments.

Ptc1 ECD1 with C-terminal His_6-tag was purified via TALON® (Clontech) metal affinity purification. The medium was collected from roller bottles (Cellmaster®, VWR) five days post-transfection, centrifuged at 3000 rpm for 15 min and filtered through a Millipore Stericup® filter unit (0.22 μm membrane, Merck). The protein sample was concentrated and diafiltrated into PBS pH 7.4 (Sigma) using a Quixstand benchtop system (GE Healthcare) connected to a 60 cm Xampler Cartridge (GE Healthcare) with a 10 kDa nominal molecular weight cutoff. The sample was incubated with 3.5 mL TALON® beads for 1 hour at 20°C and gentle shaking. The beads were collected and washed with 20 mM HEPES pH 7.5, 500 mM NaCl, followed

by washes in 20 mM HEPES pH 7.5, 500 mM NaCl, 2 mM imidazole and then 20 mM HEPES pH 7.5, 500 mM NaCl, 5 mM imidazole. The protein was eluted in 20 mM HEPES pH 7.5, 500 mM NaCl, 250 mM imidazole and subjected to SEC using a Superdex 200 16/60 column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl. The peak fractions were pooled and concentrated to 6- 6.7 mg mL^{-1} using a 10 kDa MWCO filter (Millipore). The protein was used immediately for crystallisation.

Bispecific nanobodies with C-terminal His_6-tag were purified via Ni-affinity chromatography. Five days post-transfection, the medium was collected from roller bottles (Cellmaster®, VWR). Cell debris was removed by centrifugation at 3000 rpm for 15 min after which the medium was filtered through a Millipore Stericup® filter unit (0.22 μm membrane, Merck). The protein sample was diafiltrated into PBS pH 7.4 (Sigma). The sample was incubated with 1.5 mL Ni-NTA resin (Qiagen) per litre of medium for 90 min at 14°C, 70 rpm. The beads were collected and washed with 50 mM HEPES pH 7.5, 500 mM NaCl, followed by washes in 50 mM HEPES pH 7.5, 500 mM NaCl, 10 mM imidazole. The protein was eluted in 20 mM HEPES pH 7.5, 500 mM NaCl, 250 mM imidazole and subjected to size exclusion chromatography (SEC) using a Superdex 75 16/60 column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl. The peak fractions were pooled and concentrated to 2.8- 3.1 mg mL^{-1} using a 10 kDa MWCO filter (Millipore). Proteins were flash-frozen in liquid nitrogen and stored at -80°C.

2.5.2 Purification of Ptc1 from small-scale expressions

The cell pellets were resuspended in 400 μL of ice-cold membrane buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 1% mammalian protease inhibitor) and solubilised for 90 min in 1.5% high-purity detergent (1.5% βOG 10:1 CHS, 1.5% OGNG 10:1 CHS, 1.5% DMNG 10:1 CHS, or 1.5% DDM 10:1 CHS [Anatrace]; see Appendix C for chemical structures of detergents). To remove insoluble material, the samples were centrifuged at 15.000 rpm for 10 min, before being incubated for 2 hours with 60 μL of Rho1D4 antibody-coupled sepharose beads (University of British Columbia, GE Healthcare). The beads were collected and washed twice with 1 mL of ice-cold detergent buffer (20 mM HEPES pH 7.5, 300 mM NaCl and detergent at 3 x CMC). Proteins were eluted overnight in 100 μL of detergent buffer containing 0.5 mM Rho1D4 peptide (TETSQVAPA, Genscript). The eluate was collected by centrifugation at 600 rpm for 30 sec and 20 μL were injected onto a Superose 6 3.2/300 equilibrated in 10 mM HEPES pH 7.5, 300 mM NaCl, 0.007% DMNG 10:1 CHS. To detect low quantities of fluorescently-tagged protein, the column was coupled to a Shimadzu in-line fluorescence detection system. All steps were carried out at 4°C.

2.5.3 Large-scale Ptc1 purification

Cell pellets were thawed on ice and cells were resuspended in ice-cold membrane buffer (50 mM HEPES pH7.5, 300 mM NaCl, 0.1 M L-Arginine, 10% glycerol, 1% mammalian protease inhibitor). To solubilise the proteins, a stock of 10% DDM or DMNG (in water) was added to the cell suspension to give a final detergent concentration of 1.1%. The sample was then incubated for 2 hours at 4°C before being centrifuged at 9300 x g for 30 min at 4°C. The supernatant was collected and diluted two-fold with ice-cold membrane buffer. Rho1D4 antibody-coupled sepharose

beads (University of British Columbia, GE Healthcare) were added and the sample was incubated for another 4 hours at 4°C, 70 rpm. The beads were collected and washed twice with ice-cold detergent buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 100 mM L-Arginine, 10% glycerol, detergent at 3 x CMC 10:1 CHS). The protein was eluted in detergent buffer containing 0.5 mM of Rho1D4 peptide (TETSQVAPA) overnight at 4°C. The eluate was collected and concentrated using a 100 kDa MWCO filter (Sartorius) before being loaded onto a Superose 6 10/30 column or Superose 6 Increase 10/30 column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, and 0.03% DDM 10:1 CHS (or 0.007% DMNG 10:1 CHS). For crystallisation, the peak fractions were pooled and the protein was concentrated to 2- 54 mg mL^{-1}.

For Ptc1 preparation in complex with Shh or nanobody, ligands were either added in excess before incubation with Rho1D4 antibody-coupled sepharose beads or simultaneously with the elution buffer.

2.6 Protein purification from bacterial cells

2.6.1 Shh purification

The cell pellet from 1 litre of bacterial culture was resuspended in 50 mL of lysis buffer (PBS pH 7.2, 360 mM NaCl, 10 mM β-ME, cOmplete^{TM} protease inhibitor) and sonicated on ice for 4 min at 30% amplitude applying sonication intervals of 10 sec on and 30 sec off. To remove insoluble material, the lysate was centrifuged at 18.000 rpm for 30 min at 4°C. The supernatant was diluted two-fold with lysis buffer and TALON® metal affinity resin (Clontech) was added to capture the His_6-tagged proteins. The beads were collected after 90 min at room temperature and washed with PBS pH 7.2, 360 mM NaCl and PBS pH 7.2, 360 mM NaCl, 5 mM imidazole. The protein was eluted in PBS pH 7.2, 360 mM NaCl, 250 mM imidazole and subjected to SEC on a Superdex 75 16/60 column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl. Peak fractions were pooled and concentrated to 5 mg mL^{-1} using a 10 kDa MWCO filter (Millipore). The protein was snap-frozen in liquid nitrogen and stored at -80°C.

2.6.2 Purification of nanobodies

The cell pellet from 1 litre of bacterial culture was resuspended in 23 mL of ice-cold TES buffer (0.2 M Tris-HCl pH 8.0, 0.5 mM EDTA, 0.5 M sucrose) and incubated on ice for 4 hours at 120 rpm. The lysate was diluted with 40 mL of $TES_{1/4}$ buffer (0.05 M Tris-HCl pH 8.0, 0.125 mM EDTA, 0.125 M sucrose) and incubated on ice for another 16 hours at 120 rpm. The insoluble fraction was removed by centrifugation at 6500 x g for 30 min at 4°C. Ni-NTA resin (Qiagen) was added to the supernatant and incubated for 1 hour at 4°C. The resin was washed with 50 mM HEPES pH 7.5, 500 mM NaCl and then 50 mM HEPES pH 7.5, 500 mM NaCl, 10 mM imidazole.

The protein was eluted in 20 mM HEPES pH 7.5, 150 mM NaCl, 250 mM imidazole. The sample was subsequently purified by SEC using a Superdex 75 16/60 column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl. SEC peak fractions were pooled and concentrated to 1-5 mg mL^{-1} using a 10 kDa MWCO filter (Millipore). The protein was snap-frozen in liquid nitrogen and stored in aliquots at -80°C.

2.6.3 Purification of nanodisc forming proteins

For 10 grams of cell mass, 50 mL of lysis buffer (20 mM sodium phosphate pH 7.4, 1% Triton X-100, cOmplete^{TM} protease inhibitor) was used to resuspend the pellet. Cells were lysed by mechanical disruption using a high pressure homogenizer (Constant Systems Limited) operating at 25- 28 PSI. Insoluble material was removed by centrifugation at 30.000 x g for 30 min at 4°C. Ni-NTA resin (Qiagen) was added to the supernatant and incubated for 90 min at 10°C. Beads were collected and washed with the following buffers: (1) 40 mM Tris-HCl pH 8.0, 300 mM NaCl, 1% Triton X-100; (2) 40 mM Tris-HCl pH 8.0, 300 mM NaCl, 50 mM sodium cholate; (3) 40 mM Tris-HCl pH 8.0, 300 mM NaCl; (4) 40 mM Tris-HCl pH 8.0, 300 mM NaCl, 50 mM imidazole for MSP1D1 or 20 mM imidazole for MSP2N2. The protein was eluted in 40 mM Tris-HCl pH 8.0, 300 mM NaCl, 300 mM imidazole. MSP1D1 was loaded onto a Superdex 75 16/60 column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl; MSP2N2 was loaded onto a Superdex 200 16/60 column (GE Healthcare) equilibrated in 20 mM Tris-HCl pH 7.5, 100 mM NaCl, 0.5 mM EDTA. Peak fractions were pooled and concentrated to 5 mg mL^{-1} using a 10 kDa MWCO filter (Millipore) for MSP1D1 and 30 kDa MWCO filter (Millipore) for MSP2N2. The proteins were snap-frozen in liquid nitrogen and stored at -80°C.

2.7 Biophysical characterisation

2.7.1 Protein thermostability assay

A thermostability assay is a useful tool to identify promising protein constructs for crystallisation. The more thermostable a protein, the more likely it is to crystallise. To analyse the melting point (T_M) of a sample, it is heated to different temperatures over a specific time frame and subsequently analysed on SEC. The temperature at which the protein starts to denature directly correlates with the T_M of the protein. An overlay of the SEC traces helps to estimate the T_M. In the case of Ptc1-20ΔICL3-5-bRIL in complex with Nb64, 50 μL of purified sample were used for each temperature point and heated to a temperature between 4°C- 80°C for 1 hour in a T100^{TM} thermal cycler (BioRad). Samples were then loaded onto a Superose 6 3.2/300 column equilibrated in 10 mM HEPES pH 7.5, 300 mM NaCl, 0.03% DDM 10:1 CHS and linked to a Shimadzu in-line UV detection system.

2.7.2 Multi-angle light scattering (MALS)

MALS is a method used to determine the radius of gyration and molecular mass of a protein. A light source, typically a laser is used to induce a dipole in the molecule. This dipole subsequently emits light, which is also known as light scattering. The scattered light is then detected simultaneously at multiple angles. The intensity of the detected light depends on the strength of the induced dipole (the larger the molecule, the greater the dipole) and is directly proportional to the molecular mass of the protein. In order to analyse the light scattering it is important to know the polarisability $\delta n/\delta c$ of a molecule and its 0.1% extinction coefficient ϵ at 280 nm. The polarisability is dependent on the protein buffer (solvent) and can be determined by measuring the change in the refractive index δn of the solvent with the change of

the protein concentration δc. The molecular mass can then be calculated from the intensity of the scattered light, the refractive index of the solution and the protein concentration.

For MALS experiments, detergent-solubilised protein was exchanged into Amphipols (Chapter 2.10.2) and prepared at 1 mg mL^{-1}. A five point, two-fold dilution series was prepared and samples were loaded onto a Superose 6 Increase 10/30 column (GE Healthcare) connected to a Shimadzu system with a Optilab® T-rEX refractometer and a Wyatt DAWN HELEOS 8 MALS detector (Wyatt Technologies). The column was equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl for amphipol-exchanged samples. Wyatt Technology ASTRA 6.1.2 software package was used for data analysis.

2.7.3 Bio-layer interferometry

Bio-layer interferometry (BLI) is a technique to characterise the interaction of two molecules, such as protein-protein, protein-DNA/RNA or protein-small molecules interactions. To study these biomolecular interactions, the Octet® RED384 system uses biosensors (optical fibers) that are coated with a biocompatible matrix such as streptavidin, Ni-NTA, anti-GST, etc. For instance, biotinylated proteins ("ligands") can be immobilised on a BLI biosensor with a streptavidin-coated matrix. The "analyte" (protein, DNA/RNA, small molecule) is then presented to the biosensor tip and a change in optical density can be detected upon binding of the analyte to the ligand. To detect the change in optical thickness, white light is used; the ligand and an internal reference reflect the white light which results in a specific interference pattern. Binding of the analyte shifts the pattern. This change can be detected and analysed by the Octet® system.

2.7.3.1 Sample preparation

Biotinylation is the process of covalently attaching a biotin molecule to a protein tagged with an AviTag^{TM} peptide sequence (GLNDIFEAQKIEWHE). The biotin ligase BirA regulates biotinylation, and can be used for *in-vitro* and *in-vivo* biotinylation of proteins. For BLI experiments, proteins were biotinylated while being expressed in mammalian cells. Procedures described in Chapter 2.3.1 and 2.3.3 were followed, with the modifications explained below. Proteins were purified according to the methods described in Chapter 2.5.1 and Chapter 2.5.3.

For biotinylation of the Ptc1 ECD1-ECD2 and Shh construct, cells were cultivated in roller bottles containing DMEM medium supplemented with 10% FCS, 1% NEAA, 1% L-Glutamine and 0.2 mM Biotin (Sigma). For six roller bottles, 2.4 mg of Ptc1 ECD1-ECD2 or Shh plasmid DNA (in pHLsec) and 0.6 mg of BirA biotin ligase (for ER location) plasmid DNA (kind gift of J. Elegheert, University of Oxford, UK; see Appendix A for plasmid map) were used for transfections. Proteins were expressed in DMEM supplemented with 2% FCS, 1% NEAA, 1% L-Glutamine, 0.2 mM biotin and in the presence of kifunensine for three days post-transfection.

For MP biotinylation, HEK293S GnTI^{-} cells in suspension (ATCC® CRL-3022^{TM}) were supplemented with 0.2 mM biotin (Sigma) one day prior transfection. 0.1 mg of BirA biotin ligase (for cytosolic localisation) plasmid DNA (kind gift of J. Elegheert, University of Oxford, UK; see Appendix A for plasmid map) and 0.4 mg of Ptc1-20ΔICL3-5 plasmid DNA were used to co-transfect one batch of 0.8 litres of culture. Cells were collected by centrifugation three days post-transfection.

To screen for nanobodies that specifically interact with the Shh-Ptc1 complex, biotinylated Ptc1 and non-biotinylated Shh were crosslinked using CovalX crosslinking Kit K100. Both proteins were mixed at a 2:1 molar ratio in the presence of 3

mM $CaCl_2$ and incubated on ice for 30 min. Crosslinker was added to the sample to give a final concentration of 0.2 mM and the reaction was allowed to proceed at room temperature for 2 hours before being quenched with Tris buffer (pH 8). The sample was then loaded onto a HisTrap™ High Performance column (GE Healthcare) to separate the crosslinked complex from non-crosslinked Ptc1 protein. The column was washed with 10 mM HEPES pH 7.5, 150 mM NaCl, 10 mM imidazole and eluted in 10 mM HEPES pH 7.5, 150 mM NaCl, 250 mM imidazole. The protein was concentrated using a 30 kDa MWCO filter (Millipore). The sample was subjected to SEC in 10 mM HEPES pH 7.5, 150 mM NaCl using a Superdex S200 16/60 column (GE Healthcare). The fractions were pooled and concentrated (as described above) to 0.1 mg mL^{-1}. The protein was used for BLI experiments on the same day or flash-frozen in liquid nitrogen for long term storage at -80°C.

2.7.3.2 Experimental setup

Streptavidin-coated biosensors (Pall ForteBio Dip and Read™ Streptavidin) were equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl, 3 mM $CaCl_2$, 0.005% Tween-20. Subsequently, biotinylated proteins were immobilized on the biosensors for 300 sec (60 sec baseline, 300 sec association, 60 sec baseline) in 0.1 mg mL^{-1} of sample using a Octet® RED 384 system. An initial screening was performed to assess the binding of nanobodies to the protein of interest (BLI rapid screen). For this, protein immobilised biosensors were incubated with 0.1 μM of nanobody for 300 sec at 25°C. Nanobody concentrations used for kinetic measurements were based on the quick screen results, and an eight point, two-fold dilution series was prepared accordingly. Multi-cycle kinetic measurements were carried out at 25°C for 300 sec (60 sec baseline, 300 sec association, 300 sec disassociation). Fitting of the data was carried out using the Octet® V9.0 data analysis software.

2.8 Gli transcription factor reporter assay

Gli proteins and Gli mRNA are excellent markers for Hh signalling. The Gli transcription factor reporter assay was developed to monitor Hh signalling activity by quantifying the amount of Gli1 mRNA produced by Hh responsive cells. To study the effect of a small molecule or protein on Hh signalling, cells are treated with Hh and the compound or protein of interest. Compared to Gli mRNA controls, low mRNA levels indicate inhibition of Hh signalling and vice versa.

Mouse embryonic fibroblasts (NIH/3T3 cells, kind gift from R. Rohatgi, Stanford University, USA) were seeded in 96-well plates at 20 x 10^3 cells per well and grown in DMEM supplemented with 10% FBS (Gibco), 1% NEAA (Gibco) and 1% L-Glutamine (Gibco) for 24 hours at 37°C and 8% CO_2. The medium was substituted with DMEM reduced serum medium containing 0.5% newborn calf serum (Gibco) and cells were starved for 24 hours at 37°C and 8% CO_2 to stimulate the growth of primary cilia. The following day, cells were treated with 50 nM of Shh and 0.5 μM or 5 μM of nanobody. Cells were incubated for another 24 hours at 37°C and 8% CO_2 before being collected. Cells were washed with cold PBS and prepared for real-time quantitative PCR (qPCR) using Power SYBR® Green Cells-to-CT™ Kit (Ambion by Life Technologies). The supplied lysis buffer (with RNAse inhibitor) was used to resuspend the cells. One unit of DNAse was added and the sample was incubated for 5 min at room temperature. The reaction was stopped by adding stop solution to the sample. The reaction was allowed to proceed for 2 min at room temperature. The lysate was mixed with reverse transcriptase (RT) master mix and incubated for 1 hour at 37°C and 5 min at 95°C to generate a cDNA library. A PCR cocktail was prepared by mixing Power SYBR® Green PCR Master Mix and 2 μM of Gli1 primers or 2 μM GAPDH (Glyceraldehyde 3-phosphate dehydrogenase, used as a reference

gene) primers. The cDNA sample was then mixed with the PCR cocktail in a 1:4 volumetric ratio and cDNA was amplified using the following qPCR settings: 10 min at 95°C, 40 cycles of 15 sec at 95°C and 1 min at 60°C, followed by 5 sec at 65°C and 5 sec at 95°C.

To ensure validity, biological and technical replicates were prepared for each sample. Minus-RT controls were included to ensure cDNA and not genomic DNA was amplified. No-template controls that do not contain the cell lysate were used to demonstrate that the base materials used in this experiment were not contaminated with DNA. Baseline subtraction was applied to calculate average quantification cycle values C_q. GAPDH was used as a stable reference to normalise the data.

2.9 Protein X-ray crystallography

2.9.1 The discovery and principle of X-ray crystallography

Protein X-ray crystallography is a technique used to determine the three-dimensional structure of biological macromolecules at atomic resolution. In the last 100 years, almost 30 Nobel prizes have been awarded for scientific work including X-ray crystallography. The time X-rays were discovered was in 1895 by W. Rontgen, and his discovery paved the way for X-ray crystallography. In 1912, M. von Laue postulated that crystals can diffract X-rays because crystal lattices are formed by regular, periodic arrangements of atoms (with interatomic distances of 1 Å). The same year, father and son Bragg outlined the idea of constructive interference that comes from waves that are reflected from a crystal lattice. In 1953, R. Franklin, J. Watson and F. Crick described the structure of DNA for the first time. M. Perutz and J. Kendrew then revealed the structure of myoglobin, the first protein structure determined by X-ray crystallography.

The beginning of every protein X-ray crystallography experiments lies in the formation of crystals from a purified, highly concentrated protein sample. When a focused, collimated X-ray beam interacts with the crystal, a diffraction pattern can be observed, from which an electron density map can be calculated. The pattern of diffraction stores information about the crystal packing symmetry (space group), crystal system (triclinic, monoclinic, orthorhombic, tetragonal, trigonal, hexagonal, cubic), and the size of the unit cell. The intensity of the spots in the diffraction pattern is dependent on the amplitude of the reflection and the phase angle difference of the waves diffracted by the crystal. A small phase difference leads to constructive interference, a phase difference of 180° results in complete destructive interference. While amplitudes can be calculated directly from the intensity of the spots, phases

can either be determined by isomorphous or molecular replacement. The structure factor is calculated from the amplitude and phases by using fast fourier transform. This results in an electron density map.

2.9.2 Methods for crystal structure determination

To calculate an electron density map for a given data set, the structure factor F(hkl) of each diffraction hkl is needed. The structure factor can be described mathematically as a function of the amplitudes and phases of waves diffracted by the crystal. It is possible to extract the amplitudes directly from a diffraction pattern (by taking the square root of the intensity of Bragg reflections), however, there is no straightforward approach to obtain the phases. This phenomenon is known as the 'phase problem' in X-ray crystallography. Fortunately, there are ways to overcome this difficulty. One technique that can be applied to obtain initial phases of a new data set is called molecular replacement (MR). This method can be used if a structure of high enough similarity is known. The model structure is placed into the unit cell of the new molecule by applying a rotational (for the same position) and translational (for the same orientation) function. By using the phases of the known molecule and the amplitudes of the new diffraction pattern, new structure factors can be calculated. Several cycles of refinement can take care of the initial bias towards the model structure.

A structural homologue of a protein is required when solving a structure by MR. If no homologue is available, other techniques need to be explored to overcome the phase problem of X-ray crystallography. One elegant way to acquire initial phases is to label the protein with selenomethionine or soak the crystal with heavy-atom derivatives. However, low protein yields and reduced crystal quality can hamper

the structure determination when using these methods. Sulphur single-wavelength anomalous dispersion (sulphur SAD, S-SAD) is a technique used to obtain phases for a data set from a native, non-labelled protein by making use of the anomalous diffraction of the sulphur atoms from methionine and cysteine residues. Unfortunately, the anomalous signal of sulphur atoms is weak (with an absorption at 5.02 Å) and to obtain meaningful data, beamlines with longer wavelengths of 2.5-5.0 Å and modern X-ray detectors are required. The signal-to-noise ratio can further be increased by making use of inverse beam data collection. This method divides the data collection into several small segments. Once the first segment (e.g 0°-25°) is recorded, data collection continues 180° apart (180°-205°), before going back to the end of the first segment (26°-50°) etc. Recording Friedel pairs in a shorter time frame increases the sulphur signal by reducing absorption interference but also radiation damage of corresponding wedges.

2.9.3 Crystallisation techniques

In order to determine the structure of a protein by X-ray crystallography it is a prerequisite to crystallise the protein of interest. Crystallisation in vapour diffusion is a common method used and can be set up in a sitting drop or hanging drop format. Both techniques require precipitant solutions containing various salts, precipitants and buffers with a broad range of pH values. The precipitant solution is dispensed into a reservoir and on top of the protein drop. The container is sealed and over time the precipitant and protein concentration in the drop will increase due to evaporation of water. This slow process results in an equilibrium between the precipitant in the drop and the reservoir, and aids the formation of well-ordered crystals. Temperature and additives may play a vital role for crystal formation.

Lipidic cubic phase (LCP) is a technique suitable for crystallisation of membrane proteins. To reconstitute detergent-solubilised protein in a lipidic bilayer, it is mixed with monoacylglycerol lipids such as 9.9 MAG (monoolein). Crystallisation is set up by dispensing precipitant solution on top of the protein sample before sealing the glass plate. The mesophase properties of the lipids are changed when the precipitant slowly enters the protein-lipid bolus. The protein may enrich in the phase, start nucleation and form crystals. LCP crystallisation has shown to be a valuable tool for the structure determination of numerous membrane proteins, predominantly G-protein coupled receptors (GPCRs) [40].

2.9.4 Sample preparation for crystallisation in vapour diffusion

2.9.4.1 Shh-Nb complexes

Shh at a concentration of 5 mg mL^{-1} was mixed in a 1:1 molar ratio with Shh-specific nanobody (1- 5 mg mL^{-1} depending on the protein). The sample was then incubated for 30 min on ice to promote complex formation. Crystallisation trials, using 100 nl protein sample and 100 nl reservoir solution were set up in 96-well crystallisation plates (Corning, Sigma Aldrich) using a Cartesian Technologies robot. Crystallisation plates were maintained at 20°C in a TAP Homebase storage vault and automatically imaged at regular time intervals using a Veeco imaging system.

2.9.4.2 Ptc1 ECD & Ptc1 ECD-Nb complexes

Prior to crystallisation, Ptc1 ECD1-ECD2 was treated with Endoglycosidase F1 (EndoF1) for 1 hour at room temperature. The enzyme EndoF1 cleaves simple N-linked glycosylation sites on asparagine residues, leaving a single N-acetylglucosamine moiety behind. Reducing protein heterogeneity by deglycosylation may increase the

likelihood of crystal formation. For Ptc1 ECD1-ECD2 complexes, Ptc1 was mixed with nanobody or Shh at a 1:1 molar ratio and left on ice for 30 min. The Ptc1 ECD1-Nb64 complex was formed by adding Nb64 in a 2:1 molar ratio. Excess nanobody was removed by purification via SEC. All complexes were treated with EndoF1 for 1 hour at room temperature. Crystallisation trials were set up as described for Shh-Nb complexes, with protein concentrations of 4.5- 5.5 mg mL^{-1} (Ptc1 ECD1-ECD2 complexes) or 6.0- 6.7 mg mL^{-1} (Ptc1 ECD1-Nb64). Co-crystallisation of Ptc1 ECD1-Nb64 with cholesterol hemisuccinate (CHS) was carried out with a final concentration of 10 mM CHS. Plates were maintained and monitored in a storage vault at 6.5°C or 20°C.

2.9.4.3 Ptc1 membrane protein constructs

Samples were concentrated to 1.8- 5.8 mg mL^{-1} and subjected to EndoF1 treatment for 1 hour at room temperature. 96-well 3 drop Swissci plates (Molecular Dimensions) were set up at protein:reservoir drop ratios of 1:1, 2:1 and 1:2. For a 1:1 protein:reservoir ratio, 50 nL of precipitant solution were added to 50 nL of protein sample using a Cartesian Technologies robot. Maintenance and monitoring of plates were carried out as described above.

2.9.5 Sample preparation for crystallisation in LCP

For crystallisation of Ptc1-20, Ptc1-20ΔICL3-5 and Ptc1-20ΔICL3-5-bRIL, the protein was concentrated to 10- 54 mg mL^{-1}. Crystallisation trials were either set up with the protein alone or, if purified in the presence of a nanobody, as a 1:1 complex. 9.9 MAG/cholesterol (90:10), 7.7 MAG/cholesterol (90:10) or 7.8 MAG/cholesterol (90:10) was used to reconstitute the protein in lipidic cubic phase (LCP). 9.9 MAG was heated to 42°C and mixed manually with the protein at a 2:3 ratio (w/v) using

a mechanical, coupled syringe setup. 7.7 MAG and 7.8 MAG were equilibrated at room temperature and mixed with the protein sample at a 1:1 ratio (w/v). The LCP sample was dispensed onto a 96-well LCP glass plate using a Gryphon LCP robot (ART Robbins Instruments). Each well contained 50 nL of LCP sample and 0.8 μL of precipitant solution. Plates were sealed with thin glass plates and maintained in a 4°C or 20°C incubator. LCP plates were inspected every three days under a light microscope with cross polariser.

2.9.6 Crystallisation conditions and data collection

2.9.6.1 Shh-Nb complexes

Shh-Nb69 crystals were taken out of mother liquor containing 0.1 M citric acid pH 4.0, 30% (w/v) PEG 6000; Shh-Nb70 crystals out of 10% (w/v) PEG 4000, 20% (v/v) glycerol, 0.03 M of each halide (NaF, NaBr and NaI), 0.1 M MES/imidazole pH 6.5; Shh-Nb71 crystals out of 12.5% (w/v) PEG 1000, 12.5% (w/v) PEG 3350, 12.5% (v/v) MPD, 0.03 M of each NPS ($NaNO_3$, Na_2HPO_4 and $(NH_4)_2SO_4$), 0.1 M MOPS/HEPES-Na pH 7.5; Shh-Nb77 crystals out of 2.0 M $(NH_4)_2$ tartrate, 0.1 M Na acetate, pH 4.0; Shh-Nb78 crystals from 20% (w/v) PEG 3350, 0.2 M ammonium tartrate; Shh-Nb794 crystals out of 25% (w/v) PEG 3350, 0.2 M $NH_4)_2SO_4$, 0.1 bis-Tris, pH 5.5; Shh-Nb797 crystal out of 10% (w/v) PEG 20000, 20% (v/v) PEG MME 550, 0.03 M of each NPS ($NaNO_3$, Na_2HPO_4 and $(NH_4)_2SO_4$), 0.1 M MES/imidazole pH 6.5, and Shh-Nb799 crystals out of 10% (w/v) PEG 8000, 20% (v/v) ethylene glycol, 0.03 M of each NPS ($NaNO_3$, Na_2HPO_4 and $(NH_4)_2SO_4$), 0.1 M MES/imidazole pH 6.5. Prior to flash-freezing in liquid nitrogen, crystals were treated with cryoprotectant solutions. For Shh-Nb69, Shh-Nb77, Shh-Nb78, and Shh-Nb794: 35% (v/v) ethylene glycol in mother liquor; for Shh-Nb70 and Shh-Nb71: 15% (v/v) glycerol in mother liquor; for Shh-Nb797 and Shh-Nb799: 15% (v/v) ethylene glycol in

mother liquor. Data were collected at beamline I03 at Diamond Light Source (DLS; Didcot, UK). XIA2 was used to process and scale the X-ray data [85, 136, 163].

2.9.6.2 Ptc1 ECD1-Nb64 complexes

Initial crystals of Ptc1 ECD1-NB64 were obtained in a Morpheus screen (Molecular Dimensions) columns 4 and 8. To improve crystal quality and size, crystal seeding was carried out using crystals obtained in Morpheus condition H4 (0.1 M Amino acids, 0.1 M Buffer System 1 [Imidazole, MES monohydrate (acid)] pH 6.5, 50% (v/v) Precipitant Mix 4 [25% (v/v) MPD, 25% PEG 1000, 25% (w/v) PEG 3350]). Ptc1 ECD1-Nb64 crystals from microseeding were obtained out of mother liquor containing 0.12 M ethylene glycols, 0.1 M Buffer System 2 (Sodium HEPES; MOPS acid) pH 7.5, 50% (v/v) Precipitant Mix 1 (40% (v/v) PEG 500 MME; 20% (w/v) PEG 20000) [Morpheus condition E5]. Ptc1 ECD1-Nb64:CHS crystals were taken out of 0.09 M Halogens, 0.1 M Buffer System 2 (Sodium HEPES; MOPS acid) pH 7.5, 50% (v/v) Precipitant Mix 4 (25% (v/v) MPD, 25% PEG 1000, 25% (w/v) PEG 3350) [Morpheus condition B8] or 0.12 M ethylene glycols, 0.1 M Buffer System 2 (Sodium HEPES; MOPS acid) pH 7.5, 50% (v/v) Precipitant Mix 4 (25% (v/v) MPD, 25% PEG 1000, 25% (w/v) PEG 3350) [Morpheus condition F8]. Prior to flash-freezing in liquid nitrogen, crystals were treated with 20% (v/v) ethylene glycol and 20 mM CHS in mother liquor for cryo-protection. S-SAD data were collected at the long wave-length MX beamline I23 at DLS (Didcot, UK) using an inverse beam operated at $\lambda = 2.7751$ Å [283]. XDS and XSCALE were used to process and scale the X-ray data [137]. All other data were acquired at beamlines I03 or I24 at DLS (Didcot, UK). XIA2 was used to process and scale the X-ray data [85, 136, 163].

2.9.7 Structure determination, refinement and model quality

For crystal structure determination, two main electron density maps, namely $2F_{obs}$-F_{calc} and F_{obs}-F_{calc}, are commonly used. In these maps, the observed and calculated structure factor amplitudes are represented by F_{obs} and F_{calc}, respectively. The reason to use both maps is to reduce model bias along the way. Poor quality electron density is often observed for long side-chains and flexible loop regions of the protein, but can also be affected by the solvent present in the crystal. Higher density can be achieved by applying methods such as solvent flattening, which assign a low, constant density to solvent molecules. During refinement, the electron density map is gradually improved by recalculating structure factors and phase angles until the map cannot be improved any further. During this process it is important to reassess the agreement between the experimental data and obtained model, in other words, to examine how well the crystal structure represents the observed data. Once the electron density is visualised, the polypeptide chain can be manually or automatically assigned.

R-factors and Pearson correlation coefficients (CC) are commonly used values to assess model quality after refinement. R-factors compare the calculated model to non-refined data in order to flag over-fitted data. R_{work} compares F_{obs} and F_{calc} for every single reflection of the experimental and calculated diffraction pattern. A R_{work} value of zero stands for a model that perfectly represents the data, whereas a value of 0.63 represents a random atomic array. The R_{free} value is used to cross-validate the structural model. This is necessary as the iterative process for calculating electron density maps can introduce model bias. Therefore, 5- 10% of the experimental data is removed from the data set before refinement. From this subset, the R_{free} value is calculated. An ideal, non over-fitted model has comparable to slightly higher R_{free} values than R_{work} values. The CC method is a measure for true resolution. Four CC

values exist: $CC_{1/2}$, CC^*, CC_{work}, and CC_{free}. $CC_{1/2}$ is calculated by dividing the experimental data set in two equal but random parts and comparing their intensities of reflections. By contrast, CC^* represents the CC of the full data set and is determined by $CC_{1/2}$. CC_{work} and CC_{free} derive from the correlation of observed and calculated intensities (similar to R_{work} and R_{free}). Comparison of CC_{work} and CC_{free} with CC^* is a measure on how well the model fits the observed data. Similar values indicate a good description, $CC^* < CC_{work}$ a model over-fit, and $CC^* > CC_{free}$ suggests that the model does not describe all signal observed in the experimental data. Ramachandran plot assessment is another method for validating structural models. This technique evaluates the stereochemical properties of amino acid side-chains and backbone dihedral angles. Poor stereochemistry is normally a result of polypeptide chain flexibility and disorder, but also a result of serious errors in the calculated model. Therefore, bond and torsion angles as well as bond length are regulated by most refinement programs.

2.9.7.1 Shh-Nb complexes

The structures of human Shh (PDB code: 3DIM) and Gelsonin nanobody (PDB code: 2X1O, hypervariable CDR regions omitted) were used as search models to solve the Shh-Nb complexes by MR in PHASER [184]. Iterative rounds of manual model building in Coot and refinement in REFMAC and PHENIX were used to trace the polypeptide chains of both proteins [7, 84, 196]. Stereochemical properties of the complexes were assessed with the structure-validation programme MolProbity [68]. The crystallographic and Ramachandran statistics of the data collection and processing are shown in Table 4.1.

2.9.7.2 Ptc1 ECD1-Nb64 complexes

The positions of 14 (out of 21) sulphur atoms were determined using HKL2MAP on a S-SAD data set with anomalous signal to 4 Å resolution [210, 250]. Phasing with the initial sub-structure was carried out in PHENIX.AUTOSOL [7]. Due to poor initial experimental electron density maps, automatic building could not be performed to obtain a starting model. Iterative rounds of manual model building in Coot and refinement in REFMAC and PHENIX were necessary to obtain interpretable electron density maps [7, 84, 196].

The structure obtained by S-SAD was used as a search model to solve the structure of the Ptc1 ECD1-Nb64 and Ptc1 ECD1-Nb64:CHS complexes by MR in PHASER [184]. Iterative rounds of manual model building in Coot and refinement in REFMAC and PHENIX were used to trace the polypeptide chains of both proteins [7, 84, 196]. Stereochemical properties of the complexes were assessed with the structure-validation programme MolProbity [68]. The crystallographic and Ramachandran statistics of the data collection and processing are shown in Table 6.1.

2.10 Electron microscopy of biomolecules

2.10.1 Principles of EM

For many decades, X-ray crystallography remained the gold standard for solving structures of biomolecules at atomic resolution. Recent advances in cryo-electron microscopy (cryo-EM) caused this field to alter, and cryo-EM has become an indispensable tool for structural biology research. Due to rapid freezing techniques, the protein is captured in its hydrated state which closely reflects its native environment. The 1.8 Å structure of *Bos taurus* glutamate dehydrogenase published by A. Merk *et al.* in 2016 is the highest resolution cryo-EM structure available at the time of writing this thesis [187]. This and many other examples in the PDB reflect the great potential of cryo-EM in structural biology. In 2017, the Nobel prize for Chemistry was awarded to J. Dubochet, J. Frank and R. Henderson "for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution".

To produce high voltage electron beams, electron microscopes are equipped with a field emission gun (electron gun). The electron beam is condensed by electromagnetic lenses before being exposed to the biological specimen. The electrons interact with the atoms of the protein and can either be scattered or transmitted through the biological matter. After going through magnifying lenses, transmitted electron waves will reach the detector (fluorescent screen) and a magnified image (electron micrograph) is recorded by a digital charge-coupled device (CCD) camera. In cryo-EM, as in X-ray crystallography, radiation damage of the protein is of high concern. A low electron dose will protect the sample, but at the same time decrease the signal-to-noise ratio. A good compromise and common dose used in cryo-EM is 20 é/Å^2 per recorded image. A newer approach, which allows the use of a higher electron dose, is to record a movie, a series of electron micrographs (frames). The

principle behind this is simple: The electron dose is divided among the individual frames that are later aligned and combined into one image (micrograph). The first few frames contain the highest resolution information but also suffer the most from a beam-induced particle drift called beam-induced motion. Another factor that plays a vital role in obtaining good particle contrast is applying an underfocus when collecting a data set. Images taken in focus contain almost no structural information as the contrast of the biological sample itself is very low.

2.10.2 Specimen preparation and data collection

2.10.2.1 Negative stain EM

Negative staining is a commonly used technique to visualise biomolecules for transmission electron microscopy (TEM) studies. The sample's contrast is introduced by treating the specimen grid with heavy metal salts, such as uranyl acetate or uranyl formate stain. The stain embeds the specimen and forms an even layer around it. As a result, areas with high stain density will cause more electrons to scatter, whereas a thinner stain density on the specimen itself allows more electrons to be transmitted and detected. Negative staining is a quick and easy method to assess purity and homogeneity of a sample for structural biology studies.

Samples for negative stain EM studies were prepared as described in Chapter 2.5.3. For reconstitution in Amphipols, 0.5 mg of detergent-solubilised protein was mixed with 10% A8-35 (Anatrace) in a 1:8 (w/w) ratio. The sample was incubated for 3 hours at 4°C before 200 mg of pre-equilibrated Bio-Beads^{TM} SM-2 (BioRad) were added. The Amphipol exchange was allowed to complete at 4°C, overnight. The beads were separated from the sample and the eluate was subjected to SEC using a Superose 6 Increase 10/30 column (GE Healthcare) equilibrated in 10 mM HEPES

pH 7.5, 150 mM NaCl. For reconstitution in nanodiscs, detergent-solubilised Ptc1 was mixed with MSP2N2 in a 1:8 molar ratio. Next, DMPC was added in a 1:180 MSP2N2:DMPC molar ratio. The reaction was carried out in detergent buffer for 90 min at room temperature, followed by incubation with Bio-Beads^{TM} SM-2 resin (BioRad) overnight. The detergent-free eluate was collected and 1 mL of Rho1D4 antibody-coupled sepharose beads were added. The sample was incubated for 4 hours at 4°C, 20 rpm. The beads were collected and washed 5 x with 10 mM HEPES pH 7.5, 150 mM NaCl. The protein was eluted overnight at 4°C in wash buffer containing 0.5 mM Rho1D4 peptide (TETSQVAPA). The eluate was collected and concentrated using a 100 kDa MWCO filter (Sartorius) before being subjected to SEC using a Superose 6 Increase 10/30 column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl.

Copper grids with carbon support (CF-300-Cu, Electron Microscopy Sciences) were glow-discharged for 30 sec at a high voltage setting using a single chamber glow-discharge unit coupled to a vacuum pump. 4.0 μL of freshly purified sample was directly applied onto the glow-discharged grid for 1 min at room temperature. Excess sample was removed by contact with filter paper (blotting), followed by two wash-and-blotting steps with deionised water. The grid was then washed and blotted with aqueous uranyl formate (heavy metal stain), before being incubated with stain for 30 sec. Excess stain was removed by blotting and the grid was allowed to air-dry for 1 hour. All steps were carried out at room temperature. Negative stain analysis was performed on a Tecnai T12 microscope (FEI) operating at a voltage of 120 kV using an electron dose of $\sim$10 é/Å^2. Focusing was achieved manually and images were taken at nominal underfocus. Data were collected at a nominal magnification of 52.000 x and a pixel size of 0.214 nm. Images were recorded using a FEI Eagle 4k CCD camera.

2.10.2.2 Cryo-EM

Unlike negative staining, vitrification is a non-invasive method for preparing samples for structural biology studies. Rapid freezing in liquid ethane prevents the sample from dehydration and embeds it in a thin layer of amorphous ice. In the case of vitrification, contrast is mainly introduced by adjusting the cryo-EM settings and/or by using special equipment such as energy filters and phase plates.

Samples for cryo-EM studies were prepared as for negative staining experiments, with final protein concentrations of 0.25- 0.6 mg mL^{-1}. Three different types of grids were tested: Ultra-thin lacey carbon on 400 mesh copper grids (Agar Scientific), C-Flat holey carbon on 300 mesh copper grids (Protochips), and Quantifoil® holey carbon R 1.2/1.3 on 400 mesh gold grids (EMS). Using a single chamber glow-discharge unit coupled to a vacuum pump, ultra-thin lacey carbon support films and C-Flats were glow-discharged for 30 sec and 15 sec at high voltage, respectively, Quantifoil® grids were glow-discharged for 30 sec at medium voltage. For vitrification, 4.0 μL of sample was applied onto the grid for 1 min at 80- 96% humidity. Excess sample was blotted away for 4 sec at -15 blotting force. Grids were immediately plunge-frozen in liquid ethane. Specimen grid vitrification was carried out using a Vitrobot™ from FEI. Cryo-EM data were acquired on a Tecnai F30 microscope (FEI) operating at a voltage of 300 kV using an electron dose of ∼30 é/Å^2. Focusing was achieved automatically and images were taken at nominal underfocus. Data were collected manually at a pixel size of 1.35 Å. Images were recorded using a Gatan K2 Summit Detector.

2.10.3 Data processing of cryo-EM data

Correcting for electron beam-induced motion (BIM) is crucial for obtaining high resolution single particle cryo-EM structures. MotionCorr, a programme developed by Li *et al.*, features an algorithm that aligns image frames by identifying identical features [166]. As images are not taken in focus, but at different defocus values, blurring of the images is a commonly observed problem. Therefore, correction of the contrast transfer function (CTF) is a necessity. CTFFIND uses Fourier transform algorithms and corrects the amplitude of an image by fitting a calculated CTF to the amplitude in reciprocal space [189, 239].

The process of single particle reconstruction starts with particle picking from a pre-processed data set. At first, a good number of particles are manually picked from a limited number of micrographs and used to generate a small set of 2D class averages. Each 2D class projection comprises particles with the same orientation, and a meaningful 2D class data set displays particles from many different perspectives. The obtained 2D classes are taken as an input model for subsequent automated particle picking. A commonly used programme is RELION, developed by S. Scheres at MRC LMB Cambridge (UK) in 2012, which provides a user friendly pipeline for single particle analysis [246]. The large number of particles is extracted and processed in RELION or cryoSPARC^{TM} to generate a comprehensive set of 2D class averages [226]. The 2D projections are then inspected and separated from empty or invalid classes, followed by *ab initio* 3D reconstruction. The algorithm for 3D reconstruction determines three Euler angles and two translations for each of the 2D class averages to reconstruct a 3D map of the protein. Projection matching algorithms are then applied to refine the structure (iterative refinement) of the initial 3D map.

A small data set of 300 micrographs was collected using a Tecnai F30 microscope (FEI) and the micrographs were aligned using MotionCorr. Particles were manually picked in RELION, followed by estimation of the CTF parameters using CTFFIND. A total of ∼2000 particles were extracted in RELION. Extracted particles were subjected to 2D classification in cryoSPARC^{TM}.

Chapter 3

Nanobodies target Shh and its receptor Ptc1

3.1 Abstract

Nanobodies are isolated, highly immunogenic regions of camelidae heavy chain antibodies (HCAb). HCAbs can be generated *in vivo* by immunisation of camelids with biochemically prepared protein of interest. Our goal was to identify nanobodies that specifically recognise the Hedgehog (Hh) pathway ligand Sonic Hedgehog (Shh) and its receptor Patched-1 (Ptc1). The Ptc1 construct used for llama immunisation (Ptc1 ECD1-ECD2) was deficient of the transmembrane regions and only contained the two large extracellular domains (ECDs) of the receptor. Those domains (ECD1 and ECD2) were fused by an artificial Gly/Ser (G/S)-rich linker. Immunisation of the llama with the Shh-Ptc1 ECD1-ECD2 complex, as well as the selection and cloning of nanobodies was carried out in collaboration with the Steyaert laboratory at Vrije Universiteit Brussels (VUB). Nanobodies were expressed, purified and subsequently characterised in protein-nanobody interaction experiments using bio-layer interferometry (BLI). All nanobodies showed nanomolar to low micromolar binding to its target proteins and in most cases very low off-rates. This work laid the ground for structural and cellular studies of the two Hh signalling components Shh and Ptc1.

3.2 Introduction

3.2.1 Nanobodies, a smaller type of antibodies

In mammals, antibodies (also termed immunoglobins, Ig) are inherently produced by the adaptive immune system in response to pathogen (bacteria, viruses, etc) exposure. Antibodies are specific for antigens presented on the pathogen's outer layer and either neutralise the antigen directly or opsonise the pathogen for further attack of the immune system. After infection, antibodies are retained in the system to ensure a rapid and enhanced immune response to future infections.

Antibodies are Y-shaped, heterotetrameric glycoproteins formed by a C-terminal crystallisable fragment (Fc) and an N-terminal antigen-binding (Fab) domain (Figure 3.1, A). Fc fragments are constant regions (CH regions) formed by heavy chains only and determine the class of an antibody. Classes such as IgG, IgA, and IgD consist of three CH domains (CH1-CH3), whereas IgM and IgE carry four CH domains (CH1-CH4). As part of the immune response, Fc regions bind to Fc receptors of the complement system and multiply the defense signal. The Fab domains consist of heavy and light chains, both harbouring a variable and a constant region. The heavy chain variable region (V_H) and the light chain variable region (V_L) are located at the tip of the Fab domain and represent the paratope, the antigen-binding site of the antibody. The specificity of the paratope is determined by six complementarity-determining regions (CDRs). Fab fragments can be isolated, and their variable regions V_H and V_L can be fused together (scFv paratope) to form a smaller yet powerful molecule for biological and medicinal application.

Recently, a new class of antibodies was discovered in camelids (HCAb) and cartilaginous fishes (IgNARC) [105, 114]. Interestingly, these types of antibodies lack

light chains characteristic for conventional antibodies. Instead, these antibodies contain heavy chains only that carry a variable, antigen-binding region at their tip and constant regions (Fc) at their tail (Figure 3.1, A). Similar to conventional antibodies, variable regions can be isolated as single-chain antibody fragments termed V_{HH} or V_{NAR} for camelid and cartilaginous fish antibodies, respectively. Nowadays, V_{HH} domains are better known as nanobodies, which have been used extensively for biological and medical research due to their interesting biomedical properties. Nanobodies have the same specificity as conventional antibodies but are more stable and can be expressed in various expression system yielding milligram quantities [198].

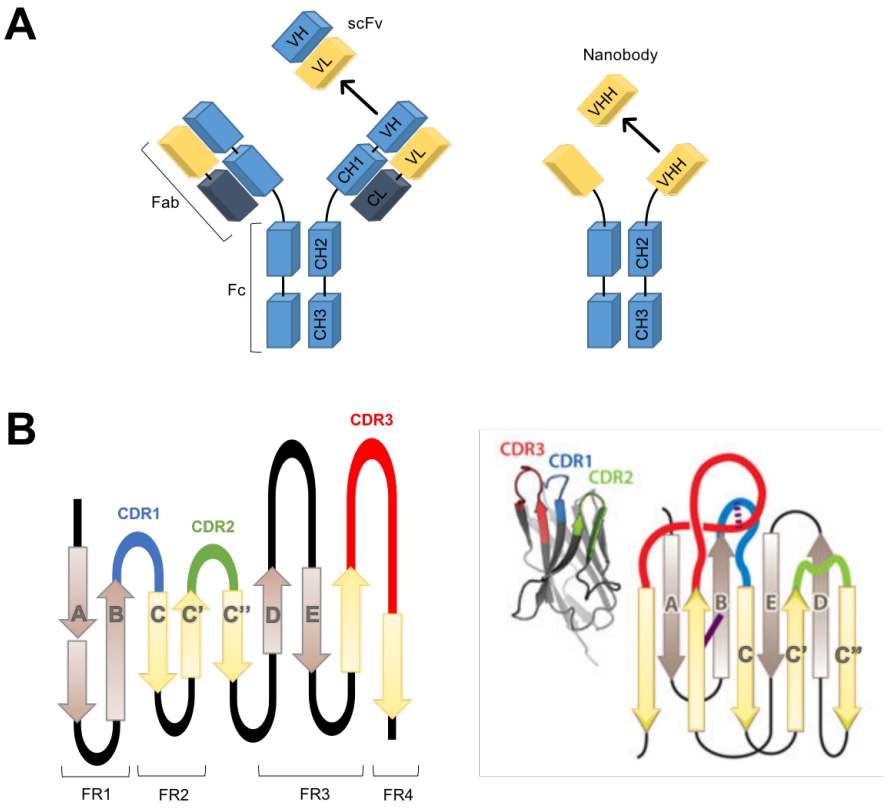


Figure 3.1: **Antibodies vs. Nanobodies. A**, Schematic representation of conventional antibodies (left) and single-chain antibodies (right) as found in camelids. Nanobodies are derived from the tip of single-chain antibodies (highlighted in yellow). **B**, A nanobody fold generally consists of nine consecutive β-strands. The antigenic region is formed by three CDR regions CDR1-3. The fold of most nanobodies is stabilised by two disulfide bonds (purple and purple dotted line). Figures adapted from [115, 178, 204].

A closer look into the global fold of nanobodies reveals two antiparallel β-sheet motifs that are arranged as a pair to give a unique, compact shape (Figure 3.1, B) [178]. In addition, the total of nine β-strands can be grouped into four frameworks (3x2 and 1x3 β-strands) along its sequence [204]. Frameworks (FRs) are classified as constant regions and are interspersed by hypervariable loops, the complementarity determining regions 1-3 (CDRs). The CDRs form the paratope of the nanobody, which is responsible and sufficient for antigen binding.

3.2.2 Current applications of nanobodies

Due to their unique characteristics, nanobodies have been extensively used for fundamental research purposes, as early diagnostic tools and for a variety of medical applications. The following paragraphs will touch on each topic in the above stated order and summarise its main points.

A chromobody is a fusion of a nanobody with a fluorescent protein, e.g. green fluorescent protein (GFP). It is an elegant tool that has been developed for direct visualisation of target proteins in living cells. The main advantage of using chromobodies is that no modification of the target protein is needed for its tracing and/or molecular imaging. Several chromobodies are already commercially available or can simply be designed and produced in a conventional laboratory setting [70]. For immunoprecipitation (IP) experiments, a novel system has recently been introduced by A. Pollithy and co-workers. In this approach, nanobodies are displayed on the surface of magnetosomes (bacterial magnetic nanoparticles) by fusing the nanobodies to magnetosome-specific membrane proteins [223]. Due to their excellent magnetic properties, these particles can be used for magnetic-based IP experiments. A quite simple method that has been developed in the laboratory of H. Leonhardt at LMU Munich,

Germany is to use anti-GFP nanobodies as GFP-nanotraps® by immobilising the nanobody on a monovalent matrix [241]. This approach is suitable for chromatin IPs using GFP-tagged DNA-binding proteins, and one-step purifications of GFP-labelled proteins and their binding partners for mass spectrometric analysis and biochemical enzyme activity assays [241]. Another novel application that has recently been presented is to use nanobodies for agricultural purposes. The company Agrosavfe has developed a system for specific nanobody-based delivery of agrochemicals to crops and weeds. The agrocapsules increase the retention time while decreasing the quantity of chemicals needed for the desired purpose. A similar system has now been developed for introducing insecticides to insects [70]. Nanobodies have also been extensively used for structural biology studies as they can stabilise a protein's native fold and hence facilitate crystallisation of difficult to crystallise target proteins. Recently, the application of nanobodies has also gained interest in the field of cryo-electron microscopy. Nanobodies not only stabilise highly flexible regions or distinct conformations of a protein, but also add additional molecular weight, which aids visualisation of smaller proteins under the microscope [78, 211].

As for diagnostics, nanobodies have been labelled with radionuclides for preclinical non-invasive imaging of tumours [277]. Compared to conventional antibodies, nanobodies show much better tumour penetration and excellent renal clearance properties, which is important when working with potentially harmful substances such as radionuclides. For pathogen diagnosis such as HIV infection, the HIV-1 p24 antigen has been fused to a red blood cell-specific nanobody [110]. This bispecific protein construct readily causes agglutination of blood from HIV-positive patients and is therefore an excellent biomarker to detect the disease. Even though nanobodies are currently extensively studied for therapeutic applications, clinical studies are still in the early stages of development. For instance, nanobodies have been raised against

extracellular targets such as membrane proteins and/or their ligands. The goal of this approach is to delay disease progression or to inhibit tumour growth by targeting members of signalling pathways with highly specific nanobodies. Another elegant approach, especially for cancer therapy, is nanobody-mediated drug delivery. The drugs are encapsulated into nanoparticles and delivered to the target cells by binding of the nanobodies to cell-surface receptors. Those systems still undergo development but have already been shown to inhibit cell-growth and proliferation in pre-clinical trials [275].

Nanobodies are versatile tools that can be used for a wide range of applications. Their production is not only economically viable, but also relatively straightforward. Nanobodies are exceptionally stable, have excellent solubility properties and exhibit high specificity and affinity to their target molecules. It will be interesting to follow future development of nanobody-based systems.

3.3 Results

3.3.1 Llama immunisation

For llama immunisation, 0.5 mg of protein was required for five sequential protein injections. Human Sonic Hedgehog (Shh) protein was expressed in *E.coli* and purified via immobilised metal affinity chromatography (IMAC), followed by size-exclusion chromatography (SEC). The secreted human Patched-1 (Ptc1) extracellular domain (ECD) fusion construct (ECD1-ECD2) was expressed in HEK293T cells and purified via Rho1D4 antibody affinity chromatography, followed by SEC. Protein preparation is discussed in detail in Chapter 2.4.1 and 2.5.1. Both proteins were mixed in a 3:1 (Shh:Ptc1 ECD1-ECD2) molar ratio in the presence of 1.4 mM Ca^{2+} for 15 min. The protein was not subjected to any other pretreatment before immunisation. Three to five days after the last immunisation, RNA was isolated from camelid peripheral blood lymphocytes and a cDNA library was generated. Nanobody regions were amplified by PCR and sequences were cloned into a suitable phage display vector. As llama immunisation was carried out with non-covalently linked complex, there was a chance that the two proteins would fall apart during incubation. Thus, nanobodies displayed on the phage envelope were screened against Shh and Ptc1 ECD1-ECD2 alone, as well as the complex. Two individual phage display experiments were performed. For the first phage display experiment, Shh was directly immobilised (solid phase-coated via aspecific adsorption) onto ELISA plates, whereas Ptc1 ECD1-ECD2 and the Shh-Ptc1 ECD1-ECD2 complex were either solid phase-coated or conjugated to Rho1D4 antibody-coated plates. The second phage display was performed with biotinylated proteins that were applied onto neutravidin-coated ELISA plates. To avoid dissociation of the complex during the second screening, the Shh-Ptc1 ECD1-ECD2 complex was covalently crosslinked using chemical crosslinkers with the objective of

finding additional complex-specific nanobodies. Screening with biotinylated Shh was performed because solid phase immobilisation can lead to partial denaturation of the protein and unfolded protein might not be sensitive for detection of interesting nanobody candidates. Before carrying out phage display experiments, all proteins were subjected to SEC. Results of the protein preparation are discussed in Chapter 3.3.3 and Chapter 4.3.1. Both ELISA screenings yielded a total of 27 unique nanobodies; 13 from the first screening and 14 from the second screening (Table 3.1).

Official Reference ID (1)	Internal Reference (2)	Family (3)	Occurance (4)	Shh (5)	Ptc1 ECD1-ECD2 (6)
CA10664	Nb 64				x
CA10665	Nb 65				x
CA10666	Nb 66				x
CA10667	Nb 67	1	263		x
CA10790	Nb 790				x
CA10791	Nb 791				x
CA10792	Nb 792			*	*
CA10793	Nb 793				x
CA10668	Nb 68	2	3		x
CA10669	Nb 69	3a	30	x	
CA10794	Nb 794			x	
CA10670	Nb 70	3b	4	x	
CA10671	Nb 71			x	
CA10795	Nb 795			*	*
CA10796	Nb 796	3c	16	x	
CA10797	Nb 797			x	
CA10798	Nb 798			x	
CA10799	Nb 799			x	
CA10677	Nb 77	4	1	x	
CA10678	Nb 78	5	1	x	
CA10672	Nb 72	6	1	*	*
CA10676	Nb 76	7	1	*	*
CA10675	Nb 75	8	1		x
CA10800	Nb 800	9	3	*	*
CA10802	Nb 802	10	2		x
CA10803	Nb 803	11	1		x
CA10804	Nb 804	12	1		x

Table 3.1: **Overview of nanobodies against Shh and Ptc1 ECD1-ECD2.** Nanobodies from the first screening and second screening are highlighted in dark blue and light blue, respectively. **1**, CA numbers represent official nanobody reference IDs as created by VUB, Belgium. **2**, Internal nanobody nomenclature derived from the last two or three digits of the official reference ID. **3**, Nanobody classification is based on CDR3 sequence similarity. **4**, The occurrence indicates how many times a family was found. **5-6**, x indicates nanobodies that bind to either Shh (5) or Ptc1 ECD1-ECD2 (6). Nanobodies that could not be verified for binding are indicated by an asterisk.

3.3.2 Production of nanobodies

Nanobody (Nb) sequences in pMESy4 background with N-terminal pelB leader sequence and C-terminal His_6-tag were received from our collaborators at VUB, Belgium [211]. All 27 constructs were expressed in *E.coli* and purified from the periplasmic extract via IMAC, followed by SEC. Proteins eluted as single monodisperse peaks, with the exception of Nb72, which displayed two overlapping main peaks. A representative SEC profile (Nb64) can be found in Figure 3.2, A. High purity of all samples was subsequently confirmed by SDS-PAGE analysis as shown in Figure 3.2, B. One litre of culture yielded milligram quantities of nanobody (1-20 mg, depending on the construct). All nanobodies were concentrated to 5 mg mL^{-1}, however, Nb77 precipitated above concentrations of 0.5 mg mL^{-1}. No significant change in protein quality was observed after flash-freezing in liquid nitrogen and storage of the proteins at -80°C. Nanobodies were used for affinity measurements, structural analysis and cellular studies.

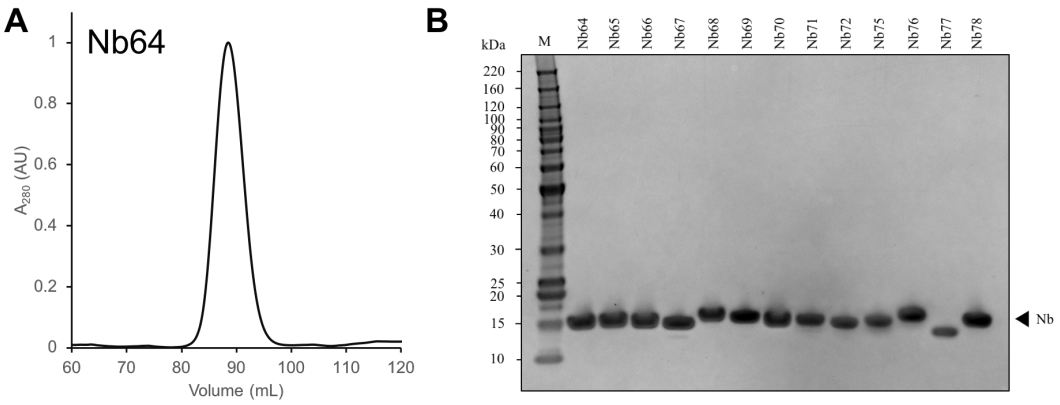


Figure 3.2: **Purification of nanobodies against Shh and Ptc1. A**, SEC profile of nanobody 64 (Superdex 75 16/60 column). The protein purifies as a single monodisperse peak. This example is representative for most of the other nanobodies purified. **B**, SDS-PAGE of purified nanobody samples confirms sample purity. Lanes from left to right: M, BenchMark^{TM} Marker; Nb64, 65, 66, 67, 68, 69, 70, 71, 72, 75, 76, 77, 78.

3.3.3 Binding studies of nanobodies using Bio-layer interferometry

In ELISA screenings, a total of 27 nanobodies were identified, which belong to 12 different sequence families. An overview of all nanobodies can be found in Table 3.1. Sequence family members share the same B-lymphocyte heritage and recognise the same site of the antigen as their CDR3 domains show 80%- 100% sequence similarity [211]. Sequence alignments of nanobodies can be found in Figure 3.5, B (Shh-specific nanobodies) and Figure 3.8, B (Ptc1-specific nanobodies).

In order to characterise the interaction between nanobodies and target proteins, bio-layer interferometry (BLI) was carried out using an Octet® Red384 instrument. Biotinylated proteins (referred to as "ligands") were immobilized on streptavidin-coated biosensors and tested against all nanobodies ("analytes") for binding. Appropriate immobilisation response was achieved when the ligand was incubated with optical biosensors for 300 sec. Successful binding to the sensors was indicated by a change in response values which is due to increasing optical thickness upon binding of the ligand. In general, a larger signal is expected from smaller molecules. Typical response values were 10 nm for Shh, 4 nm for Ptc1 ECD1-ECD2 and 5 nm for the complex. Immobilisation quickly reached a plateau, and binding was irreversible as indicated by very little dissociation of proteins. To reduce unspecific binding of proteins to the sensors, 0.005% Tween-20 was added to the running buffer. Proteins were then tested against all 27 nanobodies in a BLI rapid screen assay to identify specific ligand-analyte interactions. Ligand-coupled sensors were incubated with 1.0 μM of each analyte for 300 sec, after a 60 sec equilibration period. Using this method, 10 nanobodies were identified for Shh and 12 for Ptc1 ECD1-ECD2. Subsequently, positive ligand-analyte pairs were tested for their affinities in multi-cycle kinetic mea-

surements. The ligands were immobilised as described for the BLI rapid screen assay, and tested against various concentrations of analyte by preparing a two-fold dilution series of nanobody. Dissociation constants (K_D) values ranged from low nanomolar to low micromolar affinity. Results are discussed in detail below.

3.3.3.1 Shh-specific nanobodies

In order to immobilise proteins onto optical biosensors for BLI, proteins were labelled with a tag sequence. For this experiment, Shh was cloned into the pHLsec vector with N-terminal protein secretion signal and C-terminal AviTag^{TM} for biotinylation, followed by a His_6-tag for purification [12]. The protein was expressed in HEK293T cells and biotinylated *in vivo* by co-expressing Shh with BirA biotin ligase. Protein purification was carried out by IMAC and SEC, similar to purifications of non-biotinylated samples as described in Chapter 2.5. Shh eluted as a broad, monodisperse peak with a peak maximum at ∼70 mL retention volume on a Superdex 75 16/60 column (GE Healthcare) as shown in Figure 3.3, A. Using SEC, protein aggregate and higher molecular weight complexes were successfully separated from monomeric Shh protein (result not shown). The peak fractions were pooled and concentrated to 0.1 mg mL^{-1} for subsequent BLI experiments. Corresponding SDS-PAGE analysis confirmed the purity of the sample (Figure 3.3, A).

Figure 3.3, B illustrates a representative example for real-time immobilisation of biotinylated Shh to streptavidin-coated biosensors. Immobilisation was fast and consistent among the eight sensors, with strong response values of 10 nm. Low dissociation rates were indicators for strong and irreversible biotin-streptavidin interaction and very little unspecific binding. The Shh-immobilised biosensors were then used in a BLI rapid screen assay to detect Shh-specific nanobodies. Five nanobodies were identified in the first screening: Nb69, 70, 71, 77 and 78 (Figure 3.3, C). Nb70

and Nb77 showed the strongest and very similar response values of 1.6 and 1.7 nm, with relatively low off-rates (dissociation). Nb78 showed the lowest (0.5 nm), and Nb69 and Nb71 response values in between (1.2 nm), all with relatively good dissociation rates. Nb64, 65, 66, 67, 68, 72, 75, and 76 showed no response, indicating that these nanobodies are not specific for Shh. In a second BLI rapid screen setup, Nb794, 796, 797, 798, and 799 were identified as Shh-specific nanobodies (results not shown).

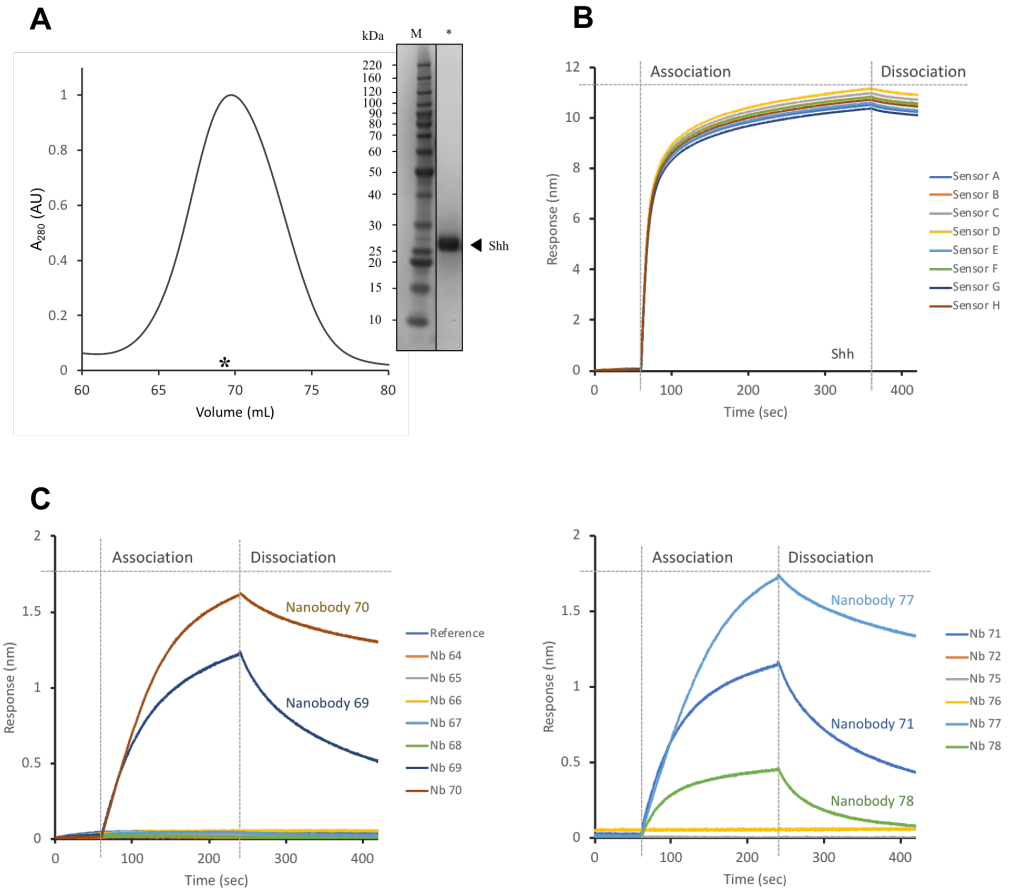


Figure 3.3: **Identification of Shh-specific nanobodies.** **A**, SEC profile of biotinylated Shh (Superdex 75 16/60 column). The protein purifies as a single monodisperse peak. SDS-PAGE of purified Shh sample confirms sample purity. BenchMark^{TM} Marker (M), SEC fraction analysed is indicated by an asterisk. **B**, Immobilisation of Shh to biosensors for BLI measurements. **C**, BLI rapid screen to identify Shh-specific nanobodies.

All Shh-Nb pairs were subsequently subjected to kinetic characterisation for dissociation constant (K_D) determination. Seven concentrations of nanobody, resulting from two-fold dilutions were tested in parallel and K_D values were calculated from the observed on- and off-rates. Figure 3.4, A-B shows the real-time sensorgrams and steady-state analysis for Nb70 and Nb77 as representative examples. Both nanobodies showed high affinity to Shh with K_D values of $4.4 \times 10^{-8} \pm 5.3 \times 10^{-9}$ M and $6.5 \times 10^{-8} \pm 8.2 \times 10^{-9}$ M, respectively. An overview of K_D values for all Shh nanobodies can be found in Figure 3.5, A. Data analysis was carried out using ForteBio's data analysis software (Version 9.0). Even though our structural studies (Chapter 4) support a 1:1 interaction between the individual nanobodies and Shh, a 1:1 binding model, which uses pseudo first order kinetics for fitting, could not be used. For most nanobodies equilibrium and baseline signal was not reached during association and dissociation steps, respectively. Instead the curves had biphasic character and thus data were fitted using a 2:1 heterogeneous binding model. This particular model is often used when the analyte recognises two binding sites on the ligand. For Shh-Nb interactions it cannot be ruled out that biosensors were over-saturated with Shh, which led to the occurrence of two binding events. The first binding might have occurred rapidly, whereas binding with lower on-rates occurred when remaining Shh sites were slowly occupied by nanobodies.

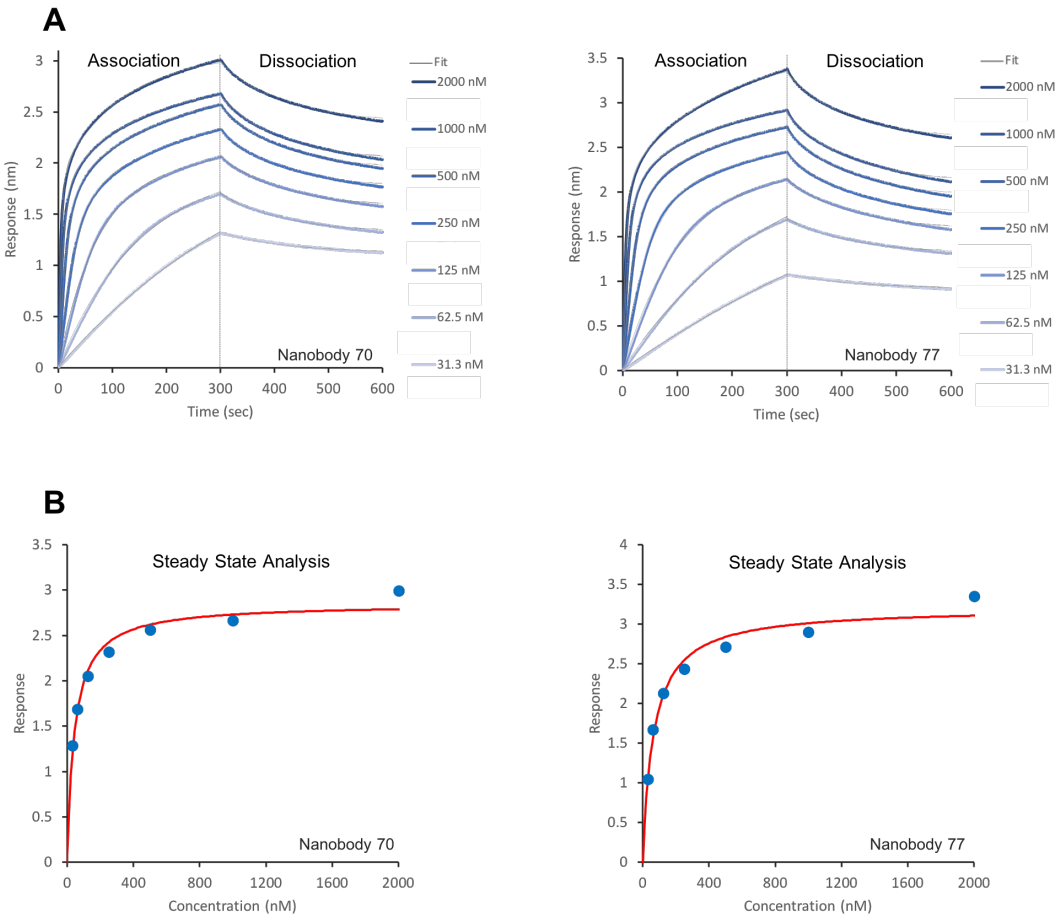


Figure 3.4: **Binding studies of Shh-specific nanobodies. A-B**, Characterisation of Shh-Nb70 (A) and Shh-Nb77 (B) interaction. **A**, Real-time sensorgrams for association and dissociation of nanobody 70 (left) and nanobody 77 (right). Statistical fitting of experimental data is shown in grey. Association and dissociation steps are divided by a black dotted line. **B**, Steady-state analysis to determine the equilibrium dissociation constant. A K_D of 4.4 x 10^{-8} ± 5.3 x 10^{-9} M and 6.5 x 10^{-8} ± 8.2 x 10^{-9} M was determined for Nb70 and Nb77, respectively.

Sequence analysis (Figure 3.5, B) of all Shh nanobodies revealed clear differences in the CDR3 regions of Nb77 and Nb78 compared to other Shh-specific nanobodies. As anticipated, these nanobodies belong to different sequence families, family 4 and 5, respectively. All other nanobodies belong to sequence family 3, which is further subdivided into categories a to c as they mainly differ in their CDR1 and CDR2 region. All Shh-specific nanobodies were used for co-crystallisation experiments with Shh. These results are discussed in detail in Chapter 4.

A

Shh nanobodies	Family	K_D (M)	R^2
Nb 69	3a	$1.2 \times 10^{-7} \pm 7.0 \times 10^{-9}$	0.995
Nb 794		$4.9 \times 10^{-8} \pm 2.5 \times 10^{-9}$	0.996
Nb 70	3b	$4.4 \times 10^{-8} \pm 5.3 \times 10^{-9}$	0.963
Nb 71		$1.6 \times 10^{-7} \pm 9.8 \times 10^{-9}$	0.996
Nb 796	3c	$7.3 \times 10^{-8} \pm 2.9 \times 10^{-9}$	0.998
Nb 797		$1.5 \times 10^{-7} \pm 2.1 \times 10^{-8}$	0.981
Nb 798		$5.2 \times 10^{-8} \pm 4.5 \times 10^{-9}$	0.988
Nb 799		$1.0 \times 10^{-7} \pm 2.5 \times 10^{-8}$	0.911
Nb 77	4	$6.5 \times 10^{-8} \pm 8.2 \times 10^{-9}$	0.969
Nb 78	5	$6.3 \times 10^{-7} \pm 4.6 \times 10^{-8}$	0.993

B

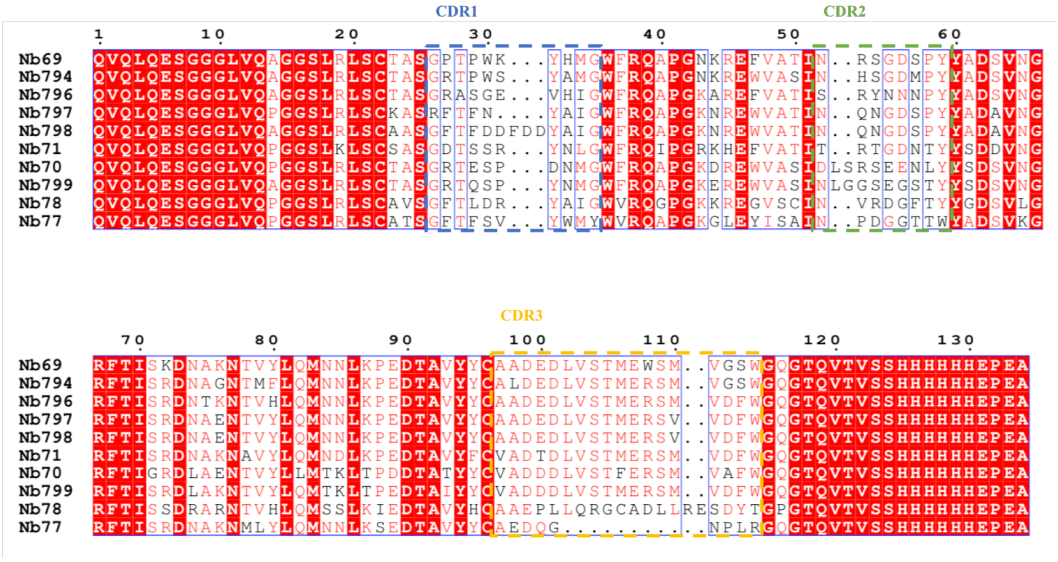


Figure 3.5: **Sequence alignment of Shh-specific nanobodies. A**, Table including dissociation constants of Shh-Nb complexes and corresponding coefficient of determination R^2. **B**, Sequence alignment of Shh nanobodies. Sequence similarity is indicated by colour coding. Hypervariable regions CDR1-3 are coloured in blue, green and yellow respectively. Alignment generated with MultAlin and ESPript 3 [63, 234].

3.3.3.2 Ptc1-specific nanobodies

For biophysical studies, Ptc1 ECD1-ECD2 was cloned into the pHLsec vector with an N-terminal protein secretion signal and a C-terminal BirA recognition sequence (AviTag^{TM}), followed by a Rho1D4 antibody epitope tag [12]. Ptc1 ECD1-ECD2 was co-expressed with BirA biotin ligase to achieve *in vivo* biotinylation in HEK293T cells. The protein was then directly purified from the tissue culture medium. Prior to SEC, the protein was purified via Rho1D4 antibody affinity chromatography. The SEC profile (Figure 3.6, A) revealed two overlapping peaks; one broader peak at ∼55 mL and one sharper peak at ∼80 mL retention volume, respectively. The first peak indicated formation of Ptc1 ECD1-ECD2 multimers or aggregates, which might happen during expression and/or purification of the protein. To ensure high quality of monomeric Ptc1 ECD1-ECD2 sample for BLI studies, only fractions to the right of the peak maximum were pooled and concentrated to 0.1 mg mL^{-1}. SDS-PAGE analysis revealed a single band of the expected molecular weight (Figure 3.6, A), which at the same time confirmed that the sample was of high purity.

Similar to Shh, association of biotinylated Ptc1 ECD1-ECD2 to streptavidin-coated biosensors occurred rapidly with a typical response rate of 4 nm and negligible dissociation after immobilisation (Figure 3.6 B). BLI rapid affinity screens identified six nanobodies (Nb64, 65, 66, 67, 68, and 75) for the first set, and another six nanobodies (Nb790, 791, 793, 802, 803, 804) for the second set tested. The results of the second screening are shown in Figure 3.6, C. Most nanobodies showed response values between 0.6-0.8 nm, however, Nb68, 802, 303, and 804 showed lower values between 0.15-0.25 nm. All Ptc1-Nb pairs were then further characterised in affinity measurements for K_D determination. A detailed overview of Ptc1-specific nanobodies and their dissociation constants can be found in Figure 3.8, A. Nb64, 65, 66, 67, 791

and 793 showed the highest affinity to Ptc1 ECD1-ECD2 with K_D values below 55 nM. All other nanobodies showed affinities in the range of 135 nM- 2.15 μM. Two representative real-time sensorgrams (Nb64, Nb802) are shown in Figure 3.7, A. Fitting of the data was carried out using a 1:1 binding model and applying either a local fitting if dissociation reached baseline values, or a partial fitting in case pre-association values were not met at the end of the measurement.

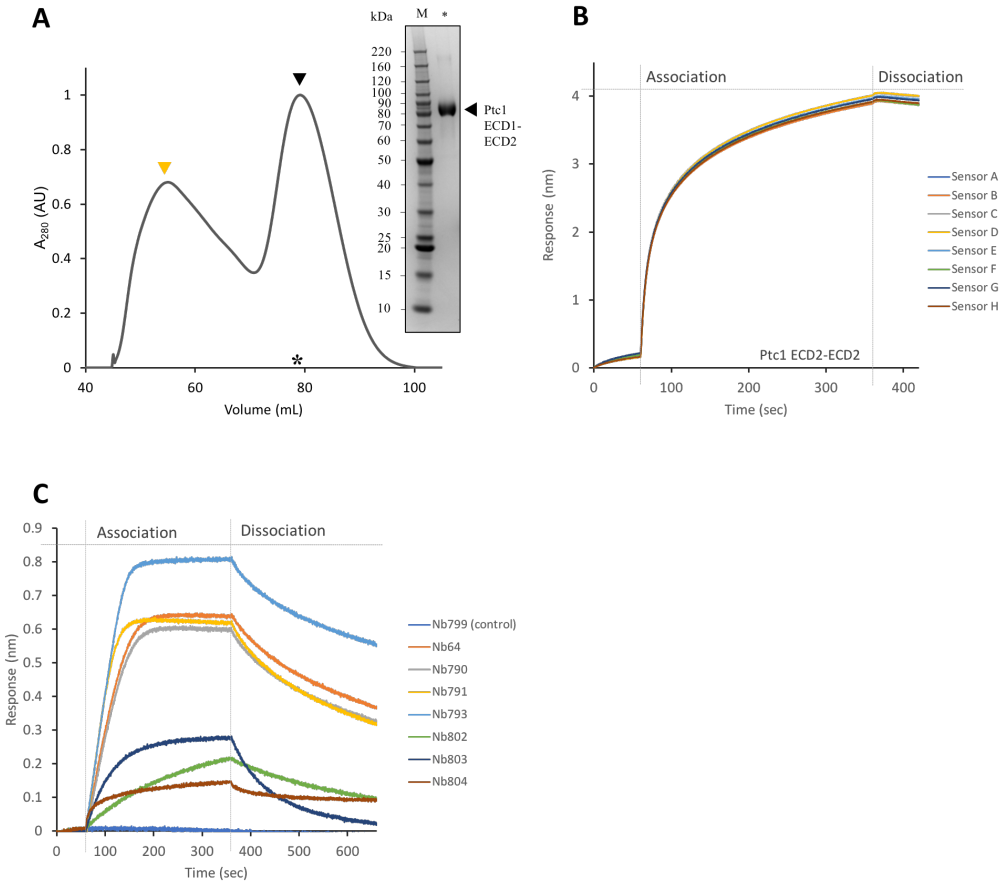


Figure 3.6: **Identification of Ptc1-specific nanobodies. A**, SEC profile and corresponding SDS-PAGE of Ptc1 ECD1-ECD2 (Superdex 200 16/60 column). Yellow arrowhead indicates the peak of aggregated or multimeric Ptc1 ECD1-ECD2, black arrowhead the peak of monomeric Ptc1 ECD1-ECD2. SDS-PAGE of purified Ptc1 ECD1-ECD2 monomeric fraction confirms sample homogeneity. Lanes from left to right: BenchMark^{TM} Marker (M), Peak SEC fraction (marked by an asterisk in the corresponding SEC trace). **B**, Immobilisation of Ptc1 ECD1-ECD2 to optical biosensors for BLI analysis. **C**, Sensorgram of BLI rapid screen to identify Ptc1-specific nanobodies. Nb799 was used as a negative control (Shh-specific nanobody).

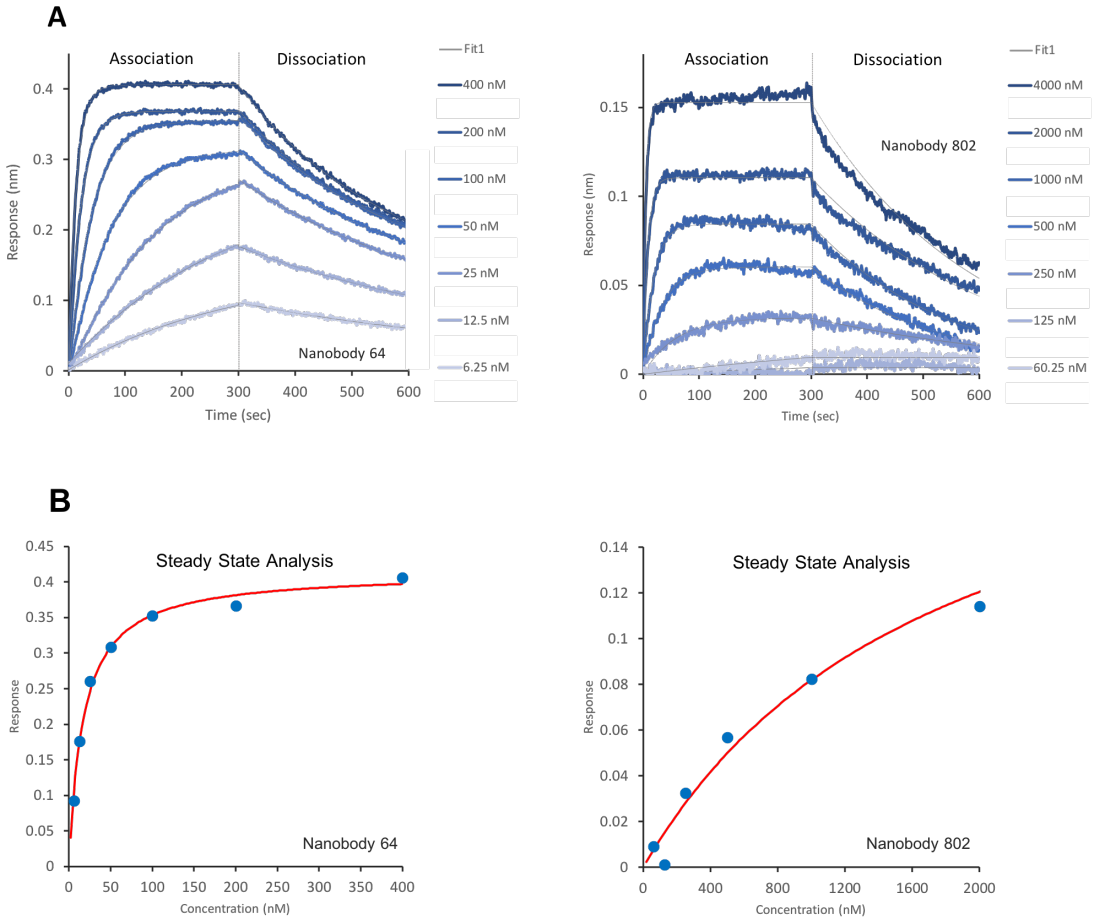


Figure 3.7: **Binding studies of Ptc1-specific nanobodies. A-B**, Characterisation of Shh-Nb64 (A) and Shh-Nb802 (B) interaction. **A**, Real-time sensorgrams for association and dissociation of Nb64 (left) and Nb802 (right). Statistical fitting of experimental data is shown in grey. Association and dissociation steps are divided by a black dotted line. **B**, Steady-state analysis to determine the equilibrium dissociation constant. A K_D of 1.7 x 10^{-8} ± 1.5 x 10^{-9} M and 1.8 x 10^{-6} ± 3.4 x 10^{-8} M was determined for Nb64 and Nb802, respectively.

Sequence analysis (Figure 3.8, B) revealed considerable differences in the CDR3 regions of Nb68, 75, 802, 803, and 804, not only among each other but also compared to other Ptc1-specific nanobodies. Interestingly, Nb64, 65, 66, 67, 790, 791, 793 showed no differences in their CDR3 and only minor differences in their CDR2 regions (with the exception of Nb67). This suggests that these nanobodies interact in a similar way with Ptc1 ECD1-ECD2 and/or their binding interface is modulated by their CDR1 region only. Nb68, 75, 802, 803, and 804 possibly recognise

different sites or interact in slightly different ways with Ptc1 ECD1-ECD2. However, further experiments, such as crystallisation of the complexes is required to test this hypothesis.

A

Ptc1 ECD1-ECD2 nanobodies	Family	K_D (M)	R^2
Nb 64		$1.7 \times 10^{-8} \pm 1.5 \times 10^{-9}$	0.989
Nb 65		$1.4 \times 10^{-8} \pm 2.8 \times 10^{-9}$	0.959
Nb 66		$1.7 \times 10^{-8} \pm 2.3 \times 10^{-9}$	0.987
Nb 67	1	$1.8 \times 10^{-8} \pm 5.0 \times 10^{-9}$	0.962
Nb 790		$3.4 \times 10^{-7} \pm 7.2 \times 10^{-8}$	0.945
Nb 791		$3.2 \times 10^{-8} \pm 9.3 \times 10^{-9}$	0.934
Nb 793		$3.7 \times 10^{-8} \pm 1.5 \times 10^{-8}$	0.879
Nb 68	2	$1.7 \times 10^{-7} \pm 1.9 \times 10^{-8}$	0.985
Nb 75	8	$1.5 \times 10^{-7} \pm 1.5 \times 10^{-8}$	0.992
Nb 802	10	$1.8 \times 10^{-6} \pm 3.4 \times 10^{-7}$	0.985
Nb 803	11	$1.0 \times 10^{-7} \pm 3.4 \times 10^{-8}$	0.800
Nb 804	12	$1.7 \times 10^{-6} \pm 3.9 \times 10^{-7}$	0.937

B

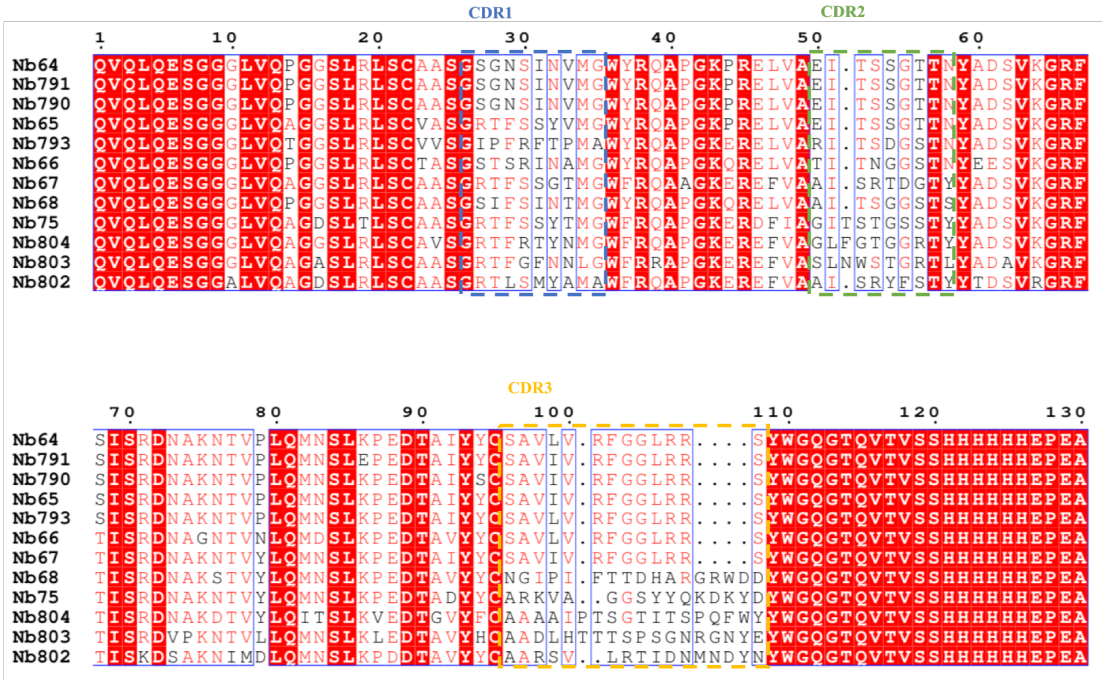
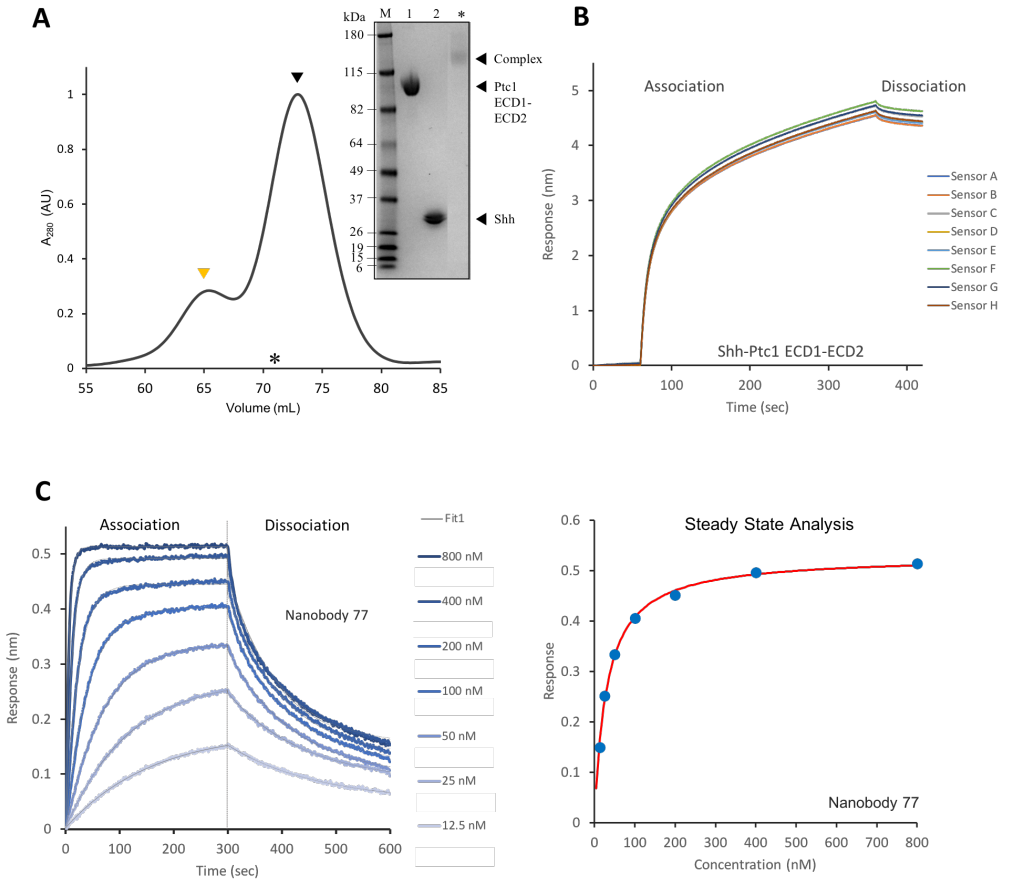


Figure 3.8: **Sequence alignment of Ptc1-specific nanobodies. A**, Table including dissociation constants of Ptc1-Nb complexes and corresponding coefficient of determination R^2. **B**, Sequence alignment of Ptc1-specific nanobodies. Sequence similarity is indicated by colour coding. Hypervariable regions CDR1-3 are coloured in blue, green and yellow respectively. Alignment generated with MultAlin and ESPript 3 [63, 234].

3.3.3.3 Nanobodies that target the Shh-Ptc1 complex

A common approach to stabilise protein-protein complexes is to introduce permanent covalent interactions between proteins. Nowadays, the most frequently used method is to crosslink lysine residues that are in close proximity to each other and freely accessible by chemical crosslinkers. The company CovalX has developed a kit with highly efficient amine residue crosslinkers, that have shown excellent results in protein-protein crosslinking experiments for mass spectrometric analysis [24]. Therefore, we were confident using these crosslinkers to produce a stable Shh-Ptc1 ECD1-ECD2 complex for BLI experiments. For the experiments, Shh was expressed and purified as explained in Chapter 3.3.3.1 without any modifications in the protein preparation, and Ptc1 ECD1-ECD2 as in Chapter 3.3.3.2 but without BirA biotin ligase being present. Both proteins were mixed in a 2:1 molar ratio, and incubated for 30 min on ice in the presence of 3 mM $CaCl_2$. Subsequently, 0.2 mM of crosslinker was added to the sample and the reaction was allowed to proceed for two hours at room temperature. The reaction was then quenched by adding Tris buffer (pH 8). The sample was subjected to SEC and analysed on SDS-PAGE before affinity measurements were carried out. SEC profiles of the chemically crosslinked Shh-Ptc1 ECD1-ECD2 complex revealed two overlapping peaks with peak maxima at ∼65 mL and 73 mL retention volume (Figure 3.9, A). On SDS-PAGE, comparison of free Shh and free Ptc1 ECD1-ECD2 with the crosslinked complex confirmed that the proteins were successfully crosslinked, mostly resulting in a 1:1 complex with little contamination of free Ptc1 ECD1-ECD2 (Figure 3.9, A).



Complex nanobodies	Family	K_D (M)	R^2
Nb 65		$3.5 \times 10^{-7} \pm 1.1 \times 10^{-7}$	0.914
Nb 66	1	$1.1 \times 10^{-7} \pm 4.3 \times 10^{-8}$	0.754
Nb 67		$1.0 \times 10^{-7} \pm 3.2 \times 10^{-8}$	0.925
Nb 69	3a	$2.3 \times 10^{-7} \pm 2.7 \times 10^{-8}$	0.985
Nb 70	3b	$2.1 \times 10^{-7} \pm 4.8 \times 10^{-8}$	0.948
Nb 71	3c	$2.9 \times 10^{-7} \pm 3.3 \times 10^{-8}$	0.988
Nb 77	4	$3.0 \times 10^{-8} \pm 1.1 \times 10^{-9}$	0.998

Figure 3.9: **Identification and characterisation of nanobodies against the Shh-Ptc1 complex. A**, SEC profile and corresponding SDS-PAGE analysis of crosslinked Shh-Ptc1 ECD1-ECD2 complex (Superdex 200 16/60 column). Yellow arrowhead indicates the peak of higher molecular weight complexes, black arrowhead the peak of monomeric Shh-Ptc1 ECD1-ECD2 complex. SDS-PAGE of samples used for crosslinking experiments. Lanes from left to right: BenchMark^{TM} Marker (M), Shh (1), Ptc1 ECD1-ECD2 (2) and crosslinked Shh-Ptc1 ECD1-ECD2 sample as used for BLI experiments (marked by asterisk). **B**, Immobilisation of Ptc1 ECD1-ECD2 to optical biosensors for BLI analysis. **C**, Real-time binding chart for association and dissociation of Nb77 to Shh-Ptc1 ECD1-ECD2 (left). Statistical fitting of experimental data is shown in grey. Association and dissociation steps are divided by a black dotted line. Steady-state analysis (right) to determine the equilibrium dissociation constant. A K_D of $3.0 \times 10^{-8} \pm 1.1 \times 10^{-9}$ M. **D**, Dissociation constants of Ptc1-Nb complexes and corresponding coefficient of determination R^2.

For BLI experiments it is important to immobilise only one species of protein onto the tip to not falsify the results. For crosslinking, biotinylated Shh was used in combination with label-free Ptc1 ECD1-ECD2. The molecular weight of Shh is lower than the Ptc1 extracellular domain and can simply be separated from the complex by SEC. As Ptc1 ECD1-ECD2 is not biotinylated, minor contamination of free protein is hardly of concern as non-tagged protein does not bind to the biosensors and very little to no unspecific binding is expected due to the presence of 0.005% Tween-20 in the running buffer. Association of the complex to the sensor tips was smooth with response values of 5 nm and little dissociation after immobilisation. At the time of the experiment, only the first set of nanobodies was available and tested against the complex. In a BLI rapid screen setup Ptc1-specific Nb65, 66, and 67 as well as Shh-specific Nb69, 70, 71, and 77 were identified. Interestingly, this excludes Nb64, 68, and 75 that have previously been identified as Ptc1-specific nanobodies, and Nb78, a Shh-specific nanobody. The antigen binding site on either Ptc1 ECD1-ECD2 or Shh might not be fully accessible when Shh is bound to Ptc1. Of all nanobodies identified, Nb77 showed the tightest binding with a K_D of $3.0 \times 10^{-8} \pm 1.1 \times 10^{-9}$ M. A real-time binding chart used for steady-state analysis of Nb77 is shown in Figure 3.9, C. All other Shh-specific nanobodies showed a weaker binding affinities between 200-320 nM, similar to Ptc1-specific Nb65. Both Ptc1-specific nanobodies Nb66 and 67 showed values of roughly 100 nM (Figure 3.9, D). In comparison to binding studies with free Shh and free Ptc1 ECD1-ECD2, all nanobodies showed a weaker binding affinity to the complex, with the exception of Nb77.

3.4 Discussion

What conclusions can be drawn from our experiments at this point and what are the consequences for our future studies? With BLI measurements we were able to identify Shh-specific and Ptc1-specific nanobodies and determine their dissociation constants (K_D) when interacting with the ligand Shh or the receptor Ptc1 extracellular domain, as well as their crosslinked complex. Of the nanobodies that can bind either to Shh or Ptc1 ECD1-ECD2 alone, most, but not all, interacted with the complex. This observation indicates that accessibility of the binding interface and possibly conformational changes of the components during complex formation, modulate or prevent binding of the nanobody to its target. The following paragraphs will discuss our findings in the light of the Shh-Ptc1 ECD1-ECD2 complex.

The binding affinities of most Shh-specific nanobodies (e.g. Nb69, 70, and 71) to the Shh-Ptc1 ECD1-ECD2 complex were lower compared to Shh alone. This indicates that the nanobody's binding interface on Shh is partially inaccessible in the complex and only partial interaction between the nanobody and its antigen is feasible. In contrast, when Shh is in complex with Ptc1 ECD1-ECD2, Nb77 (also a Shh-specific nanobody) showed a two-fold increase in binding affinity. This observation points towards a potential conformational change of the Shh ligand when bound to Ptc1 ECD1-ECD2, which seems to have a positive effect on binding of Nb77. Is this fact also reflected in sequence similarities and differences of these nanobodies? Indeed, the CDR3s of Nb69, 70, and 71 are nearly identical, whereas the CDR3 of Nb77 revealed striking differences. The region is not only distinct from the other nanobodies in terms of sequence composition but also much shorter in comparison. This fact, combined with the observed increase in binding affinity to the complex suggests that Nb77 targets a different antigen on Shh (Figure 3.10). To support this

statement, BLI measurements will need to be repeated as in all experiments binding of nanobodies to Shh-immobilised biosensors could not reach equilibrium; the signal continued to rise. This might have an effect on the calculated K_D values. The reason why equilibrium binding could not be reached might be that initial complex formation occurred rapidly and residual Shh binding sites were occupied at a much slower rate after most Shh-Nb complexes were already formed. To reach equilibrium binding, experiments should be repeated with lower amounts of Shh presented on the tip. For the repeats it is likely that a similar trend of binding affinities would be observed.

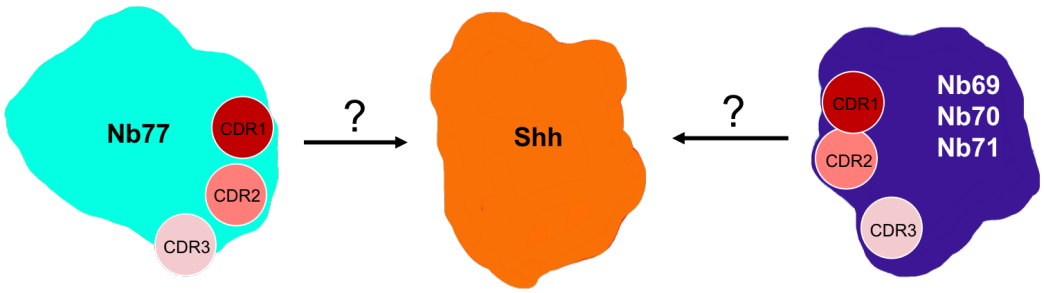


Figure 3.10: **Shh-Nb interactions: A binding model.** Schematic illustration. Minor differences in the CDR3 loop sequences of Nb69, 70 and 71 suggest that the nanobodies target the same binding site on Shh. The CDR3 loop sequence of Nb77 does not align with those of Nb69, 70 and 71 and may therefore interact with a different side of the antigen.

Interestingly, no binding of Ptc1-specific Nb64, 68 and 75 to the Shh-Ptc1 ECD1-ECD2 complex was observed. This suggests that as soon as the ligand-receptor complex is formed, the nanobody-binding interface on Ptc1 becomes inaccessible. But why do Nb65, 66 and 67 still interact with the complex? A closer look into the sequence alignment of Ptc1-specific nanobodies can help to interpret and explain this observation (Figure 3.8, B). The CDR3 loops of Nb64, 65, 66 and 67 are completely identical, however, their CDR1 loops differ considerably. This suggests that these nanobodies either use their CDR1 loop alone, or a combination of both regions to

interact with Ptc1 ECD1-ECD2. In the event that binding is carried out by the CDR1 loop only, it is tempting to speculate that all nanobodies (Nb64, 65, 66, 67) recognise a different antigen-binding site. The binding site of Nb64 might not be available after Shh-Ptc1 complex formation, whereas the binding interface of Nb65, 66 and 67 may still be accessible. However, this scenario is rather unlikely, as it has been shown that nanobodies with similar CDR3 loops recognise similar binding interfaces on the target protein [211]. In the case that the nanobodies use a combination of CDR loops, the target site for the nanobodies on Ptc1 ECD1-ECD2 might differ only slightly. Thereby, the CDR3 loops of Nb64, 65, 66 and 67 may not be able to access their binding interface on Ptc1 in the Shh-Ptc1 complex. The CDR1 loop of Nb64 alone may not be sufficient for establishing an interaction; on the contrary, this could be the case for Nb65, 66, and 67 (Figure 3.11). This scenario is further supported by the fact that these nanobodies show a lower binding affinity to the complex compared to the Ptc1 ECD1-ECD2 itself. That the CDR1 loop is not the only region used for interaction becomes obvious when having a closer look at the sequences of Nb65 and Nb75. Both have very similar CDR1 loops, yet they differ in their CDR2 and CDR3 loops. How can this be interpreted? Nb65 and Nb75 might have different binding interfaces on Ptc1 or binding is only sufficient when at least two out of three CDR loops can still interact with the protein. In any case, it is important that the interaction between the nanobody and the antigen is strong enough to form the nanobody-antigen complex. In summary our results suggest that binding of Ptc1-specific nanobodies to the target protein is utilised by the interplay of all three, but at least two CDR loops. The CDR3 loop might have a prominent function in interacting with the antigen, and binding is supported by the CDR1 and CDR2 loops.

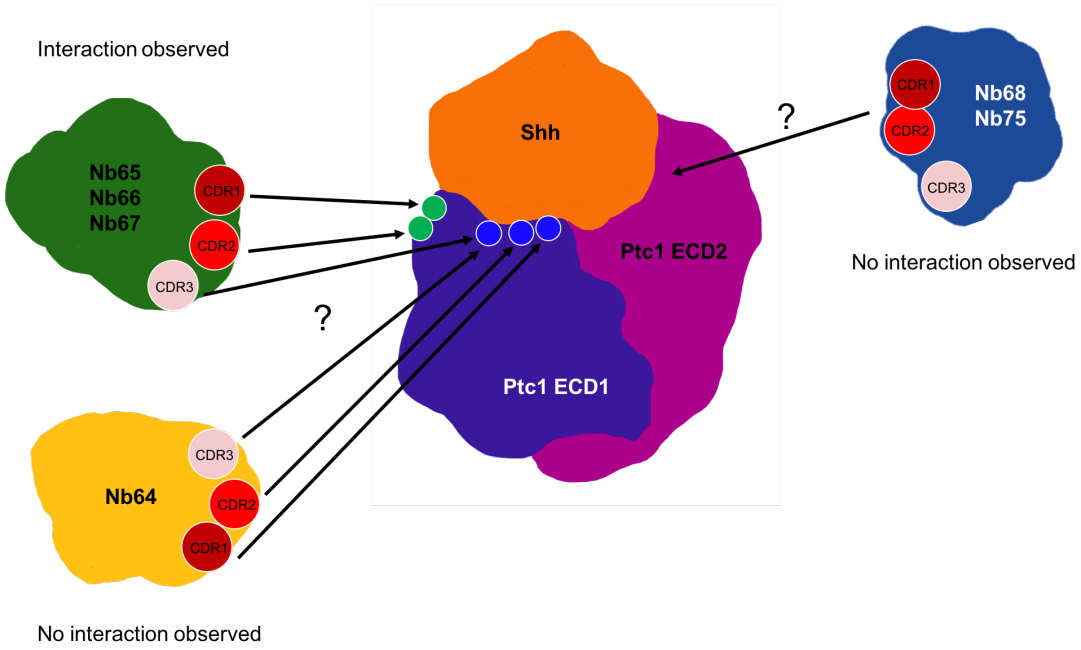


Figure 3.11: **Recognition of the Shh-Ptc1 complex by nanobodies: A binding model.** Schematic illustration. The binding interfaces of Nb64, 68 and 75 on Ptc1 may not be accessible in the Shh-Ptc1 complex as no interaction was observed in BLI experiments. In addition, their CDR3 loop sequences show no similarity, which suggests that they target different binding sites on the antigen. On the contrary, the Ptc1 binding interfaces of Nb65, 66 and 67 might be partially accessible. Their CDR1 and CDR2 loops may still be able to interact with Ptc1 in the Shh-Ptc1 complex, whereas the binding site of their CDR3 loop (which is identical to Nb64) may not. Accessible and inaccessible binding interfaces on Ptc1 are shown as green and blue dots, respectively.

Even though the second set of nanobodies was not tested against the complex, some conclusions can be drawn from their sequence analysis. The sequence of Nb790 and 791 is almost identical to Nb64, differing by a single amino acid residue substitution. Therefore it is very likely that these nanobodies do not show binding to the Shh-Ptc1 ECD1-ECD2 complex. Interestingly, Nb790 has a tyrosine to serine mutation close to its CDR3 loop, and also shows lower binding affinity to Ptc1 ECD1-ECD2 compared to Nb64 and Nb791. This indicates that this residue is important for Ptc1-Nb interaction. The CDR1 loops of Nb802, 803, and 804 share about 50% sequence identity with Nb75 and exhibit longer CDR3 loops, but are otherwise distinct

from each other. These nanobodies might have a different binding interface on Ptc1 ECD1-ECD2 and binding to the complex would depend on whether the interface is still accessible or not. Nb793 does not align with other nanobodies, and it is difficult to predict from sequence comparison whether it would still be able to interact with Ptc1 in the complex or not. Its CDR3 loop is identical to Nb64, but also Nb65. However, the CDR1 loop of Nb793 is unique. As for the second set of Shh-specific nanobodies, it is tempting to hypothesise that some of them would still be able to weakly interact with Shh even when the ligand is bound to Ptc1 ECD1-ECD2. This hypothesis is mainly based on the strikingly high sequence similarity of their CDR3 loops to the CDR3 loop of Nb71. To confirm this hypothesis, it is essential that BLI measurements of these nanobodies are carried out.

With 27 nanobodies against the Hh signalling components Shh and the extracellular domain of Ptc1, we have had a good repertoire of nanobodies for binding studies. These studies laid the ground for subsequent cellular studies to identify nanobodies that can activate or inhibit Hh signalling. Furthermore, this was the first step forward to use these nanobodies for structural studies of Shh and Ptc1. To date, it has been shown in many cases that nanobodies can facilitate crystallisation of difficult to crystallise protein targets. Our findings from cellular and structural studies are discussed in detail in the following chapters.

Chapter 4

Nanobodies are inhibitors of Hh signalling

4.1 Abstract

Hedgehog (Hh) signalling plays an important role in early embryonic development and fulfills important functions in renewal and maintenance of adult stem cells. Aberrant Hh signalling not only leads to developmental abnormalities but also to the occurrence of several types of cancer. Drug development to target Hh signalling in disease has mainly focused on the G protein-coupled receptor Smoothened (Smo) for treatment of basal cell carcinoma (BCC) and medulloblastoma (MB). Robotnikinin is the only known small-molecule inhibitor that targets the ligand Sonic Hedgehog (Shh) directly. Although not in the clinic yet, this small molecule modulator has the potential to target Hh ligand-dependent cancers (e.g. breast, colon, liver, lung, pancreatic, and prostate cancer) and developmental disorders by blocking signal relay to Patched-1 (Ptc1).

Targeting disease with antibodies or antibody fragments is not a new concept but a rather well-established strategy. The goal of this study was to identify nanobodies (Nbs) which show impeding effects on Hh signalling. We co-crystallised Shh in complex with several nanobodies, and verified their inhibitory potential in cellular assays. The structures gave insights into the molecular basis of the Shh-Nb interaction and revealed two distinct nanobody-binding sites on the antigen; one close to the zinc-binding site and one opposite the zinc-binding site. Based on structural analysis of eight Shh-Nb complexes, kinetic characterisation, and cellular assays, we selected two nanobodies, Nb70 and Nb77 to design a bispecific Shh nanobody. Nb70 and Nb77 recognise a different epitope on Shh, demonstrated the tightest binding, and reduced Hh signalling activity in embryonic fibroblasts. This study presents the first step towards the design of a novel nanobody-based Shh inhibitor and may lay the ground for treatment of Hh ligand-dependent malignancies.

4.2 Introduction

4.2.1 Nanobodies: From structures to inhibitors

The utilisation of antibody fragments as crystallisation chaperones dates back several decades [132, 149, 206]. This approach has often been employed if crystallisation efforts have proven unsuccessful. In many cases, antibody fragments rigidify a particular fold and thus reduce conformational heterogeneity, a phenomenon which often hampers crystallisation of proteins. In addition, antibody fragments assist in the formation of crystal contacts by increasing the hydrophilic surface of a protein. Not unexpectedly, this is particularly important in membrane protein crystallography. Commonly used antibody fragments are Fab fragments, F_v fragments and nanobodies. In contrast to Fab/F_v fragments, nanobodies possess the ability to insert one of their hypervariable regions into protein clefts without distorting the antigen's native fold [69, 146]. The spectrum of published nanobody-bound protein structures is broad, ranging from soluble ligands (e.g. Vitamin B12-binding protein BtuF [190], human cytokine Interleukin (IL)-23 [75], cellular tumor antigen p53 [23], human prion protein HuPrP [6]) and enzymes (e.g. E3 ubiquitin ligase Parkin PINK1 [248], trypsin-like serine protease urokinase-type plasminogen activator [151], Bloom's syndrome RecQ like DNA helicase [202], *Leishmania donovani* tyrosyl-tRNA synthetase [16]) to membrane proteins (e.g. Lipid-linked oligosaccharide flippase PglK [218], lactose permease LacY [134], μ-opioid receptor [123], human chemokine G protein-coupled receptor [37], human M2 muscarinic acetylcholine receptor [152], human β_2 adrenergic receptor [229]) and multidomain protein complexes (e.g. β_2 adrenergic receptor-G_s protein complex [229], *saccharomyces cerevisiae* endosomal phosphatidylinositol 3-kinase Vps34 complex [240], metazoan Nup62-Nup58-Nup54 nucleoporin complex [59]).

Nanobodies that target transmembrane receptors and/or their ligands have not only proven successful as crystallisation chaperones, but also have the potential to function as activators or inhibitors of cellular processes. S. Mireku *et al.* (2017) have identified a nanobody which targets BtuF, the ligand of *E.coli* ABC importer BtuCD-F [190]. The nanobody prevents vitamin B12 binding to BtuF, leading to inhibition of BtuCD-F function and therefore uptake of essential nutrients. The authors suggest that nanobodies that function as inhibitors of bacterial ABC transporters might evolve as "next generation antibiotics". With regard to multi-drug resistance of bacteria, alternatives to conventional antibiotics are critically needed. A. Desmyter and co-workers have designed a trivalent, bispecific anti-Interleukin (IL)-23 nanobody based on structural findings. Individual structures of human cytokine IL-23 in complex with three different nanobodies revealed three distinct binding sites on the antigen [75]. Neutralising effects on IL-23 were reported for a construct which places an anti-HSA (human serum albumin) nanobody in between two anti-IL-23 nanobodies. This multivalent nanobody construct has potential for treatment of autoimmune and inflammatory diseases such as Crohn's disease, rheumatic arthritis, and psoriasis.

The straightforward and cost-efficient pipeline for nanobody production, their small size, high solubility, and thermal robustness make them excellent tools for protein crystallographic studies. Structural analysis, combined with biophysical and cellular studies, present a useful toolkit for in-depth characterisation of nanobody-target interactions. Based on scientific findings, nanobodies can be tailored for biological and medical applications.

4.2.2 Targeting disease with bispecific antibodies

Bispecific antibodies (bsAbs) are molecules with dual specificity which recognise two distinct epitopes on the same or different antigen(s). Their potential to target two molecules simultaneously enables them to guide immune cells (T-cells) to tumour cells or to concurrently inhibit more than one signalling pathway. The first prototype of a bsAb was presented in 1997 by M. Coloma and S. Morison, who modified the C-terminus of an IgG-class antibody with a single-chain antibody of different specificity [61]. Since then, bsAbs have presented themselves in various facets. More than 60 different formats of bsAbs have emerged over the past 20 years and were tailored to fit a desired application [147, 255]. BsAbs can be divided into two major groups, those including partial or complete Fc regions, and those excluding Fc regions entirely [147]. BsAbs bearing Fc regions exhibit antibody-dependent cellular cytotoxicity (ADCC), complement fixation (CDC) and a longer half-life. In contrast, bsAbs lacking Fc regions are favoured for their small dimensions, simplified production, and low immunogenicity. Most newer generation bsAbs are represented by Fab fragments, fusions of single-chain variable fragments (scFv) or nanobodies. As of 2017, more than 130 bsAbs had been developed and forwarded into clinical trials [249]. To date, two success stories have been reported; Catumaxomab (Removab®) and Blinatumomab (Blincyto®) are bsAbs that have been approved in the EU and USA. Catumaxomab was launched to treat malignant ascites in epithelial cell adhesion molecule (EpCAM)-positive tumours when standard therapy is not feasable, whereas Blinatumomab was introduced to treat acute B-cell lymphoblastic leukemia [255]. Both drugs are T-cell retargeting agents that bring T-cells close to tumour cells. Catumaxomab belongs to the class of bsAbs with Fc domains. One arm of the antibody recognises the T-cell receptor cluster of differentiation 3 (CD3), whereas the

other arm identifies EpCAM-expressing tumour cells. The Fc region further interacts with Fc receptors presented on macrophages, dentritic cells or natural killer (NK) cells. Blinatumomab belongs to the class of bispecific tandem scFv antibodies, also called bispecific T-cell engager (BiTE®) which recognises T-cells and B-cells expressing CD3 and CD19, respectively. Even though Blinatumomab shows excellent renal clearance, its disadvantage comes with neurotoxicity and drug resistance for some patients [88]. In 2014, Blinatumomab was sold as one of the most expensive cancer drug on the market [2].

Beside the application of bsAbs for T-cell recruitment, bsAbs have been developed to target signalling pathways in disease. BsAbs have the advantage of reaching two targets simultaneously. This is important when multiple pathways exert overlapping functions. Unfortunately, blocking one signalling pathway often leads to the activation of another with similar function, thus introducing drug resistance. The use of escape mechanisms is commonly observed in cancer cells and are a known problem in cancer therapy. BsAbs reduce the possibility of switching on another signalling pathway and at the same time lower resistance development. Prominent drug targets are receptor tyrosine kinases, such as human epidermal growth factor receptor (HER) and insulin-like growth factor 1 (IGF-1) receptor which both influence the growth of tumour cells. The bsAb MM-111 was developed to target HER2-positive solid tumours by targeting HER2 and HER3 simultaneously. MM-111 prevents the dimerisation of the two receptors which leads to inhibition of PI3K/AKT signalling [185]. Similar to Blinatumomab, MM-111 belongs to the BiTE® family of antibodies, but carries an additional modified human serum albumin (HSA) domain to increase protein stability. MM-111 is currently in clinical phase II with the vision of treating breast cancer, as well as gastric and esophageal cancers. MM-141 (Istiratumab) is a tetravalent bsAb; stable scFv are fused to an IgG Fc region. Dual inhibition of HER3

and the IGF-1 receptor leads to inactivation of the PI3K/AKT/mTOR in advanced solid tumours such as hepatocellular carcinomas [93].

Bispecific nanobodies (bsNb) Ozoralizumab (ATN-103) and Vobarilizumab (ALX-0061) from Ablynx are currently being investigated for the treatment of rheumatoid arthritis. Ozoralizumab, a trivalent bsNb with two anti-tumour necrosis factor α (TNF-α) domains and one anti-HSA unit, neutralises TNFα and binds to HSA to prevent early renal clearance. A similar approach was taken for Vobarilizumab, which also utilises an anti-HSA domain. However, Vobarilizumab targets the interleukin 6 receptor (IL6R) via a single anti-IL6R nanobody domain. The phase II study on Vobarilizumab was completed in early 2018, with favourable safety findings. ALX0761 is another bsNb which targets the pro-inflammatory cytokines IL-17F and IL-17A. Both bind to the cell-surface receptor IL-17R and are linked to autoimmune as well as inflammatory disease. ALX0761 is currently in clinical phase I (healthy volunteers).

BsAbs and bsNbs are novel, protein-based T-cell recruiters and signalling pathway inhibitors. While bsAbs stand out for their stability, bsNbs come in a preferred size for medical applications, have been engineered for increased protein stability and exhibit low immunogenic potential. With two bsAbs already on the market and dozens more in the pipeline, bsAbs and bsNbs have high potential for treatment of cancers, autoimmune and inflammatory diseases. Those that are able to cross the blood-brain barrier could even target neurological diseases [89]. With the rise of multi-drug resistant bacteria and the failure of classical antibiotics, bsAbs and bsNbs have great potential in circumventing this problem [77]. BsAbs and bsNbs are rapidly evolving as next-generation antibodies.

4.3 Results

4.3.1 Preparation of Shh & Shh-Nb complexes for structural studies

In cells, active Shh is produced from a 45 kDa precursor protein. Internal intein cleavage results in a 20 kDa Shh N-terminal fragment (ShhN) and simultaneous attachment of a cholesterol moiety to its C-terminus. The 23 amino acid long signal peptide at the N-terminus is cleaved and subsequently modified with a palmitoyl residue by the Hedgehog O-acyltransferase (Hhat) [34]. Adapting this mechanism to produce the dual-lipidated active form of Shh *in vitro* is rather complicated. Thus, a genetically engineered construct was designed that replaces the N-terminal palmitoyl moiety and its attachment residue Cys^{24} (C24) with two consecutive Isoleucine (Ile, I) residues (ShhC24II). No additional modifications were introduced to the C-terminus of ShhN. Even though ShhC24II does not exhibit the same activity as its dual lipidated equivalent, it has been shown that introducing two hydrophobic amino acid residues to the N-terminus increases its potency by 8-fold when compared to the unmodified protein [265]. For crystallisation experiments, ShhC24II was the most suitable construct as it can be readily produced in high quantities and was also used for immunisation of llamas to obtain nanobodies against the Shh-Ptc1 ECD1-ECD2 complex. In addition, ShhC24II has previously been shown to crystallise and its crystal structure has been determined by the Siebold group [27, 287]. For reasons of simplicity, ShhC24II is referred to as Shh unless otherwise noted.

Human Shh with C-terminal His_6-tag was expressed in *E.coli* cells from a pET22b plasmid and subsequently purified via IMAC and SEC. The protein eluted as a single, slightly asymmetric peak with a peak maximum at ∼77 mL retention volume on a Superdex 75 16/10 column (GE Healthcare). All samples were analysed

on SDS-PAGE to assess sample purity (Figure 4.1, A). The analysis revealed a band at approximately 25 kDa, slightly higher than the expected molecular weight of 20.5 kDa. The peak fractions from SEC were pooled, concentrated to 5 mg mL^{-1}, and stored at -80°C until crystallisation experiments were carried out. The typical Shh purification yielded approximately 40 mg of protein from one litre of *E.coli* culture.

DNA sequences of Shh-specific nanobodies were received in pMESy4 expression vectors from our collaborators at VUB, Belgium [211]. Proteins were equipped with an N-terminal pelB leader sequence to direct the protein to the periplasm, and a C-terminal His_6-tag for purification. All nanobodies were produced by periplasmic expression in *E.coli* and purified from the periplasm via IMAC, followed by SEC. The SEC profiles displayed a single symmetrical peak for all Shh-specific nanobodies and SDS-PAGE analysis revealed a single band of the expected molecular weight of the respective protein. Representative SEC profiles and corresponding SDS-PAGE analysis of Nb70 and Nb77 are shown in Figure 4.1, B-C. SEC peak fractions were pooled and concentrated to 5 mg mL^{-1} except for Nb77, which precipitated at concentrations above 0.5 mg mL^{-1}.

For crystallisation experiments, Shh was prepared at a concentration of 5 mg mL^{-1} and the respective nanobody was added to achieve a 1:1 molar ratio. Prior to crystallisation, complexes were incubated on ice for 30 min in the presence of 3 mM $CaCl_2$. Most samples had a clear appearance, however, when adding Nb77 to Shh, the proteins precipitated. The precipitate was removed by centrifugation before crystallisation trials were set up. Hence, the remaining quantity of protein in the sample was lower than anticipated.

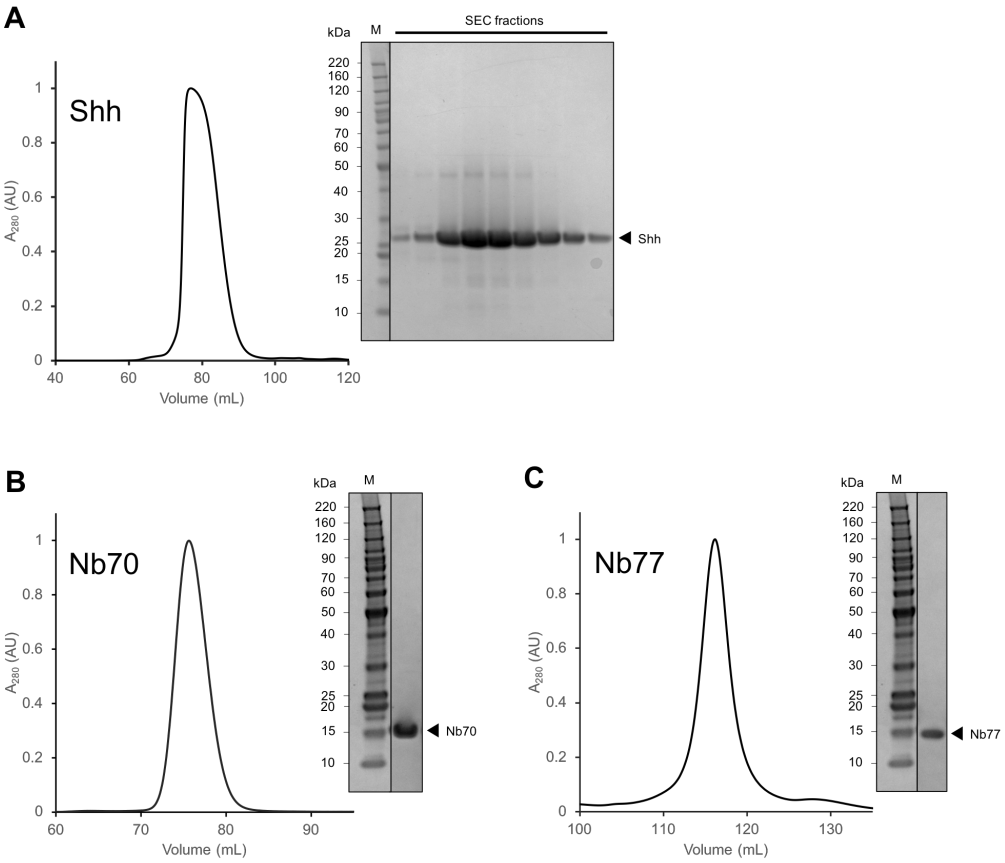


Figure 4.1: **Purification of Shh and Shh-specific nanobodies. A-C**, SEC profiles and corresponding SDS-PAGE analysis of Shh (A), Nb70 (B), and Nb77 (C). Proteins eluted as single monodisperse peaks and SDS-PAGE confirmed sample purity. M, BenchMark^{TM} ladder.

4.3.2 Shh crystallises readily in complex with nanobodies

Protein crystallisation is an excellent tool for identifying binding sites between two or more proteins. To date, several structures of Shh complexes have been solved and deposited in the Protein Data Bank (PDB). Our goal was to co-crystallise Shh with Shh-specific nanobodies to shed light on their molecular mechanism of interaction. For crystallisation, protein complexes were prepared as stated in Chapter 4.3.1 and crystallisation was set up using the sitting-drop vapour diffusion (VD) method [285]. To maximise the chance of crystallisation, a total number of four crystallisation

screens per Shh-Nb complex were set up in parallel at a 1:1 (protein:reservoir) drop ratio. Crystallisation screens used were Block 1 and Block 3 formulations, Morpheus® (Molecular Dimensions) and Index™ (Hampton Research). Crystals appeared after a few hours to several weeks depending on the Shh-Nb complex and crystallisation condition. Detailed information about the composition of mother liquour can be found in Chapter 2.9.6.1. Data sets of eight different Shh-Nb complexes were each collected from a single protein crystal on beamline I03 at the Diamond Light Source (Didcot, UK).

4.3.3 Nanobodies target Shh at two distinct binding sites

All eight crystal structures of Shh-Nb complexes were determined by the molecular replacement technique using structures of Shh (PDB code: 3D1M) and Gelsolin (PDB code: 2X1O) as templates. Depending on the Shh-Nb complex, structures were refined to 1.4- 2.8 Å resolution. Data collection and refinement statistics are shown in Table 4.1. The globular fold of Shh consisted of two motifs; one formed by two α-helices and another one comprised of six β-sheets (see Figure 4.3, A as a representative example). This sandwich-like fold is interconnected by an extensive network of loops. Furthermore, the structure revealed a zinc-binding site; two β-strands were involved in coordination of the Zn^{2+} ion. Zinc ions bound to proteins are not uncommon (e.g. in zinc hydrolases) and have important functions for the activity, stability and interactions between proteins. This observation also stands up for Shh, although Shh did not show hydrolase activity [27, 95]. The previously reported Ca^{2+} ions were not visible in the structures [27]. The overall fold of the nanobodies was comprised of two anti-parallel β-sheet motifs, arranged opposite of each other (Figure 4.3, A). The framework of β-strands was interspersed by loop regions, including the hypervariable loops CDR1-CDR3.

	Shh-Nb69	Shh-Nb70	Shh-Nb71	Shh-Nb77	Shh-Nb78	Shh-Nb794	Shh-Nb797	Shh-Nb799
DATA COLLECTION								
Beamline	DLS-I03	DLS-I03	DLS-I03	DLS-I03	DLS-I03	DLS-I03	DLS-I03	DLS-I03
Wavelength [Å]	0.97625	0.97625	0.97625	0.97626	0.97625	0.97626	0.97626	0.97626
Resolution [Å]	46.18-1.98 (2.03-1.98)	94.26-2.70 (2.77-2.70)	86.86-1.41 (1.45-1.41)	75.51-2.40 (2.44-2.40)	48.89-1.63 (1.67-1.63)	62.30-2.50 (2.54-2.50)	52.05-2.10 (2.10-2.09)	86.04-1.70 (1.73-1.70)
Space group	C2	C2	P1	P212121	P21	C2	C2	P21
Cell dimensions	a = 153.1 Å b = 67.8 Å c = 76.3 Å α = 90° β = 99.9° γ = 90°	a = 67.3 Å b = 92.3 Å c = 188.6 Å α = 90° β = 91.12° γ = 90°	a = 55.4 Å b = 67.1 Å c = 92.3 Å α = 106.6° β = 96.3° γ = 104.0°	a = 51.9 Å b = 98.5 Å c = 117.6 Å α = 90° β = 90° γ = 90°	a = 57.8 Å b = 67.8 Å c = 71.3 Å α = 90° β = 97.8° γ = 90°	a = 136.2 Å b = 45.9 Å c = 116.9 Å α = 90° β = 113.9° γ = 90°	a = 51.1 Å b = 51.4 Å c = 104.1 Å α = 90° β = 90° γ = 90°	a = 60.1 Å b = 86.0 Å c = 66.1 Å α = 90° β = 112.0° γ = 90°
Unique reflections	51558 (3006)	29854 (6510)	222071 (14197)	24336 (1183)	67922 (5013)	23205 (1172)	16746 (165)	67686 (2923)
Completeness [%]	96.0 (75.9)	93.9 (95.9)	94.8 (81.7)	100.0 (99.7)	99.8 (99.6)	99.4 (99.4)	99.5 (94.3)	98.7 (85.6)
R_{merge} [%]	4.5 (81.0)	8.4 (100.1)	7.0 (69.1)	23.5 (174.6)	5.5 (67.7)	19.6 (156.6)	19.4 (96.8)	8.9 (127.2)
R_{pim} [%]	3.5 (60.1)	7.8 (80.2)	5.4 (57.7)	10.7 (86.6)	4.1 (51.9)	13.0 (103.5)	11.6 (63.0)	5.8 (109.9)
$CC_{1/2}$	0.998 (0.562)	0.987 (0.315)	0.996 (0.443)	0.986 (0.326)	0.996 (0.597)	0.984 (0.323)	0.993 (0.711)	0.997 (0.240)
I/σI	11.7 (1.1)	7.8 (0.8)	8.8 (1.4)	5.7 (0.8)	8.7 (1.3)	5.8 (0.8)	9.4 (2.2)	8.1 (0.5)
Multiplicity	3.3 (2.6)	2.9 (2.9)	3.4 (2.9)	5.6 (4.8)	3.2 (2.7)	3.2 (3.2)	6.4 (5.7)	3.2 (1.9)
REFINEMENT								
Resolution range [Å]	46.18-1.98	94.26-2.70	86.86-1.41	75.51-2.40	48.89-1.63	62.30-2.50	52.05-2.09	61.35-1.70
Number of reflections (test set)	49016 (2542)	28419 (1435)	211042 (11029)	24111 (1201)	67412 (3356)	22034 (1170)	15857 (840)	67571 (3293)
R_{work} [%]	18.9	23.1	17.1	23.8	19.1	22.9	19.9	18.4
R_{free} [%]	22.4	26.3	19.3	27.7	22.2	27.4	25.0	21.8
Number of atoms								
Protein	4340	6647	8730	4569	4760	4455	2400	4841
Ligand	2	3	34	34	26	12	11	41
Water	253	-	1289	121	495	29	89	303
B factors [Å^2]								
Protein	48.5	102.2	17.1	39.4	29.2	58.2	32.5	39.2
Ligand	40.4	86.2	25.0	49.2	54.4	42.0	36.1	51.4
Water	47.4	-	29.0	32.8	36.6	35.0	26.4	40.7
RMSD bonds [Å]	0.018	0.019	0.016	0.002	0.003	0.014	0.006	0.017
RMSD angles [deg]	1.807	2.008	1.791	0.609	0.649	1.671	0.787	1.340
Ramachandran statistics [%]								
Favoured	97.4	94.1	99.4	96.9	97.6	94.6	97.4	97.3
Disallowed	0.2	0.7	0	0.2	0.0	0.9	0.0	0.5

Table 4.1: **Crystallographic table for Shh-Nb complexes.** Crystal structures of eight Shh-Nb complexes were obtained and solved at resolutions between 1.4 and 2.7 Å. Numbers in parentheses refer to the highest resolution shell. R_{pim}, precision indicating merging R-factor [286]; $CC_{1/2}$, correlation coefficient between random half data sets [141]; R_{free} equals the R-factor against 5% of the data. RMSD, root-mean-square deviation from ideal geometry. Ramachandran statistics were calculated using MolProbity [51].[3]

Seven out of eight nanobodies that co-crystallised with Shh interacted with an interface that is located near the zinc-binding site, close to where binding of Shh co-receptor Cdo and negative Hh regulator Hhip take place (Figure 4.2, A-G). These nanobodies recognised an almost identical surface on Shh, which can be seen in the

[3]Preliminary structure refinement statistics.

superposition of these seven structures (Figure 4.2, H). Nanobodies (excluding Nb77) showed an average root-mean-square deviation (RMSD) value of 1.56 Å when aligned over 120 C_α residues and mostly differed in their loop regions. Interestingly, the structure of the Shh-Nb77 complex revealed a completely different binding mode, with a binding site opposite of all other nanobodies (Figure 4.2, I). Alignment of Shh molecules over 144 C_α residues (Thr^{40} - Val^{185}) displayed a RMSD value of 0.37 Å. The low RMSD value indicated that the nanobodies did not alter the overall fold of Shh by much.

Of all Shh-specific nanobodies, Nb70 and Nb77 displayed the tightest binding to Shh in affinity measurements (Chapter 3.3.3.1) and strong inhibitory potential in cellular assays (Chapter 4.3.4). The following section will therefore pay particular attention to structural features of Shh-Nb70 and Shh-Nb77 complexes. The crystal structures of these complexes and the details of their binding interface are presented in Figure 4.3, A-D. The structures revealed the commonly observed fold of Shh and both nanobodies utilised their CDR loops to establish contact to Shh. The crystal structure of the Shh-Nb70 complex was solved at 2.8 Å resolution. The binding interface was located near the zinc-binding site of Shh and was formed by four salt bridges and seven hydrogen bonds, supported by 65 non-bonded contacts such as van der Waals and hydrophobic interactions. A detailed overview of interactions can be found in Figure 4.4. Residues involved in hydrogen bonding were located on CDR2 (Glu^{58}) and CDR3 (Asp^{103}, Glu^{109}, Ser^{111}), as well as the constant framework region (Tyr^{61}, Asp^{64}) of the nanobody. The four salt bridges were formed between CDR3 residues Asp^{103}, Glu^{109}, Arg^{110} and Shh residues Leu^{45}, Leu^{54}, Glu^{136}, and Glu^{137}. The highest number of atomic contacts was observed between residues Asn^{50} on Shh and Tyr^{61} on the nanobody framework, as well as Thr^{55} (Shh) and Asp^{64} (Nb70). In addition, a hydrophobic pocket was detected between Nb70 residues Tyr^{61}, Phe^{108}

and Shh residue Phe^{47} (Figure 4.3, C).

The crystal structure of the Shh-Nb77 complex was solved to a resolution of 2.4 Å. The nanobody formed its interaction network by using all three CDR loops and their surrounding framework regions (Figure 4.3 D, Figure 4.4). Four hydrogen bonds were observed between CDR1 (Thr^{28}, Tyr^{32}, Trp^{33}) and Shh (Asp^{83}, Gln^{100}, Asp^{104}), and another four between CDR3 (Asp^{99}, Asn^{102}) and Shh (Thr^{77}, Pro^{78}, Tyr^{80}, Lys^{103}). The CDR3 loop residue Asp^{99} formed an additional salt bridge to Lys^{103} of Shh. In addition, a total of 92 non-bonded contacts were found; strongest interactions were observed between CDR1 and CDR3 residues Val^{31}, Tyr^{32}, Trp^{33}, Asp^{99} and Shh residues Pro^{82}, Tyr^{80}, Asp^{83}, Asp^{104}, and Asn^{107}. In contrast to Nb70, no framework residues were involved in the formation of hydrogen bonds or salt bridges.

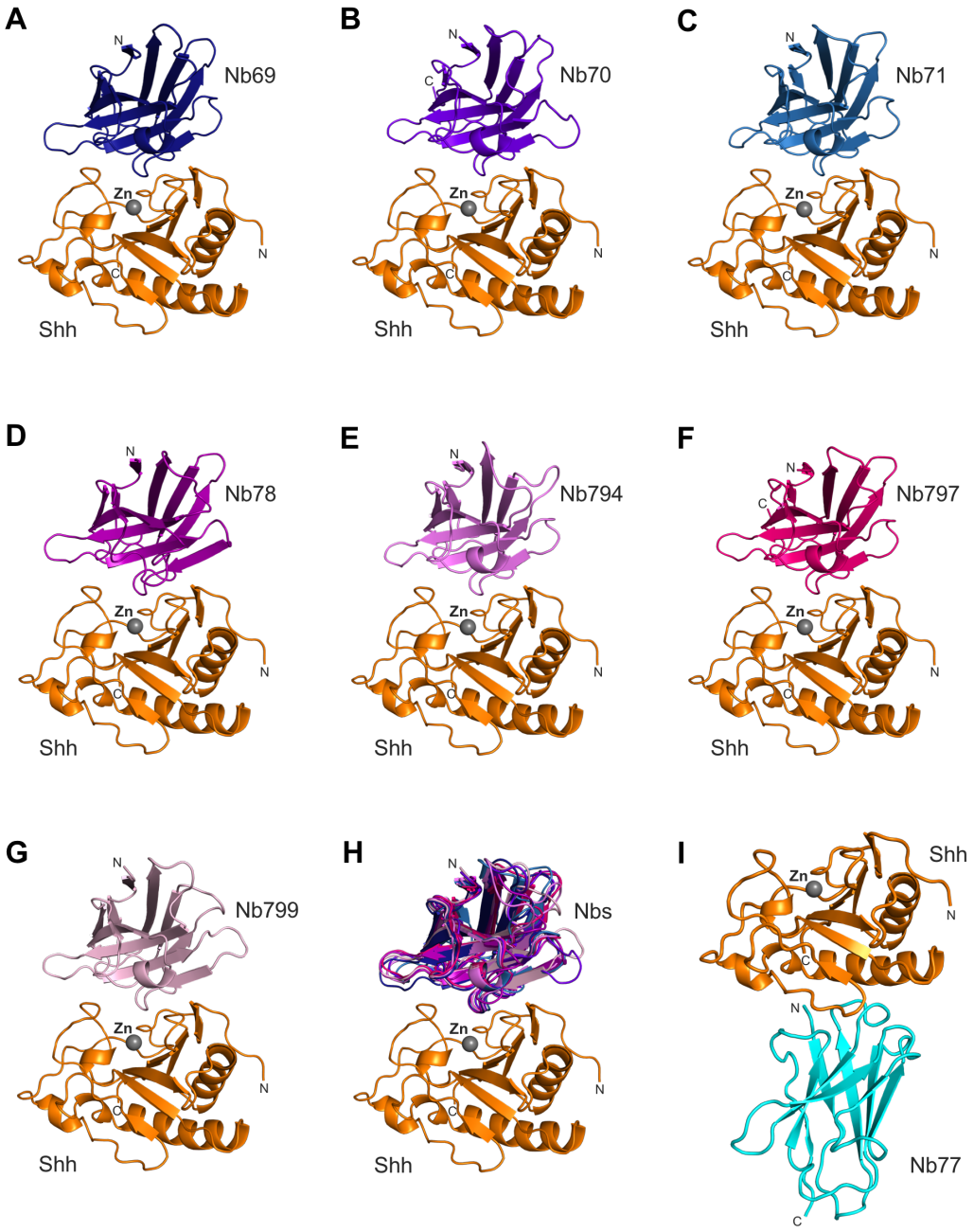


Figure 4.2: **Crystal structures of Shh-Nb complexes. A-G, I**, Structures are represented as cartoon diagrams with N- and C-termini indicated by N and C, respectively. Shh is represented in orange, Zn^{2+}-Ions as grey spheres. Shh in complex with Nb69 in density (A), Nb70 in purpleblue (B), Nb71 in skyblue (C), Nb78 in magenta (D), Nb794 in violet (E), Nb797 in hotpink (F), Nb799 in lightpink (G), and Nb77 in cyan (I). **H**, Structural alignment of nanobodies projected on Shh.

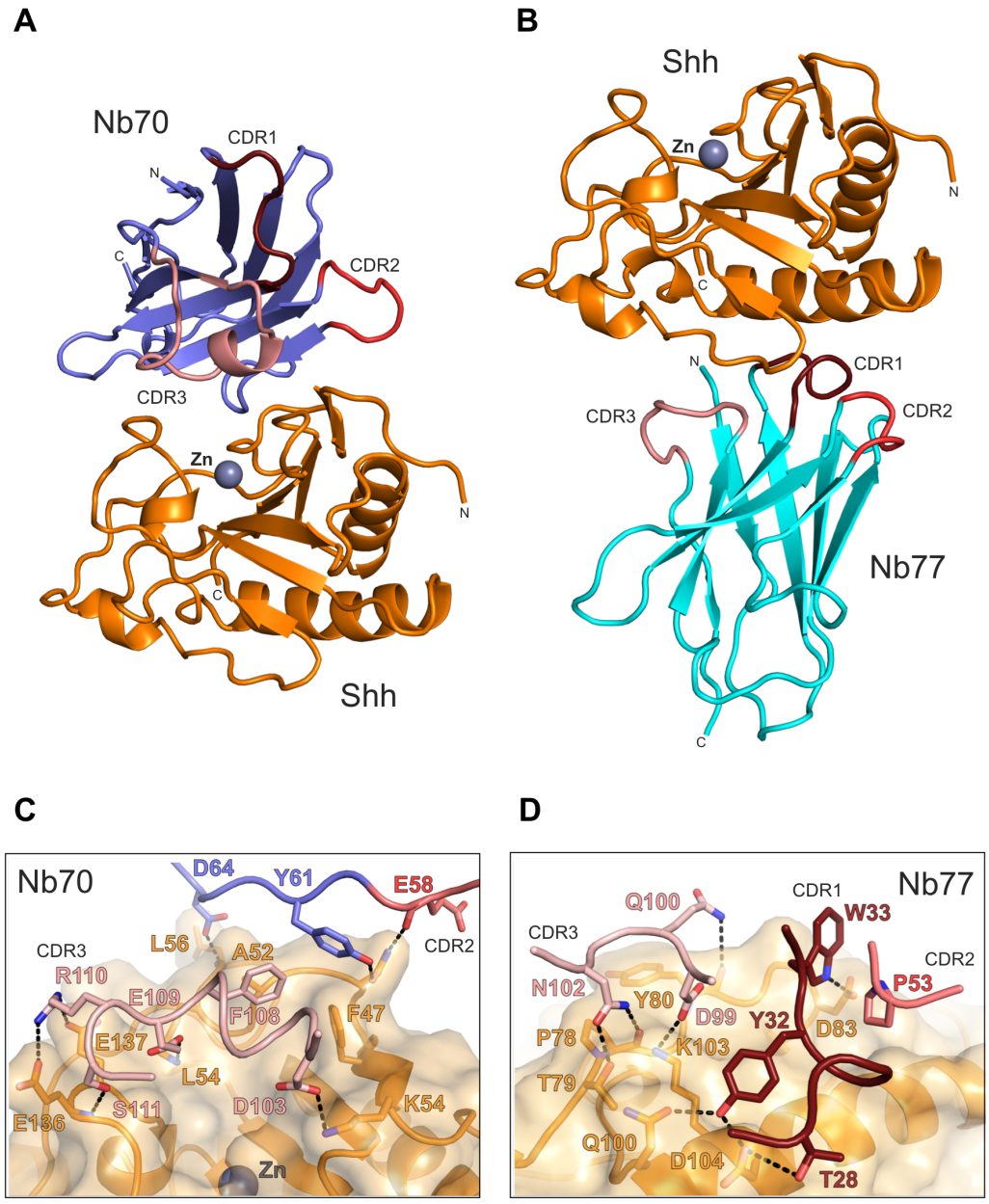


Figure 4.3: **Comparison of Shh-Nb70 and Shh-Nb77 complexes. A-B**, Crystal structures of Shh-Nb complexes represented as cartoon diagrams. Shh is coloured in orange, Nb70 in slate, and Nb77 in cyan. Zn^{2+} ions are shown as grey spheres. CDR loops are marked in ruby (CDR1), red (CDR2), and salmon (CDR3). N- and C-termini are indicated by N and C, respectively. **A**, Nb70 interacts near the Shh zinc-binding site. **B**, Nb77 interacts with Shh opposite to the zinc-binding site. **C-D**, Close-up view of Shh-Nb interfaces. Hydrogen bonds are indicated as black dotted lines. Oxygen atoms are colored in red, nitrogen atoms in blue. **C**, Nb 70 interacts with Shh via its CDR2 (red) and CDR3 (salmon) loops, as well as framework residues (slate). **D**, Nb77 utilises all three CDR loops for Shh-Nb interface formation.

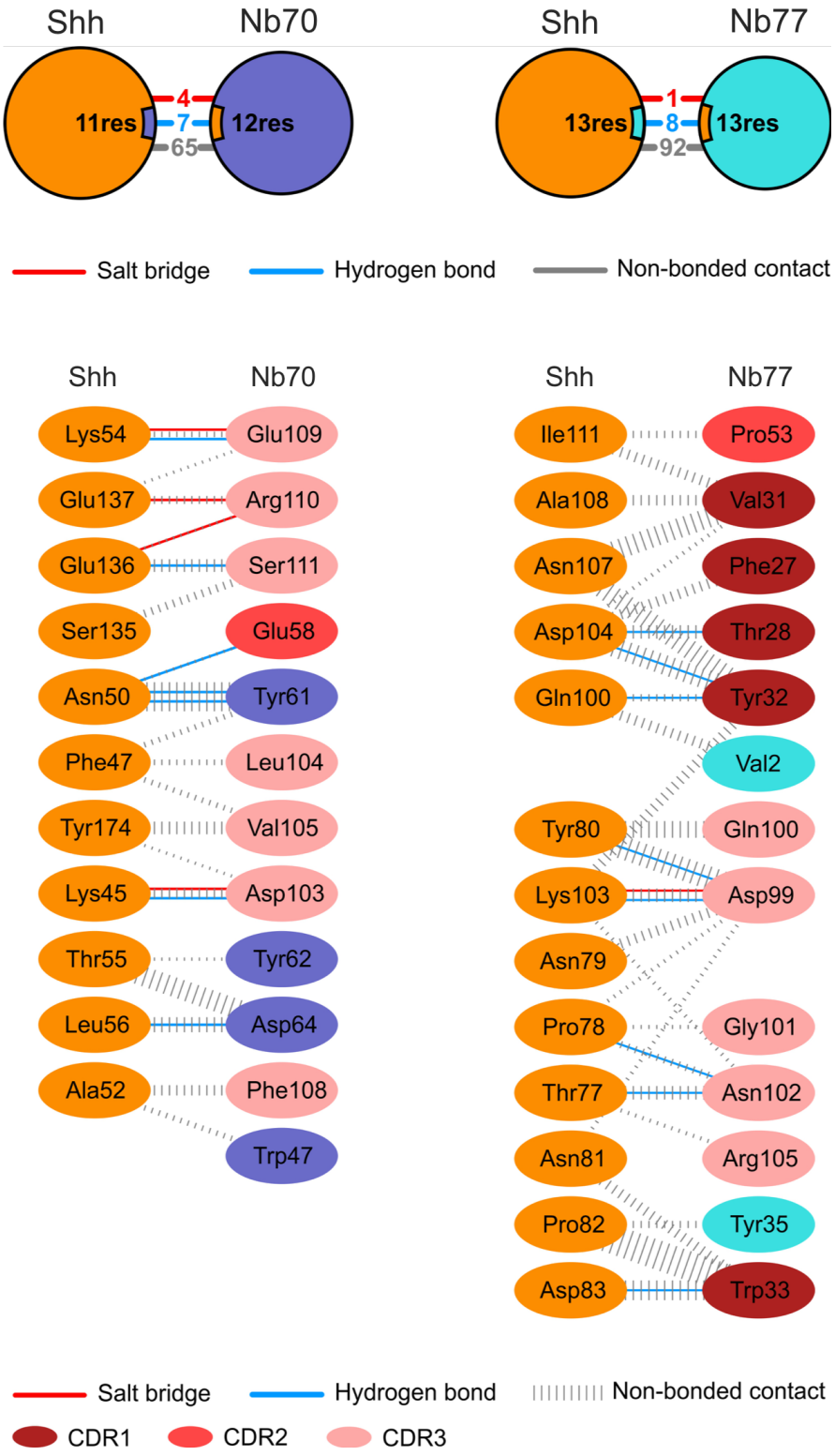


Figure 4.4: **Detailed overview of Shh-Nb70 and Shh-Nb77 interface residues.** A total of 11 residues of Shh and 12 residues on Nb70 are involved in interface formation. Shh and Nb77 are each contributing 13 residues for interaction. Interfaces are formed by salt bridges (red), hydrogen bonds (blue) and non-bonded contacts (grey/grey vertical dashed line). Figures adapted from PDBsum [157].

4.3.4 Shh-specific nanobodies inhibit Hh signalling in cellular studies

Upregulation of the Hh-specific Gli1 transcription factor is an indicator of Hh signalling activity. When Shh binds to its receptor Ptc1, downstream signalling is initiated by firstly relieving the inhibition of Ptc1 on Smo and secondly by upregulating activator forms of Gli transcription factors. Of all three Gli proteins (Gli1-3), Gli1 functions as a signalling activator only, whereas the other two Gli proteins, Gli2 and Gli3, additionally contain repressor domains. Therefore, measuring the transcriptional activity of Gli1 is a good indicator for signalling activity; high Gli1 levels indicate activation of Hh signalling. For Gli1 transcription factor reporter assay experiments, which are suitable for recording Gli1 mRNA levels, mouse NIH/3T3 fibroblasts were grown in a 96-well plate format. After starving the cells for one day, cells were treated with 50 nM Shh and 25 μM Shh-specific nanobody and incubated for 24 hours. Cells were lysed and real-time quantitative reverse transcription-PCR (Real-time qRT-PCR) was used to generate a cDNA library. Gli1 cDNA was amplified by real-time PCR using a Gli1 specific set of primers. Gli1 mRNA levels were normalised against the house-keeping gene GADPH.

Shh-specific nanobodies, previously identified to bind to Shh in BLI kinetic measurements (Chapter 3.3.3.1), were tested in Gli1 transcription factor reporter assays. In order to characterise the nanobodies in terms of their inhibiting or activating potential, Gli1 mRNA levels of Shh-responsive cells were measured after treatment with nanobody or Shh-Nb complexes (Figure 4.5, A). Cells were exposed to Shh and the Smo small molecule-inhibitor Sant-1 (positive control) [49], nanobodies or Shh-Nb complexes. The negative control was represented by treatment with protein buffer only. As anticipated, endogenous Gli1 mRNA levels were low prior to treatment

with Shh, and increased drastically after Shh exposure, indicating that Hh signalling was activated. Treatment with the Smo high-affinity antagonist Sant-1 or a Sant-1/Shh combination yielded a low Gli1 mRNA response, suggesting inhibition of Hh signalling.

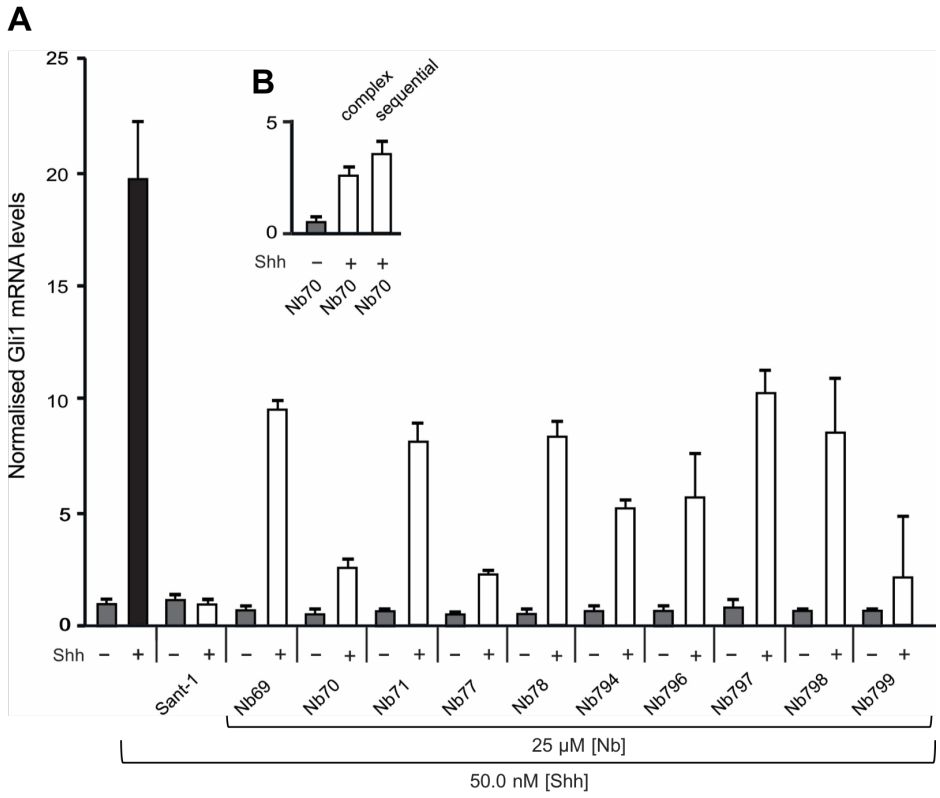


Figure 4.5: **Assessment of Gli1 mRNA levels in response to nanobody administration.** **A**, NIH/3T3 cells were treated with nanobody or Shh-Nb complexes and Gli1 mRNA response levels were assessed 24 hours post-exposure. Smo small molecule-inhibitor Sant-1 was included as positive control. Gli1 mRNA response values in absence or presence of Shh are shown in grey and white, respectively. Control experiment with Shh is shown in black. **B**, Comparison of treatment of NIH/3T3 cells with Shh-Nb70 complex and sequential administration of Shh and Nb70. No significant difference was observed. Assays were performed in collaboration with R. Siebold-Schwab.

All Shh-specific nanobodies in complex with Shh caused Gli1 mRNA levels to decrease, indicating that the nanobodies had inhibitory potential (Figure 4.5, A). On the contrary, they had no effect without the addition of Shh. Shh-Nb70 and Shh-

Nb77 complexes showed the strongest effect on Hh signalling ability, with Gli1 mRNA levels being reduced approximately 8-fold. This corresponds to the observation that Nb70 and Nb77 showed the tightest binding in Shh-Nb binding studies (compare to Chapter 3.3.3.1). Nb69, 71, 78, 797 and 798 decreased the Gli1 mRNA output two-fold, whereas Nb794, 796, and 799 could be allocated in between these two groups. To investigate whether simultaneous or sequential treatment of cells would lead to differences in Gli1 mRNA response, a control experiment with Nb70 was performed. The protein was either presented to the cells as a Shh-Nb70 complex, or Shh was added prior to Nb70 (Figure 4.5, B). Importantly, no significant difference in Gli1 mRNA response was observed.

4.3.5 The design of bispecific nanobodies to target disease

Our structural analysis of Shh-Nb complexes revealed two distinct binding sites on Shh (Figure 4.3). Most nanobodies favoured interaction close to the zinc-binding site, whereas Nb77 targeted a different epitope on Shh. Interestingly, those epitopes were not overlapping, indicating that two nanobodies can bind simultaneously to one Shh molecule. Of all nanobodies that interacted with Shh near the zinc-binding site, Nb70 showed the highest binding affinity ($4.4 \times 10^{-8} \pm 5.3 \times 10^{-9}$ nM). Furthermore, Nb70 and Nb77 were both able to significantly reduce Gli1 mRNA levels in Gli1 mRNA transcription reporter assays, suggesting that both nanobodies interfere with Hh signalling, potentially acting as inhibitors. Based on these results, we designed a bispecific nanobody by introducing a Gly/Ser (G/S)-rich linker to interconnect Nb70 and Nb77. We assessed a total of six nanobody fusion constructs; Nb70 and Nb77 were interlinked by a 10-, 20-, or 30-residue G/S-linker with either of the nanobodies first in sequence (see Appendix B for details). To predict a sufficient length for the linker, we superimposed the Shh structures of the Shh-Nb70 and Shh-Nb77 complexes

in PyMOL and measured the distance between the C-terminus of Nb70 and the N-terminus of Nb77 (Nb70/Nb77) and vice versa (Nb77/Nb70). This revealed distance values of 52 Å for Nb70/Nb77, and 84 Å for Nb77/Nb70 (Figure 4.6, A). On average each amino acid contributes 3.5 Å to the linker length. Hence, we estimated that a 20-residue linker should enable the Nb70/Nb77 construct to bind with both nanobodies to Shh simultaneously. For the reverse (Nb77/Nb70) we speculated that a 30-residue linker might be sufficient for cooperative binding. The 10-residue linker constructs were included as negative controls because the linker length of approximately 35 Å would not be sufficient to facilitate binding of both nanobodies at the same time.

Constructs were synthesised by overlapping PCR and cloned into a pHLsec vector background with N-terminal secretion signal and C-terminal His_6-tag. All proteins were expressed in HEK293T cells and purified by IMAC, followed by SEC. Proteins appeared as symmetric monodisperse peaks, and retention times (directly proportional to retention volumes) gradually increased for constructs with longer linker length. For constructs with Nb70 first, proteins eluted at retention volumes of ∼67 mL, 66 mL, 65.5 mL for the 10-, 20-, and 30-residue linker, respectively, and for constructs with Nb70 second at ∼72 mL, 70 mL, 68.5 mL (Figure 4.6, B-C). The IMAC purification step and selected SEC fractions were analysed on SDS-PAGE, revealing a band at the expected molecular size of roughly 30 kDa (Figure 4.6, D). Proteins appeared as single bands in western blot analysis (Figure 4.6, E). No degradation products were observed.

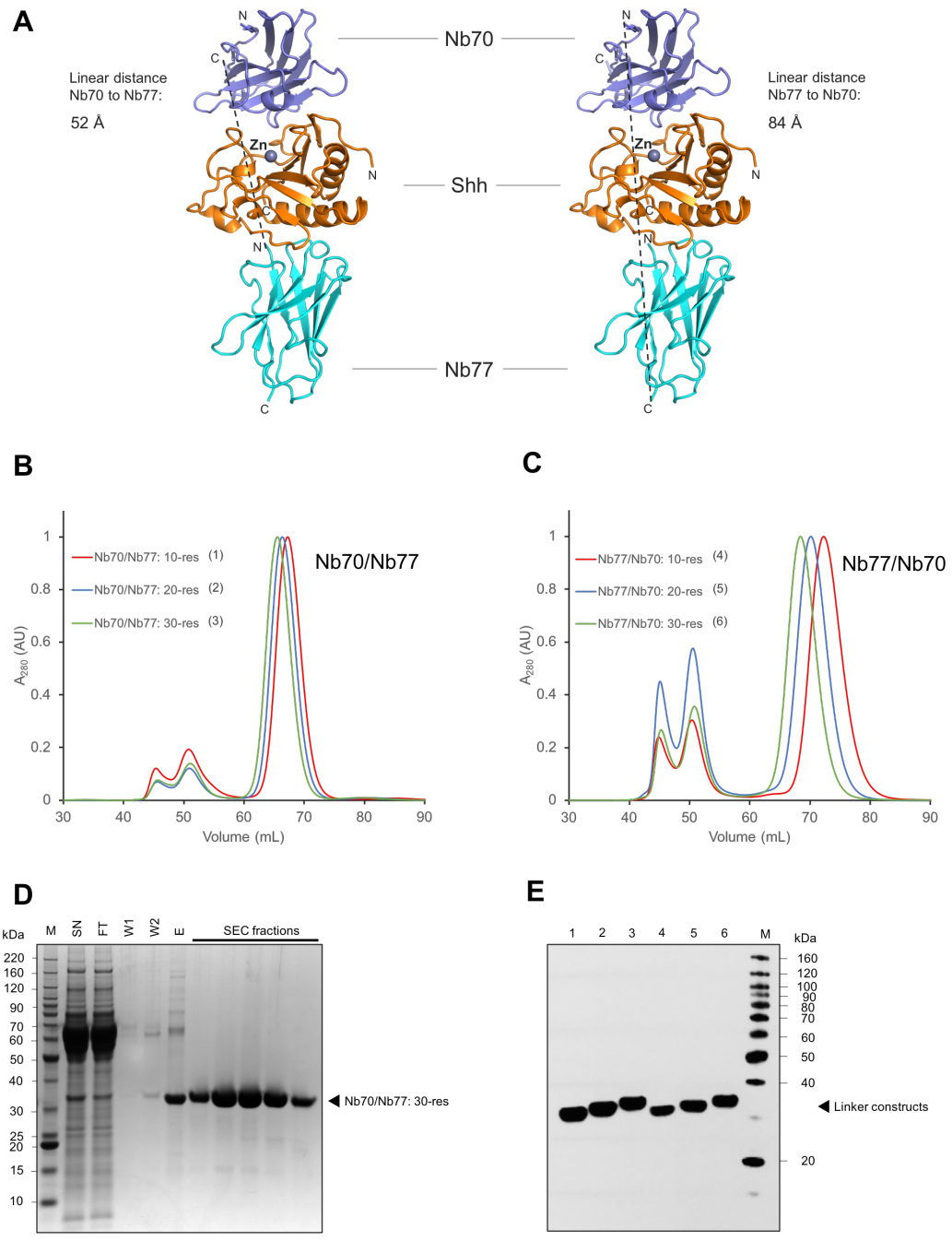


Figure 4.6: **Design and production of bispecific nanobodies.** **A**, Structural model of Shh in complex with Nb70 and Nb77 as used for N- to C-termini distance measurements in PyMOL. Linear measurements between the termini revealed distances of 52 Å and 84 Å. **B-C**, SEC profiles of Nb70/Nb77 (B) and Nb77/Nb70 (C) constructs with different linker lengths. Higher molecular weight complexes were successfully separated from monomeric protein. **D-E**, Sample purity was assessed on SDS-PAGE (D) and Western Blot (E). M, BenchMark^{TM} ladder; SN, Supernatant; FT, Flow-through; W1, Wash 1; W2, Wash 2; E, Eluate.

4.3.6 Bispecific nanobodies increase the inhibition potential

To confirm our hypothesis, bispecific nanobodies were tested for their inhibitory function in Gli1 mRNA transcription reporter assays. Endogenous Gli1 mRNA levels were measured after treatment of NIH/3T3 cells with 50 nM Shh and either 0.5 μM or 5.0 μM of bispecific or monospecific nanobody (Figure 4.7). Controls included treatment with Shh alone, Smo small molecule-inhibitor Sant-1, and treatment with protein buffer only. As anticipated, exposure of cells to either protein buffer or Sant-1 yielded in no activation or almost complete inhibition of Smo, respectively. Exposure of cells to Shh and 0.5 μM of either Nb70 or Nb77 led to similar Gli1 mRNA levels compared to the Shh control. This suggests that at this concentration of nanobody, Hh signalling could not be switched off. Treatment of cells with Shh and both nanobodies simultaneously caused only a minor reduction of Gli1 mRNA levels and thus could not suppress downstream signalling either. Administration of bispecific nanobody constructs Nb70/Nb77, however, gradually decreased Gli1 mRNA levels with increasing linker length. The 10-residue linker construct lowered the levels by approximately half, whereas the 20-residue and 30-residue linker constructs decreased Gli1 mRNA levels by 4.5-fold and 6-fold, respectively. These results suggest that in this orientation, the bispecific nanobodies have the ability to reduce Hh signalling activity. For bispecific nanobodies Nb77/Nb70, all three linker construct showed a two-fold decrease in Gli1 mRNA levels compared to the control. Values were comparable to those of the 10-residue linker construct in the reverse orientation. Increasing the concentration of bispecific nanobody to 5.0 μM almost suppressed Gli1 mRNA levels completely. Single nanobodies showed comparable, but very minor changes in Gli1 mRNA response. Taken together, our results suggest that bispecific nanobodies with sufficient linker length and flexibility represent potential inhibitors of Hh signalling

and that inhibition is concentration dependent.

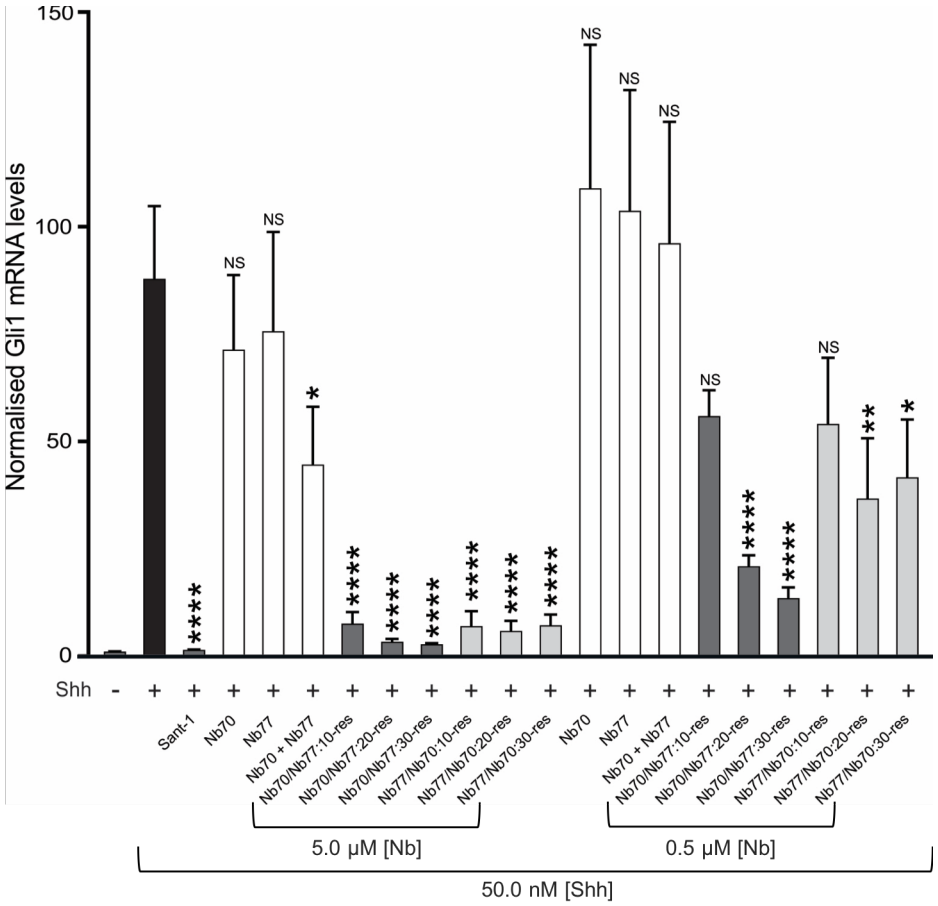


Figure 4.7: **Influence of bispecific nanobodies on Gli1 mRNA response levels in Hh responsive cells.** Assessment of Gli1 mRNA levels after exposure of NIH/3T3 cells to Shh in complex with mono- or bispecific nanobodies at two different concentrations. Statistical significance was assessed by the one-way ANOVA test combined with a Holm-Sidak post-test for the correction of multiple comparisons. Statistical significance (p-value) is denoted as follows: NS (not significant); $p>0.05$, $*$; $p\leq 0.05, **$; $p\leq 0.01, ***$; $p\leq 0.001, ****$; $p\leq 0.0001$. Assays were performed by R. Siebold-Schwab.

To investigate whether bispecific nanobodies induce oligomerisation and/or precipitation of Shh, protein complexes were analysed by analytical SEC using a Shimadzu chromatographic system. Shh was prepared at 1 mg mL^{-1} and the respective nanobody was added at a 1:1 molar ratio to form the Shh-Nb complexes. Proteins were incubated on ice for 15 min before being applied on a Superdex 200 3.2/300

column (GE Healthcare). Nb70/Nb77:20-residue and Nb70/Nb77:30-residue linker constructs displayed a shift in peak position when in complex with Shh. This indicated that these two bispecific nanobodies form a 1:1 complex with Shh. Decreasing the linker length to 10 residues (Nb70/Nb77:10-residue) caused the proteins to precipitate, suggesting that the bispecific nanobodies induced oligomerisation and thus formation of higher molecular weight complexes. In the case of Nb77/Nb70 constructs, all complexes precipitated in the presence of Shh. Figure 4.8, A-B shows an overlay of respective SEC profiles.

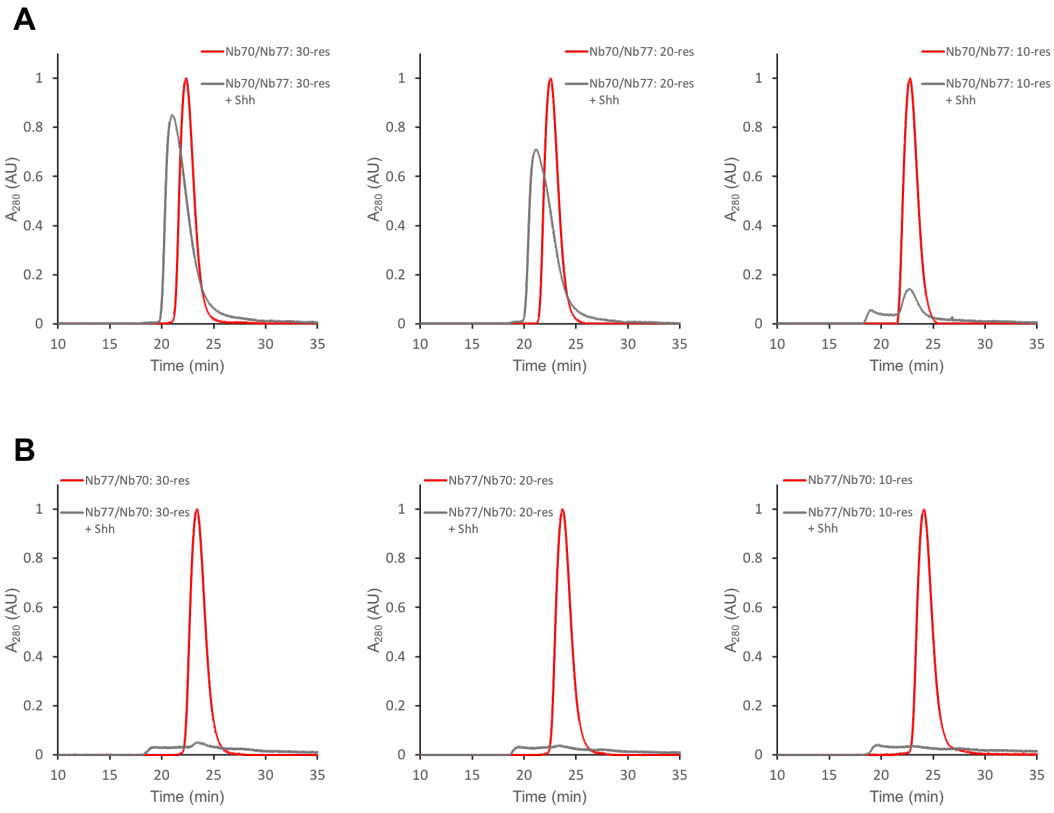


Figure 4.8: **Analysis of Shh-bispecific nanobody complexes by SEC. A-B**, Bispecific nanobodies designed with different G/S-linker length (10-, 20-, or 30-residues) and either Nb70 (A) or Nb77 (B) first in sequence. Constructs were analysed on analytical SEC with and without Shh.

4.4 Discussion

The Shh-Nb structures presented in this chapter are the first Shh structures determined in complex with a nanobody. We investigated the molecular interaction between Shh and nanobodies with the aim of designing a novel nanobody-based inhibitor of Hh signalling. Judging from our results, nanobodies are promising drugging agents for targeting Hh ligand-dependent malignancies.

As expected, the nanobodies contacted the Shh surface via at least two of their three CDR loops, and the interaction was supported by a few framework residues. Nanobodies, in contrast to other antibody fragments, are known for inserting one of their loops into protein clefts of the antigen [69]. However, this was not observed for the Shh-Nb complexes. Instead the nanobodies recognised a rather flat, slightly convex surface on Shh, either close to the zinc-binding site or opposite. Recognition of convex paratopes is not uncommon for nanobodies and has been described in the literature many times before [197]. A combination of hydrogen bonds, salt bridges and non-bonded contacts formed the interaction platform between Shh and the nanobody, achieving binding affinities in the low nanomolar range. Of the eight nanobodies that crystallised in complex with Shh, all indicated inhibitory function in cellular assays. The lowest Gli1 mRNA response levels were achieved when treating the cells with Nb70 and Nb77. However, relatively high concentrations of nanobody (25 μM) had to be used to reduce the Hh signalling response to 12.5%- 15.0% relative to the Shh control. How can this observation be explained? The first possible explanation is that nanobodies can only block the interaction between Shh and its receptor Ptc1 and/or its co-receptors to a limited extent. When the nanobody is bound, the binding interface is not fully accessible. As a result, the output signal is weakened but can still be passed on to Smo, which initiates downstream intracellular signalling. Second,

Shh might carry more than one epitope for protein-protein interaction. Both epitopes can either be recognised by Ptc1 alone or by Ptc1 and a co-receptors simultaneously. In fact, our findings suggest that Ptc1 functions as a dimer. Provided that Shh is in complex with a Shh-specific nanobody, only one epitope is fully or partially accessible and can be recognised by Ptc1 or the co-receptors. In addition it has been postulated that Ptc1 regulates Smo via a small molecule such as cholesterol [264]. When partially blocked, Ptc1 might not be able to sufficiently deliver to or take away cholesterol from Smo, resulting in partial pathway inactivation. Even though all Shh-specific nanobodies (except Nb77) revealed a similar binding site on Shh as Nb70, they achieved lower inhibitory effects and also displayed lower binding affinities and faster off-rates. This might lead to continuous dissociation of the nanobodies from Shh, giving the ligand the possibility to interact with Ptc1 and Shh co-receptors. Is this theory also supported by sequence similarities or differences of the nanobodies? Indeed, Nb70 displays an extended CDR2 loop, created by an insertion of two amino acid residues. In addition, among all CDR3 loops of Shh-specific nanobodies, Nb70 displays the most striking difference. Nb70 residue Phe^{108} is involved in the formation of a hydrophobic pocket together with Tyr^{61} of the nanobody framework and Phe^{47} of Shh. A Phe^{108}Met substitution as observed in all other nanobodies that target the zinc-binding site, preserves, but lowers the hydrophobicity of the pocket. Nb78 differs significantly in its CDR3 region and displays a Phe^{108}Cys substitution. Not only is cysteine less hydrophobic, our structural data also revealed a disulfide bridge between Cys^{108} and Cys^{50}. This may explain its 15-fold decrease in binding affinity.

Can the inhibitory function of Shh-specific nanobodies be explained on the basis of existing structures of Shh complexes? Partly - Structures of Hh pathway members Cdo and Hhip displayed binding to the zinc-binding site and its adjacent positively charged surface area [27, 186]. Cdo is a Hh signalling agonist; ITC affinity

measurements between its third fibronectin repeat (CDOFn3) and Shh displayed a K_D in the micromolar range (1.3 μM) [186]. The rather low binding affinity suggests that the nanobodies could (at least partially) outcompete Cdo. Hhip acts as a negative regulator of Hh signalling and has a low nanomolar K_D (6.0 nM) [27]. The fact that Nb70 displays overlapping regions with the Hhip binding site on Shh points towards a similar inhibitory mechanism. This observation suggests that nanobodies fully or partially prevent interaction of Shh with Cdo and assist Hhip in negatively regulating Hh signalling. Figure 4.9 compares binding sites of Cdo and Hhip to the binding interfaces of Shh-specific nanobodies.

Ptc1 might recognise a similar surface area on Shh as Cdo and Hhip, but also Nb70. Steric interference studies suggest that residues Asn^{50}, Asn^{115}, Ser^{135}, Ser^{156}, Gly^{169}, and Ser^{177} are close to the Ptc1 binding site [216]. Interestingly, Shh interacted with Nb70 via residues Asn^{50} and Ser^{135}, forming three hydrogen bonds and two non-bonded contacts. This indicates that parts of the Ptc1 binding interface can be obstructed by the nanobody. Residues Ala^{52}, Glu^{136}, and Glu^{137} on Shh are in close proximity to the predicted interface residues and might interfere further with the ability of Ptc1 to interact with Shh. The antibody 5E1, a known inhibitor of Hh signalling that is proposed to compete with Ptc1 for Shh binding [182], uses a slightly bigger but overlapping surface area on Shh compared to Cdo and Hhip (Figure 4.9, C). This supports the theory that the zinc-binding site and surrounding areas have essential functions for initiation of Hh signalling [182]. Shh residue Lys^{54} that is targeted by 5E1 is also involved in binding to Nb70 via one hydrogen bond, one salt bridge and also a non-bonded contact.

Even though these observations can explain the inhibitory function of Nb70, one question remains unanswered. The inhibitory function of Nb77 cannot be suffi-

ciently explained.[4] However, the results may suggest that Ptc1 functions as a homo- or heterodimer.

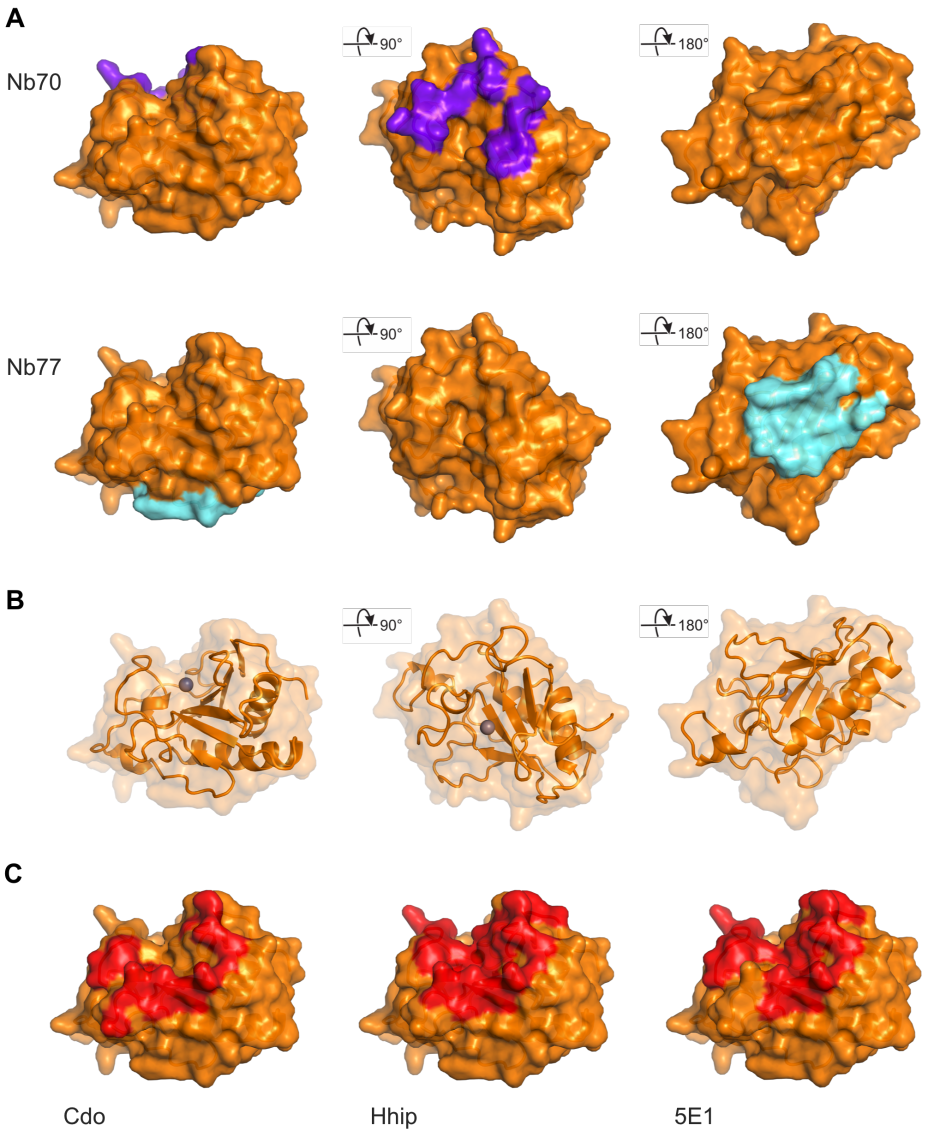


Figure 4.9: **Interface comparison of Shh binding partners.** Shh is shown as surface representation. Nanobody and Shh co-receptor interfaces mapped on Shh. **A**, Interface residues are coloured in purpleblue for Nb70, and in cyan for Nb77. View rotated by 90° and an additional 180° around the horizontal axis. **B**, Ribbon diagram of Shh fitted into surface representation. Rotation around horizontal axis to fit views in A. **C**, Interface residues of Shh co-receptor Cdo (PDB code: 3D1M), negative regulator Hhip (PDB code: 3HO5) and Shh-specific antibody 5E1 (PDB code: 3MXW) are coloured in red.

[4]During the time of writing this thesis, two full-length structures of the Shh-Ptc1 complex have been published, which shed light on the inhibitory function of Nb77 [227, 228]. Our results in context of these findings are discussed in Chapter 7.

Based on our structural analysis of Shh-Nb complexes, high affinity binding of nanobodies to Shh, and the satisfactory inhibitory function of Nb70 and Nb77 in cellular assays, we suspected that the design of a bispecific nanobody could further improve the nanobody's inhibitory effect. The model used for the design of the bispecific nanobodies was based on a structural overlay of Shh-Nb70 and Shh-Nb77 complexes in PyMOL. We estimated that a 20-residue linker should allow both nanobodies of the Nb70/Nb77 construct to bind to Shh simultaneously. As C- and N-termini of the nanobodies are further apart in the Nb77/Nb70 construct a minimum of 30 residues would be required. To verify our hypothesis, we tested the bispecific nanobody constructs in cellular assays and found that for Nb70/Nb77 constructs the inhibitory potential increased with increasing linker length. For Nb77/Nb70 constructs, no correlation between linker length and inhibitory potential was observed. As the distance between the C-terminus of Nb77 and the N-terminus of Nb70 is much larger than for the reverse, we suspect that a 30-residue linker might not have been sufficient to allow binding of both nanobodies to one Shh molecule simultaneously. When we investigated the behaviour of bispecific nanobodies in complex with Shh on analytical SEC, precipitation was observed for all Nb77/Nb70 constructs. Therefore, further experiments should include Nb77/Nb70 constructs with longer linkers. The additional length and flexibility of a 40- or 50-residue linker might enable the Nb77/Nb70 bispecific nanobody to inhibit signalling to a larger extend, possibly without introducing oligomerisation of the proteins. As to whether oligomerisation is induced by bispecific nanobodies also depends on entropy and avidity. If the stability of the construct is enhanced by simultaneous binding of both nanobodies to Shh, oligomerisation should be reduced. In summary, our results indicate that the potency of the bispecific nanobodies is not only dependent on the linker length but also on the sequence order of the nanobodies.

The expectations of monomeric, as well as bispecific or multivalent nanobodies are high, in particular for their application in diagnosis and treatment of disease. Due to their small size, nanobodies show much better tumour penetration and can reach epitopes on target molecules that are sterically inaccessible to conventional antibodies. Compared to the production of (bispecific) antibodies in hybrid hybridomas, production costs for nanobodies are much more economically viable. Based on our results, we envisage development of a bispecific nanobody with inhibitory function towards Hh signalling that could be relevant for clinical applications.

Chapter 5

Characterisation of Ptc1 for structural studies

5.1 Abstract

In most cases, protein crystallography requires large quantities of protein and the use of protein engineering approaches to increase crystallisation probability. The goal of this study was to improve the expression and stability of the human transmembrane protein Patched-1 (Ptc1). The wildtype construct was poorly expressed and therefore not suitable for structural studies. In order to elevate expression levels and to reduce flexibility, we engineered a construct containing all domains of full-length Ptc1 but with truncations of the N- and C-termini and a deletion in the third intracellular loop (ICL3). Secondly, we substituted parts of the ICL3 with a thermostabilised cytochrome b562 RIL (bRIL) domain to expand the protein's intracellular surface and therefore increase the likelihood for crystal contact formation. Thirdly, we used nanobodies (Nbs), single-chain camelid antibody fragments, which are known for their ability to operate as crystallisation chaperones. We tested binding of the nanobodies to the transmembrane domain (TMD)-containing Ptc1 construct and identified Nb64 which bound with nanomolar affinity. Our protein engineering efforts yielded in a sample of high quality and quantity, with high thermostability up to 60°C, potentially suitable for crystallisation in vapour diffusion (VP) and lipidic cubic phase (LCP). Amphipol-exchanged and nanodisc-reconstituted protein was used to prepare specimen for electron microscopy (EM) studies. This study laid the ground work for structural studies of Ptc1 and possibly other membrane proteins with similar domain architecture.

5.2 Introduction

5.2.1 (Non)native environments for membrane proteins

Of all genomes that have been sequenced to date, almost one third of the coding regions represent sequences for membrane proteins [168]. Membrane proteins (MPs) are versatile; their domain architecture ranges from single- to multi-pass transmembrane domains (TMDs) with one or more extracellular (EC) or intracellular (IC) domains. They fulfill important functions in many cellular processes such as signal transduction, ion-exchange, small molecule transport, or cell-cell communication. Structure determination of MPs is of high interest as most MPs are potential drug targets. One of the main challenges of purifying MPs from their native environment for structural and functional studies is to retain the protein's native fold and stability. The most common method for MP extraction from cell membranes is the use of detergents. Amphiphilic molecules have the ability to enclose the MP and thus keep it soluble. The hydrophobic tail points towards the lipophilic parts of the protein whereas the hydrophilic head group faces the aqueous environment. A suitable detergent is able to solubilise the protein without denaturing it. Various detergents are readily available on the market, with nonionic and zwitterionic detergents being the most popular, and n-Dodecyl-β-D-Maltopyranoside (DDM) being the most successful in protein X-ray crystallography to date [191, 259]. Still, more detergent types for studying MPs are greatly anticipated, in particular mild detergents such as maltosides and glucosides [45]. Newer types include neopentyl glycol- and cholic acid-based detergents, which exhibit a sugar or steroid group, respectively [44, 297]. After solubilisation, the protein is often transferred into a buffer with detergent concentrations two to three times higher than the detergent's critical micelle concentration (CMC). The CMC is the degree of detergent concentration required to form detergent micelles; below the

CMC the detergent exists as monomers [253]. Free detergent micelles must always be present in a MP sample as detergent monomers rapidly exchange between the detergent micelles and the MP-detergent complexes.

Over preceding decades, detergent-solubilised proteins have been successfully used in protein crystallography and cryo-electron microscopy (cryo-EM); however, detergents can also impede structural studies. For instance, detergents which form larger micelles can complicate the integration of the MP-detergent complex into the crystal lattice and hence interfere with the crystal packing density and crystallisation ability of the protein [171]. While free detergent micelles are important for keeping the MPs soluble, they complicate the structure determination by cryo-EM. The number of free detergent micelles often exceeds the number of MP-detergent complexes. Their similar size makes them indistinguishable from one another and hence leads to problems during the particle alignment and classification process [116, 267]. Detergents can also influence the surface tension which affects the sample's behaviour during vitrification [116]. The following paragraphs will introduce alternative methods for studying MP for crystallographic and microscopic studies.

A new method for MP crystallisation was introduced by Landau *et al.* in 1996 [156]. By mixing the detergent-solubilised protein with lipids, lipidic cubic phases (LCPs) are formed. LCPs are bicontinuous lipid bilayers with an integrated three-dimensional aqueous channel system. The addition of precipitants slowly dehydrates the system, which initiates protein-protein interaction [191]. Crystallisation in lipidic sponge phase is another method for MP crystallisation and most suitable for larger MP complexes. Sponge phase formation is initiated by reagents that cause swelling of the lipidic mesophase [55, 282]. Besides crystallisation in LCP, MPs have been crystallised in bicelles, disc-shaped bilayers formed by phospholipids and deter-

gents [86]. While crystallisation in LCP requires special equipment such as a manual protein:lipid syringe mixer and a LCP crystallisation robot, crystallisation in bicelles can be set up using a conventional crystallisation robot.

As crystallisation of MPs remains a challenging process, cryo-EM has come into focus; recent technical advances has led to a few high-resolution structures of MPs being solved by cryo-EM. While a few MP structures have been solved in detergents, the detergent micelle can hinder structure determination (see above). Detergent alternatives that have been presented include Amphipols, nanodiscs, and styrene maleic acid (SMA)-lipid particles (SMALPs). The amphiphatic polymers (Amphipols) are able to substitute detergent molecules as they have higher affinity to hydrophobic surfaces compared to detergents. A8-35 is the most popular Amphipol in use and is comprised of an acidic polyacrylate backbone, with some side-chains substituted by octylamine (25%) or isopropylamine (40%) residues [268]. Molecular dynamics simulations predict that A8-35 forms small globular micelle-like particles with a lipophilic core and a hydrophilic surface [219]. Amphipol-exchanged MPs can tolerate various buffer conditions, which is advantageous for cryo-EM studies. However, some MPs in Amphipols are prone to aggregation and denaturation. As an alternative, MPs can be reconstituted in nanodiscs, which enables the study of MPs close to their native environment. The lipidic bilayers are formed by amphiphatic proteins that create a belt-like structure around the lipids thus shielding the hydrophobic layer from the aqueous environment. In the presence of MP-detergent complexes, the nanodiscs spontaneously self-assemble around the MP once detergent is removed. Nanodiscs create large structures, that facilitate their detection under the microscope; however, the formation of nanodisc stacks and reconstitution of one or more MPs per nanodisc increases the heterogeneity of the sample [66]. SMALPs consist of alternating styrene (hydrophobic) and maleic acid (hydrophilic) moieties which enclose the MP and a

small fraction of the native lipidic bilayer. Instead of detergents, SMA co-polymer can be used to directly extract the MP from the cell membranes, keeping the protein's natural environment (e.g. lipid composition) [214]. SMALPs are not only interesting for cryo-EM studies, but have also been proven successful in LCP crystallisation [33].

To date, a large repertoire of methods is available for studying MPs on a functional and structural level. Each MP is unique, and a one-method-fits-all approach is not applicable. However, with a growing number of MP structures available, structures of proteins with a similar domain architecture may help finding suitable conditions for the protein of interest. For most MPs, however, it will remain a trial and error approach until an appropriate system is found.

5.2.2 Structural studies of membrane proteins

As of today, structures of 817 unique MPs are known [3]. The first MP structure, the photosynthetic reaction centre of *Rhodopseudomonas viridis*, dates back to 1985 and was solved at 3 Å resolution [71]. The MP structures available include monotopic/bitopic proteins that exhibit a single TM helix, and polytopic proteins that have multiple TM segments (α-helical, β-barrel). As for X-ray crystallography, the nearly exponential growth of solved MP structures is mainly the result of crystallisation of MP-detergent complexes using the vapour diffusion (VD) method. In 2012, crystallisation in LCP gained much popularity, with a contribution of 25- 30% structures per year [131]. The most popular expression systems include bacteria, yeast, and insect cells, as mammalian expression systems normally require a specialised laboratory setup [131]. Most MP structures deposited are of eukaryotic or bacterial origin [131]. Those proteins constitute prominent targets for therapeutics and their structures are indispensable for structure-based drug design (SBDD). Of all available

MP structures, channels and transporters account for more than 50% of the structures [213]. The class of proteins that has seen the most striking increase between 2012 and 2015 is G protein-coupled receptors (GPCR). This might be attributed to the advances in LCP crystallisation, but also microfocus beamlines, which are more suitable for analysing small crystals. LCP crystals also exhibit less solvent content compared to traditional MP-detergent crystals from VD experiments. This makes them less prone to anisotropic X-ray scattering and radiation damage [138].

Two prominent examples of membrane protein structures that have been solved by either LCP or VD are Smoothened (Smo) and Niemann-Pick C1 (NPC1), respectively. Smo is a member of the Frizzled-class GPCR family. The protein fulfills an important function in the Hedgehog (Hh) signalling pathway and has been a target for drug discovery due to its implication in cancer development. E. Byrne and coworkers (2016) obtained the crystal structure of the Smo cysteine-rich domain (CRD) and TM domain (TMD) in complex with cholesterol and vismodegib at 3.2 Å and 3.3 Å resolution, respectively. Whereas cholesterol is proposed to be the natural ligand for Smo, vismodegib is a therapeutic drug for the treatment of basal cell carcinoma (BCC) [39]. These results gave insights into the drug binding site, which can be used for further SBDD. X-ray crystallography has also proven successful in tackling structure determination of larger systems. Niemann-Pick disease type C1 (NPC1) is a transporter with 13 TMDs, which is involved in the export of low-density lipoprotein (LDL)-derived cholesterol. Dysfunction leads to lysosomal lipid storage disease, also known as Niemann-Pick disease. X. Li *et al.* presented the NPC1 crystal structure to 3.3 Å resolution in 2016 [165].

While for many decades X-ray crystallography remained the gold standard for determining protein structures, cryo-EM is slowly catching up and changing the

field of MP structural biology. The method has proven extremely successful for structure determination of large protein complexes. Ryanodine receptor 1 (RyR1) is a 2.2 megadalton intracellular calcium release channel. In 2005, the structure was determined at 9.6 Å resolution [174]. Back then, the structure of RyR1 was the highest resolution structure of a membrane protein solved by single-particle cryo-EM. Ten years later, the resolution was improved to 4 Å due to detector and software development [290, 295]. Another breakthrough was reported in 2015 when X. Bai and co-workers published the cryo-EM structure of human γ-secretase at 3.4 Å resolution [14]. With a size of approximately 170 kDa, γ-secretase was the smallest protein solved by cryo-EM near atomic resolution.

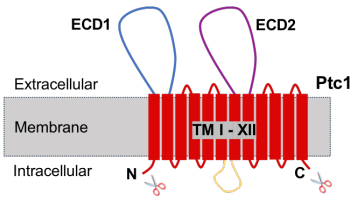
A few (10- 15%) MP structures have been solved by nuclear magnetic resonance (NMR) spectroscopy. Advances in the field have made it possible to study small MPs (30- 50 kDa) and their dynamics in solution or by solid-state NMR (ssNMR) [139]. One of the most striking examples is the backbone assignment of the potassium channel KcsA (68 kDa) in detergent micelles [56]. Recent studies have also shed light on the dynamics of GPCRs [60, 266]. Furthermore, newer approaches include NMR studies of MPs in nanodiscs. However, their large size complicates the structure determination by NMR spectroscopy and requires specific isotope labelling schemes [111]. Nevertheless, NMR spectroscopy remains the dominant method for studying protein dynamics. Combining structural information obtained by X-ray crystallography and cryo-EM with NMR spectroscopy, electron paramagnetic resonance (EPR) spectroscopy and molecular dynamics (MD) simulations can lead to a full description of MPs.

5.3 Results

5.3.1 Truncation of termini improves Ptc1 expression

Initial studies on wildtype human Ptc1 showed only minor expression and substantial aggregation of the protein. To improve the construct properties for structural studies, 30 different constructs with variable N- and C-termini were designed (Ptc1-1 - Ptc1-30, Appendix 1). These constructs can be divided in five subgroups; in each subgroup the N-terminal starting residue was kept constant, while the C-terminus was varied (Figure 5.1). All constructs were cloned into the pHLsec vector and fused C-terminally with a fluorescent tag (monoVenus, mV) for detection in fluorescent size-exclusion chromatography (FSEC), followed by a Rho1D4 antibody epitope tag (1D4-tag) for affinity purification [12]. For protein expression screening in small scale, HEK293T cells were transiently transfected with one of the 30 engineered Ptc1 constructs. All proteins were expressed in a 6-well plate format (2 wells/constructs) for 24 hours post-transfection and extracted in four different detergents (βOG, OGNG, DMNG, DDM; see Appendix C for the chemical structures of the detergents) to identify optimal conditions for protein solubilisation. Initial protein extraction was carried out at a final detergent concentration of 1.5% in 20 mM HEPES pH 7.5, 300 mM NaCl, 0.15% cholesteryl hemisuccinate (CHS). After Rho1D4 antibody affinity chromatography, proteins were analysed on FSEC using a Superose 6 3.2/300 column equilibrated in 10 mM HEPES pH 7.5, 300 mM NaCl, 0.007% DMNG 10:1 CHS. FSEC profiles revealed multiple peaks for all samples, corresponding to aggregate, a dimeric and a monomeric form of Ptc1, as well as cleaved mV-tag (Figure 5.2, 5.3, and 5.4). Overall, expression was improved for constructs with shorter N- and C-termini. The best results were achieved for constructs with domain boundaries Ala^{75} to Gly^{1179}, Ser^{1185}, or Ser^{1235} (Ptc1-19, Ptc1-20, Ptc1-21 [Group 4]), followed by constructs Ala^{58} or Ala^{75} to Ala^{1265}

(Ptc1-16 [Group 3] and Ptc1-22 [Group 4]). These five constructs displayed two main peaks, a dimeric and monomeric form of Ptc1. For constructs with C-termini ending at Ser^{1235} or after, an additional peak was observed. The peak was assigned to free mV-tag, suggesting an internal cleavage site in between Ser^{1185} and Ser^{1235}. Among the detergents tested, DMNG and DDM consistently represented the most suitable detergents for solubilisation of Ptc1 constructs. OGNG either failed to solubilise or yielded in mostly lower amounts of extracted protein compared to DMNG and DDM, in particular for constructs Ptc1-16, Ptc1-20 to Ptc1-22, and Ptc1-27. The detergent βOG could extract the same constructs from the membrane as DDM and DMNG, but displayed unusual behaviour; FSEC analysis resulted in less symmetric peaks and lower amounts of protein especially for constructs that yielded the best result for DMNG and DDM. Based on these results, it was decided to continue structural studies on Ptc1-20 (Ala^{75} to Ser^{1185}). This construct yielded the highest amounts of protein in screening and excluded the C-terminal cleavage site. An overlay of FSEC traces of Ptc1-6 (wildtype) and Ptc1-20 extracted in DMNG and DDM displays the improvement of protein yields after construct engineering (Figure 5.5, A-B). The detergents DMNG and DDM were selected for future experiments as they not only yielded the best results in screening, but DDM has also been shown to be most successful for protein crystallisation in detergent [212].



Group 1 (Ptc1-1 to Ptc1-6)	**Group 2** (Ptc1-7 to Ptc1-12)	**Group 3** (Ptc1-13 to Ptc1-18)	**Group 4** (Ptc1-19 to Ptc1-24)	**Group 5** (Ptc1-25 to Ptc1-30)
N-terminal residue: Met^{1}	Asp^{48}	Ala^{58}	Ala^{75}	Gly^{99}
C-terminal residue:				
1: Gly^{1179}	7: Gly^{1179}	13: Gly^{1179}	19: Gly^{1179}	25: Gly^{1179}
2: Ser^{1185}	8: Ser^{1185}	14: Ser^{1185}	20: Ser^{1185} (Ptc1-20)	26: Ser^{1185}
3: Ser^{1235}	9: Ser^{1235}	15: Ser^{1235}	21: Ser^{1235}	27: Ser^{1235}
4: Ala^{1265}	10: Ala^{1265}	16: Ala^{1265}	22: Ala^{1265}	28: Ala^{1265}
5: Val^{1381}	11: Val^{1381}	17: Val^{1381}	23: Val^{1381}	29: Val^{1381}
6: Asn^{1447} (Ptc1-6, WT)	12: Asn^{1447}	18: Asn^{1447}	24: Asn^{1447}	30: Asn^{1447}

Figure 5.1: **Initial screening of Ptc1: construct overview.** Five subgroups were designed; in each subgroup the N-terminal starting residue was kept constant, while the C-terminus was varied. The number in front of the C-terminal residue refers to the construct name, e.g. 1, Ptc1-1; 2, Ptc1-2; and so forth. WT indicates the wildtype construct.

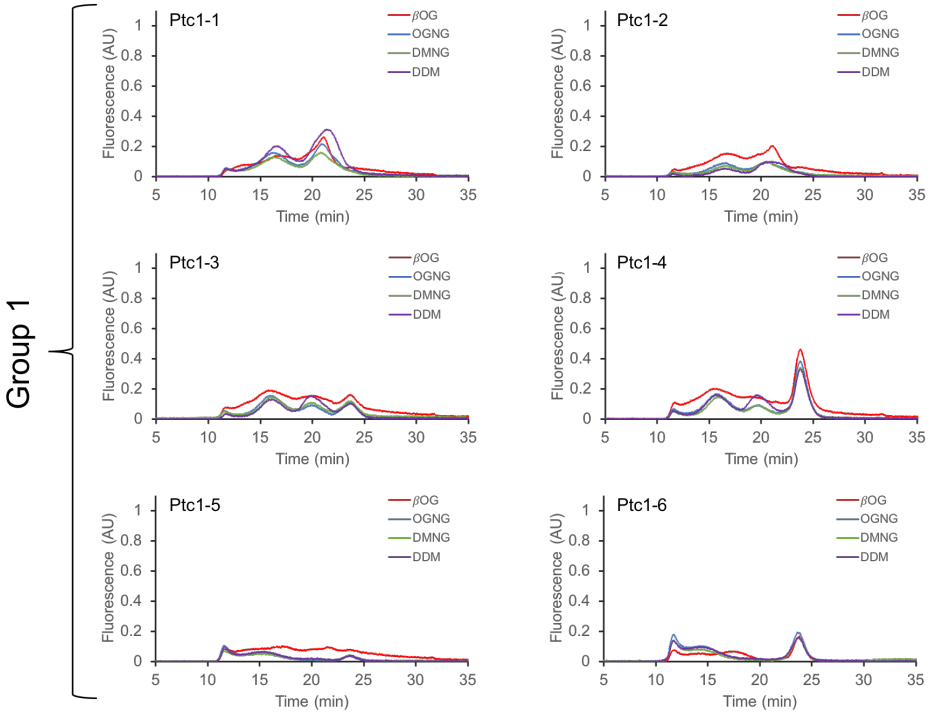


Figure 5.2: **Initial screening of Ptc1 constructs by FSEC: Group 1.** FSEC profiles of Ptc1-1 to Ptc1-6 constructs extracted in βOG, OGNG, DMNG, and DDM (Superose 6 3.2/300 column). All profiles displayed multiple peaks, indicating that Ptc1 was present as a monomer and dimer. Constructs showed minor to severe aggregation and in some cases cleavage of the mV-tag. For comparison, FSEC results were normalised to the monomer peak of Ptc1-20 (Figure 5.4) in DMNG.

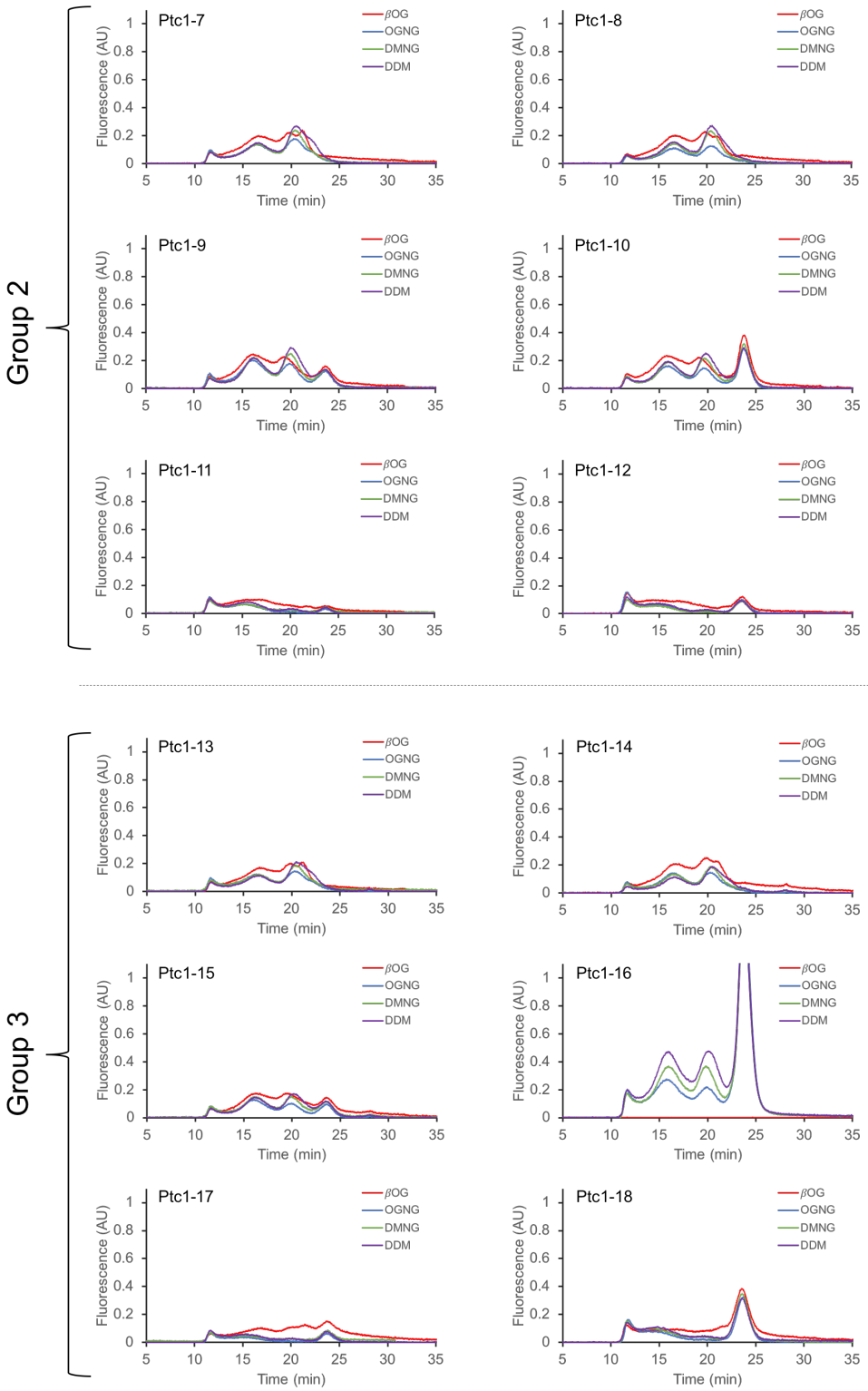


Figure 5.3: **Initial screening of Ptc1 constructs by FSEC: Group 2 and 3.** FSEC profiles of Ptc1-7 to Ptc1-18 constructs extracted in βOG, OGNG, DMNG, and DDM (Superose 6 3.2/300 column). All profiles displayed multiple peaks, indicating that Ptc1 was present as a monomer and dimer. Constructs showed minor to severe aggregation and in some cases cleavage of the mV-tag. For comparison, FSEC results were normalised to the monomer peak of Ptc1-20 (Figure 5.4) in DMNG.

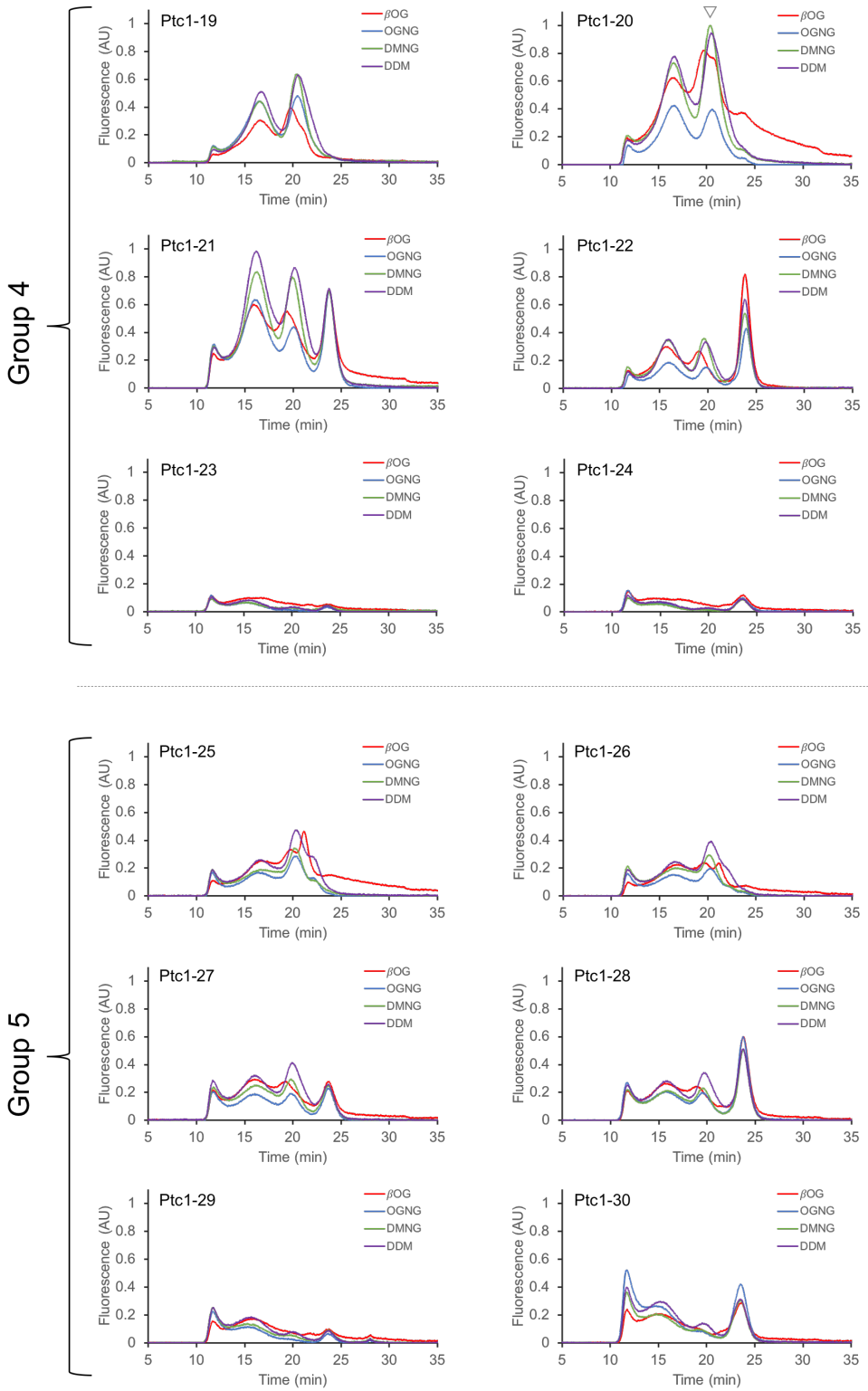


Figure 5.4: **Initial screening of Ptc1 constructs by FSEC: Group 4 and 5.** FSEC profiles of Ptc1-19 to Ptc1-30 constructs extracted in βOG, OGNG, DMNG, and DDM (Superose 6 3.2/300 column). All profiles displayed multiple peaks, indicating that Ptc1 was present as a monomer and dimer. Constructs showed minor to severe aggregation and in some cases cleavage of the mV-tag. For comparison, FSEC results were normalised to the monomer peak of Ptc1-20 (Group 4) in DMNG.

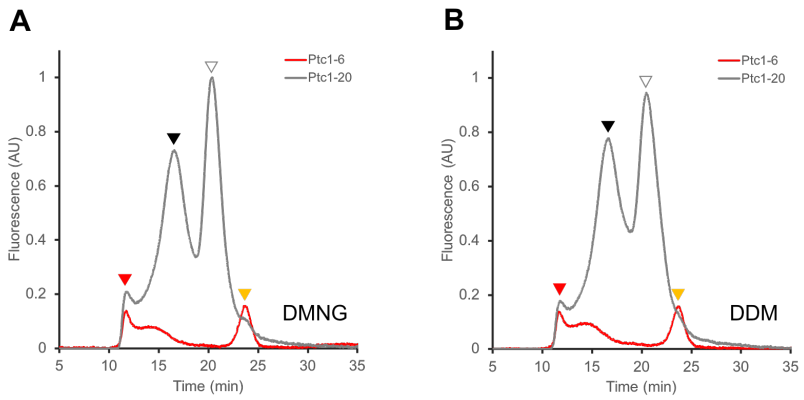


Figure 5.5: **Comparison of FSEC traces of Ptc1-6 and Ptc1-20 extracted in different detergents. A-B**, Results were normalised to the monomer peak of Ptc1-20 in DMNG for comparison. Ptc1-6 and Ptc1-20 extracted in DMNG (A) or DDM (B). Red arrowhead indicates the peak of aggregates, black and black outlined arrowhead the peak of dimeric and monomeric Ptc1, respectively. Yellow arrowhead indicates cleaved mV-tag. DMNG and DDM were equally efficient at extracting Ptc1-20. Ptc1-6 was poorly expressed, forming mainly aggregates. Free mV-tag indicated the existence of a cleavage site in Ptc1-6, which is not present in Ptc1-20.

Ptc1-20 was cloned into the pHLsec vector with C-terminal 1D4-tag (no mV-tag) and expressed in HEK293S GnTI$^-$ suspension cell culture for 60- 72 hours post-transfection. The protein was extracted from the cell membranes with 1.5% DMNG 10:1 CHS, and purified by Rho1D4 antibody affinity chromatography, followed by SEC on a Superose 6 10/30 column equilibrated in 20 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, 0.007% DMNG 10:1 CHS. While aggregation of the protein was less pronounced in the initial FSEC screening of the construct, preparative SEC resulted in a higher degree of aggregation, and only minor, rather broad peaks for Ptc1 monomer and dimer (Figure 5.6, A) with peak maxima at ∼15 mL and 11.5 mL retention volume, respectively. SDS-PAGE analysis of the peak fraction corresponding to the dimer revealed two bands, one at approximately 105 kDa and another one above 220 kDa (Figure 5.6, B). As the theoretical molecular weight of Ptc1-20 is 124.6 kDa [96], these bands could correspond to the monomeric and dimeric form of Ptc1. Indeed, it is not unusual that proteins appear on SDS-PAGE at a slightly lower

than predicted molecular weight. Furthermore, as SDS-PAGE was carried out under reducing conditions, intra- and intermolecular interactions, such as hydrogen bonds or salt bridges were not preserved, which could explain the presence of monomeric Ptc1 in the dimeric SEC fractions. The SEC peak corresponding to the monomer revealed three bands on SDS-PAGE. One at 105 kDa and two bands between 50- 60 kDa. The observation of two bands at approximately half the size of Ptc1-20 suggests the presence of a cleavage site in or near the third intracellular loop (ICL3) of Ptc1.

To increase the stability of Ptc1-20, co-purification with Shh was carried out. In addition, L-Arginine (L-Arg) was added to the solubilisation and purification buffer to increase the solubility and stability of the protein. The Shh-Ptc1 complex displayed a similar SEC profile as the Ptc1-20 apo-form (Figure 5.6, C). Again, significant aggregation occurred, however, more protein was obtained; nearly a three-fold increase in the absorption peak height of the monomer was achieved. The additional peak at ∼18.5 mL retention volume originates from Shh, which was added to the sample in molar excess. SDS-PAGE analysis confirmed that the Shh ligand co-purified with Ptc1-20 as it was detected at approximately 22 kDa in all SEC fractions (Figure 5.6, D). The peak fraction of the dimer showed a band for Ptc1 at 105 kDa and significant contamination with two low molecular weight species (50- 60 kDa), which were also present in the monomeric protein fractions. Western blot analysis using Rho1D4 antibody revealed that the contamination was due to degradation of the protein and that both fragments were C-terminal fragments of Ptc1-20 (Figure 5.6, E). This suggests that not one, but two cleavage sites are present in or close-by the Ptc1-20 ICL3 region. Analysis of the aggregation peak by western blot revealed two additional bands at higher molecular weight. The band at approximately 200 kDa could correspond to the dimer, whereas the band at 150 kDa could be a dimer formed by monomeric Ptc1 and one of the degradation products.

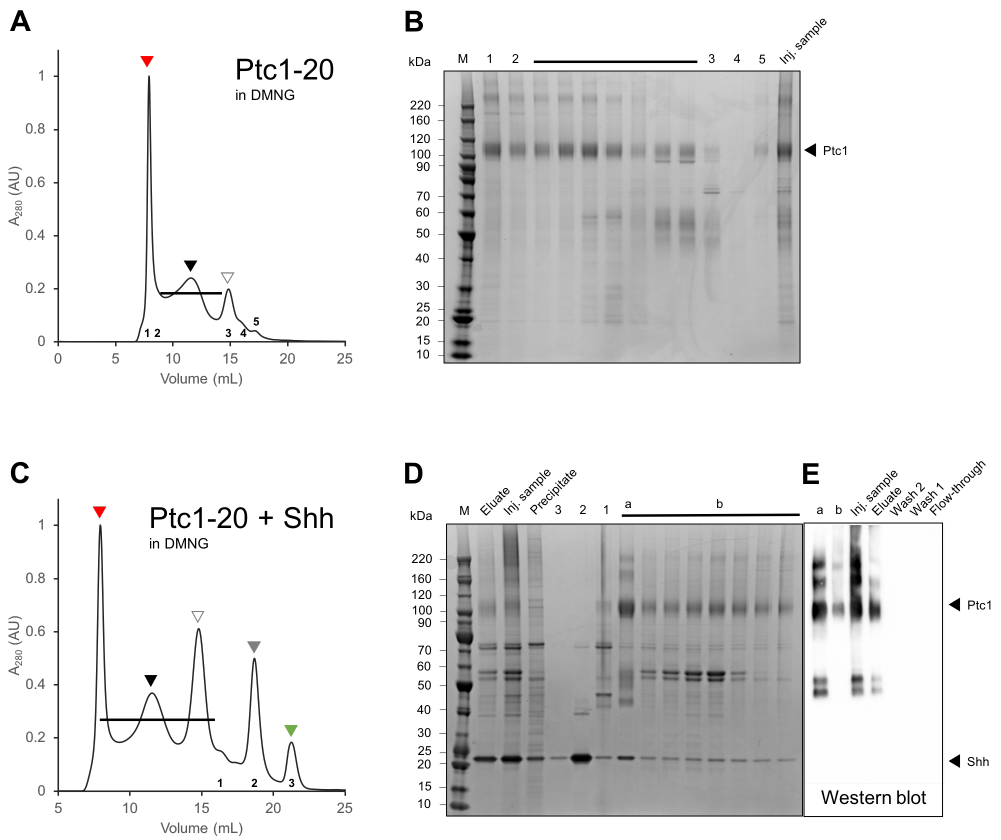


Figure 5.6: **Purification and analysis of Ptc1-20 in DMNG. A-B**, SEC profile (A) and corresponding SDS-PAGE analysis (B) of Ptc1-20 extracted in DMNG (Superose 6 10/300 column). **A**, Red arrowhead indicates the peak of aggregates, black and black outlined arrowhead the peak of dimeric and monomeric Ptc1, respectively. SEC fractions analysed on SDS-PAGE are marked by numbers and a black line. **B**, SDS-PAGE analysis confirmed the presence of Ptc1-20 and degradation products. Lanes from left to right: BenchMark^{TM} ladder (M), aggregated protein species (1, 2), monomeric Ptc1 (3), and degradation products (4, 5). SEC fractions across the dimeric peak fraction are indicated by a black line. **C-E**, SEC profile (C) and corresponding SDS-PAGE (B) and western blot (E) of the Ptc1-20-Shh complex in DMNG. **C**, Red arrowhead indicates the peak of aggregates, black and black outlined arrowhead the peak of dimeric and monomeric Ptc1, respectively. Grey arrowhead indicates excess Shh. **D**, SDS-PAGE analysis confirmed the presence of the Ptc1-20-Shh complex and degradation products. Lanes from left to right: BenchMark^{TM} ladder (M), eluate of affinity purification, sample subjected to SEC (Inj. sample), precipitate removed by centrifugation before injection, degradation products (3), Shh (2), degradation products (1). SEC fractions across the dimeric peak fraction are indicated by a black line. **E**, From left to right: SEC fractions a and b as analysed on SDS-PAGE, purification steps of affinity purification.

Although the protein yield of both samples was rather low due to significant protein aggregation, a few crystal plates using the VD method were set up. The dimeric species of Ptc1-20 alone or in complex with Shh was concentrated to 1.9 mg mL^{-1} and 1.7 mg mL^{-1}, respectively. All samples were subjected to treatment with Endoglycosidase F1 (EndoF1) to reduce N-linked glycosylation sites. For Ptc1-20, most conditions showed precipitation in the drops, which however, did not develop into protein crystals. For the complex, needle-like crystals were obtained in one condition, which were later proven to be salt crystals (Figure 5.7). Despite significant protein precipitation during concentration, monomeric Shh-Ptc1 could be concentrated to 1.2 mg mL^{-1}; however, no crystals were observed. Due to recent advances in cryo-electron microscopy (cryo-EM) and the suitable molecular size of Ptc1-20, the sample was further investigated under the microscope. The particles had a tendency to adsorb to carbon, and protein was only observed near the edge of the grid hole. This phenomenon is often observed for samples in detergent.

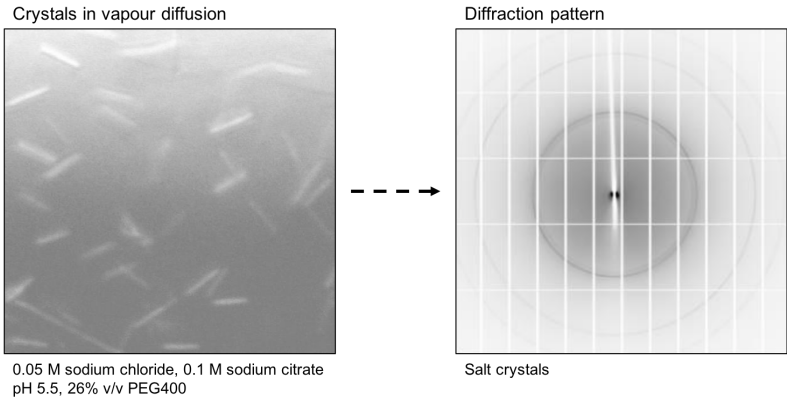


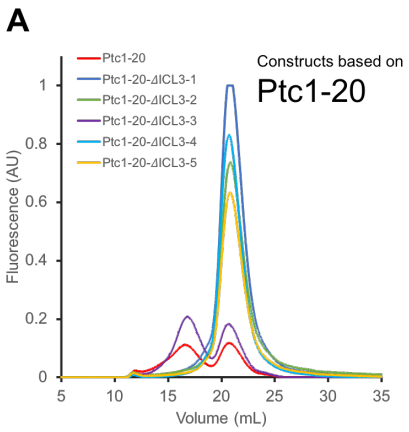
Figure 5.7: **Crystallisation of Ptc1-20 in vapour diffusion.** Needle-like crystals were obtained in one condition. X-ray diffraction experiments confirmed that the crystals originated from salt.

5.3.2 Removal of the PPXY motif stabilises Ptc1

Even though Ptc1-20 could be expressed and purified in large scale, the protein yield was unsatisfactory, and unsuitable for crystallisation in lipidic cubic phase (LCP). To improve protein yield and stability, we went a step back to explore further protein engineering approaches. Two proline-rich, endocytic sorting signals have been identified in the Ptc1 sequence; one in the third intracellular loop (ICL3) and one in the C-terminal tail. The PPXY motif in the ICL3 (Pro^{642} to Tyr^{645} in human Ptc1) is recognised by Smurf1 and Smurf2 E3 ubiquitin ligases, which target Ptc1 for degradation via the lysosomal pathway [294]. The two ligases ubiquitinate Ptc1 upon Hh signalling activation, and are responsible for clearance of Ptc1 from the primary cilium. The C-terminal PPXY motif of human Ptc1 is located between Pro^{1313} and Tyr^{1316}, and is also predicted to be targeted by HECT ubiquitin ligases which partly initiate the fast turnover of Ptc1 [119, 142, 172]. Interestingly, Ptc2, a Ptc1 homolog, does not exhibit the C-terminal PPXY motif and tends to be relatively stable [142]. As the Ptc1-20 construct was already missing the C-terminal PPXY motif (ending at Ser^{1185}), we focused on screening constructs lacking the PPXY motif in the ICL3.

Five Ptc1-20 constructs with deletion mutations in the ICL3 were designed and synthesised using overlapping PCR (Figure 5.8, B). Wildtype residues were substituted by an eight residue Gly/Ser (G/S)-rich linker (GSGGSSGG). Constructs were cloned into the pHLsec expression vector with C-terminal mV-tag (for detection in FSEC), followed by a 1D4-tag (for affinity purification). Small-scale expression tests in HEK293T cells were carried out as described earlier. The protein was extracted from the cell membrane using 1.5% DMNG 10:1 CHS in 20 mM HEPES pH 7.5, 300 mM NaCl and purified via Rho1D4 antibody affinity chromatography before being subjected to FSEC using a Superose 6 3.2/300 column equilibrated in 10 mM HEPES

pH 7.5, 300 mM NaCl, 0.007% DMNG 10:1 CHS. An in-line fluorescence detection system was used to detect the mV-tagged proteins.



B

Construct name	Residues deleted in Ptc1-20		
Ptc1-20ΔICL3-1	Phe^{611}	to	Val^{692}
Ptc1-20ΔICL3-2	Phe^{611}	to	Asp^{717}
Ptc1-20ΔICL3-3	Tyr^{630}	to	Ser^{649}
Ptc1-20ΔICL3-4	Tyr^{630}	to	Val^{692}
Ptc1-20ΔICL3-5	Tyr^{630}	to	Asp^{717}

Figure 5.8: **Screening of Ptc1 ICL3 deletion mutants. A**, FSEC analysis of Ptc1-20 based constructs, each with one of five deletion mutations in the ICL3. The control and deletion mutant missing residues Tyr^{630} to Ser^{649} (Ptc1-20ΔICL3-3) displayed two peaks in FSEC corresponding to dimeric and monomeric Ptc1. All other constructs displayed one monodisperse peak at retention times typical for Ptc1 monomer. Results indicated that regions of ICL3 participate in dimer formation. **B**, Overview of constructs tested in FSEC. Deleted residues are indicated.

Depending on the construct, FSEC profiles either revealed one or two main protein peaks (Figure 5.8, A). The control construct Ptc1-20 displayed two peaks, one corresponding to Ptc1 dimer and one to Ptc1 monomer. The shortest deletion in the ICL3 excluded a total of 20 amino acid residues, Tyr^{630} to Ser^{649} (Ptc1-20ΔICL3-3), and behaved similarly to the Ptc1-20 construct; however, protein yield could be increased approximately two-fold. All other constructs, with deletions starting at Tyr^{630} (Ptc1-20ΔICL3-4 [Tyr^{630} to Val^{692}], Ptc1-20ΔICL3-5 [Tyr^{630} to Asp^{717}]) or Phe^{611} (Ptc1-20ΔICL3-1 [Phe^{611} to Val^{692}], Ptc1-20ΔICL3-2 [Phe^{611} to Asp^{717}]) displayed a single, monodisperse peak at retention volumes characteristic for monomeric Ptc1. Ptc1-20ΔICL3-1 increased protein yield ∼10-fold, followed by Ptc1-20ΔICL3-4 (∼8-fold), Ptc1-20ΔICL3-2 (∼7-fold), and Ptc1-20ΔICL3-5 (∼6-fold). These results

suggest that the region between Ser^{649} and Val^{692} harbours a motif which is important for dimerisation of Ptc1. Furthermore, expression was decreased when the region between Val^{692} and Asp^{717} was deleted. This region may have a stabilising effect on Ptc1 and might explain the substantial increase in protein yields for the Ptc1-20ΔICL3-1 and Ptc1-20ΔICL3-4 construct.

Our collaborators R. Rohatgi and co-workers (Stanford University, USA) examined the Ptc1 subcellular localisation after transient transfection of NIH/3T3 $Ptc1^{-/-}$ fibroblasts with Ptc1-20ΔICL3 mutants (Figure 5.9). Expression of mV-tagged Ptc1 was then monitored by fluorescence microscopy. Most importantly, and as expected, expression was observed for all constructs. In addition, two interesting observations were made. Firstly, constructs lacking the region Phe^{611} to Tyr^{630} (Ptc1-20ΔICL3-1 and Ptc1-20ΔICL3-2) fail to locate in the shaft of the primary cilium. Constructs Ptc1-20ΔICL3-3, Ptc1-20ΔICL3-4, and Ptc1-20ΔICL3-5, which include this region, were detected in the cilium. This indicates that the region between Phe^{611} to Tyr^{630} possesses a motif that is important for cilia localisation. Secondly, constructs missing residues Val^{692} to Asp^{717} showed increased membrane localisation, which suggests that this region could potentially carry a motif for Ptc1 internalisation. Untransfected cells displayed Smo in cilia, indicating that the pathway was activated in the absence of Ptc1.

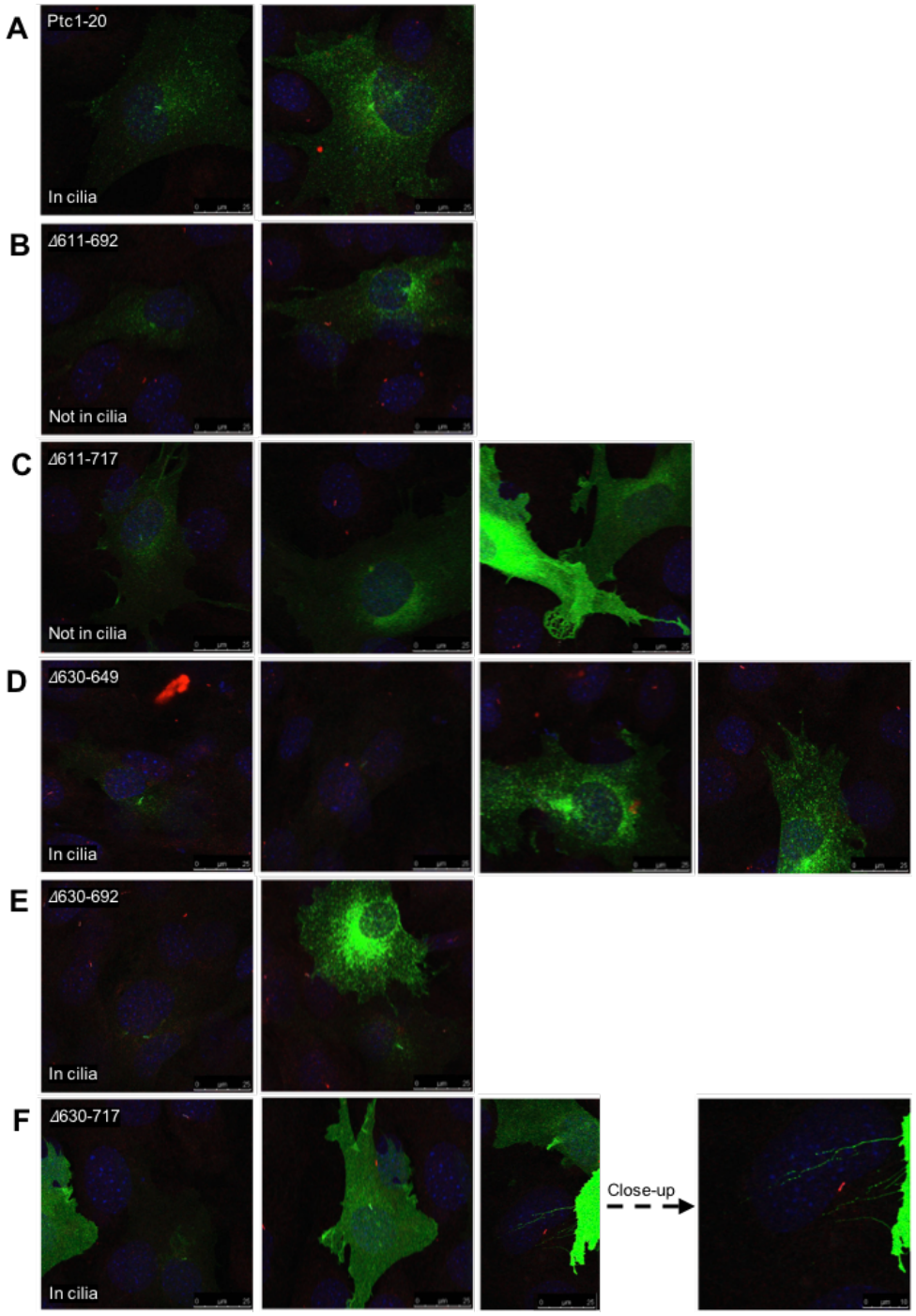


Figure 5.9: **Ptc1 mutants and their localisation in the primary cilium. A-F**, NIH/3T3 cells transfected with fluorescently-tagged Ptc1-20 mutants (see Figure 5.8, B for details on constructs). In the absence of Shh, Ptc1 was located in the primary cilium when region Phe^{611} to Ala^{629} was present (A, D-F), but failed to locate in cilia when this region was missing (B, C). Constructs lacking residues Val^{692} to Asp^{717} (C, F) showed higher membrane localisation, possibly lacking an internalisation motif located in this region. Untransfected cells showed Smo localisation in cilia as expected in a $Ptc1^{-/-}$ background. Ptc1 was detected by mV fluorescence (green), Smo by immunofluorescence (red) using anti-Smo antibody (Josman, LLC). Anti-acetylated tubulin antibody (Sigma) was used as a cilia marker (cyan). Fluorescence microscopy experiments were performed in collaboration with R. Rohatgi and co-workers.

The fact that Ptc1 mutants which still carry the region Phe^{611} to Tyr^{630} were found in primary cilia further indicates that these constructs are biologically active. However, from these studies it is difficult to say whether Ptc1 and Smo localisation in cilia is inversely correlated. Addition of Shh to the cells and monitoring the response by fluorescence microscopy, could shed light on whether Ptc1 mutants are responsive to Shh and initiate localisation of Smo into and Ptc1 out of the primary cilium. Another way to determine whether the Ptc1-20ΔICL3 mutants are biologically active, is to carry out Gli1 mRNA transcription reporter assays on cells expressing the mutants.

Based on these results, we decided to produce Ptc1-20ΔICL3-5 on a large scale for structural studies. Construct screening indicated that the protein expresses well and elutes as a monodisperse species, with very little aggregation. Furthermore, microscopy studies confirmed high expression levels and membrane localisation of this particular mutant. Besides, flexible regions can hamper crystallisation of proteins and loop regions are renowned to be unstructured. The Ptc1 ICL3 region is predicted to span residues Asp^{599} to Lys^{748} and Ptc1-20ΔICL3-5 lacks most of this region (Tyr^{630} to Asp^{717}). For large scale protein expressions, Ptc1-20ΔICL3-5 was cloned into the pHLsec expression vector with C-terminal 1D4-tag (no mV-tag). Expression was carried out in HEK293S GnTI^{-} suspension tissue culture as described in Chapter 5.3.1. Following expression, proteins were extracted from the cell membranes using 1.5% DMNG and 0.15% CHS in 10 mM HEPES pH7.5, 300 mM NaCl, 5% glycerol, and 100 mM L-Arg. Affinity chromatography was carried out in the presence of Shh, after which the complex was subjected to SEC on a Superose 6 10/30 column equilibrated in 10 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, 50 mM L-Arg, and 0.007% DMNG 10:1 CHS. In contrast to Ptc1-20, protein aggregation during purification was greatly reduced, and similar to small-scale expression trials, Ptc1-20ΔICL3-5 appeared as a single monodisperse peak in SEC; however, a small shoulder was observed, potentially

resulting from Ptc1 dimer formation (Figure 5.10, A). As the protein was purified in complex with Shh, the SEC trace displayed an additional peak at ∼18 mL retention volume resulting from excess molar quantities of Shh. SDS-PAGE analysis revealed that Shh formed a complex with the Ptc1 mutant as the proteins co-eluted on SEC (Figure 5.10, B). In addition, the sample was of high purity, and no degradation products were observed. The aggregation peak could not be analysed as the protein did not enter the gel. Proteins appeared at a higher molecular weight than estimated (102.2 kDa for Ptc1-20ΔICL3-5, 20.5 kDa for Shh [96]), which might partially be due the inaccuracy of the pre-stained protein molecular weight marker [169]. Western blot analysis with Rho1D4 antibody confirmed that the band at approximately 120 kDa belongs to Ptc1 (Figure 5.10, C). Moreover, western blot analysis of the purification steps confirmed the efficiency of the affinity purification. No protein was detected in the flow-through and wash steps. The concentrated sample, which was used for crystallisation trials after EndoF1 treatment, was free of degradation products, but revealed minor amounts of Ptc1 dimer in the sample. The fact that no degradation products were observed supports the hypothesis of a cleavage site being present in the Ptc1 ICL3. The protein was concentrated to 4.5 mg mL^{-1} and subjected to EndoF1 treatment for 1 hour at room temperature. The deglycosylated protein sample was used to set up crystallisation in VD using 96-well 3-drop Swissci plates. Even though the linker deletion reduced sample heterogeneity, no crystals have been observed to date.

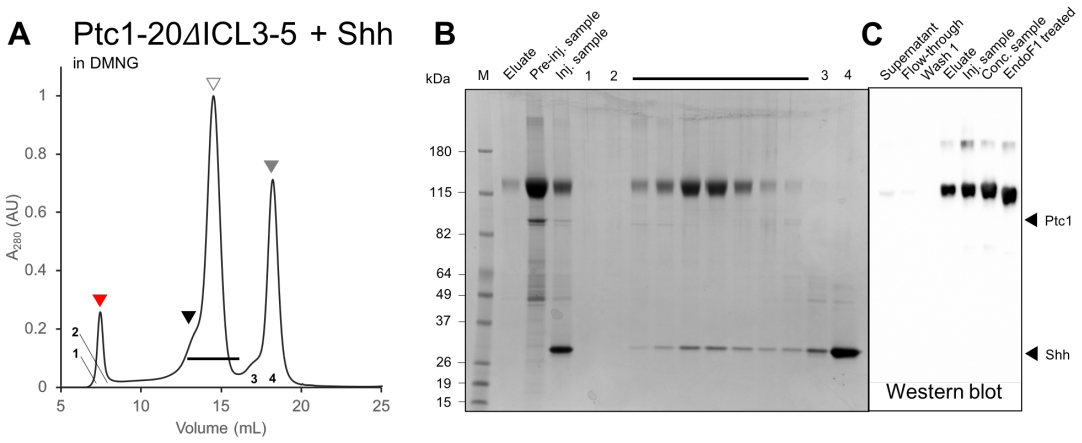


Figure 5.10: **Purification and analysis of Ptc1-20ΔICL3-5 in DMNG. A-C**, SEC profile (A) with corresponding SDS-PAGE (B) and western blot (C) of Ptc1 ICL3 deletion mutant Ptc1-20ΔICL3-5 in complex with Shh and purified in DMNG. **A**, Red arrowhead indicates the peak of aggregated protein species, black outlined arrowhead the peak of monomeric Ptc1. The minor shoulder on the left side of the peak originated from Ptc1 dimer. Grey arrowhead indicates excess Shh. **B**, Lanes from left to right: BenchMark^{TM} ladder (M), eluate of affinity purification, sample subjected to SEC (Inj. sample), aggregated protein species (1, 2), impurities (3), Shh (4). SEC fractions across the Ptc1 peak fraction are indicated by a black line. **C**, Purification steps analysed on western blot. Lanes 1-4: affinity purification, lane 5: sample subjected to SEC, lanes 6-7: sample for crystallisation before and after EndoF1 treatment.

Most success has been reported using DDM for crystallisation of membrane proteins. An analysis of all membrane protein structures listed in the "Membrane Proteins of Known Structures" database (as of December 2016) revealed that more than 36% of the proteins were crystallised in DDM [259]. Therefore, we decided to establish an extraction and purification protocol for Ptc1 using this detergent. From our small-scale screening trials we already knew that DDM was equally efficient at extracting Ptc1 as DMNG. The extraction of Ptc1-20ΔICL3-5 from cell membranes in DDM was carried out in a similar way to DMNG. Before crystallisation trials were set up, the protein was exchanged into 10 mM HEPES pH7.5, 300 mM NaCl, 10% glycerol, and 0.03% DDM 10:1 CHS via SEC. The SEC profiles displayed very little aggregation and a monomer peak with a minor shoulder on the left side (Figure 5.11, A). The shoulder most likely represented Ptc1 dimer as SDS-PAGE and western blot

analysis revealed a band at approximately double the size of Ptc1 monomer (Figure 5.11, B). Furthermore, it confirmed the presence of monomeric Ptc1 and high purity of the sample. The protein was concentrated to 5 mg mL^{-1} and crystallisation in VD was set up using 96-well 3-drop Swissci plates. Different membrane protein screens (Molecular Dimensions and Hampton Research) and drop ratios were explored. Crystals observed were either salt or detergent crystals.

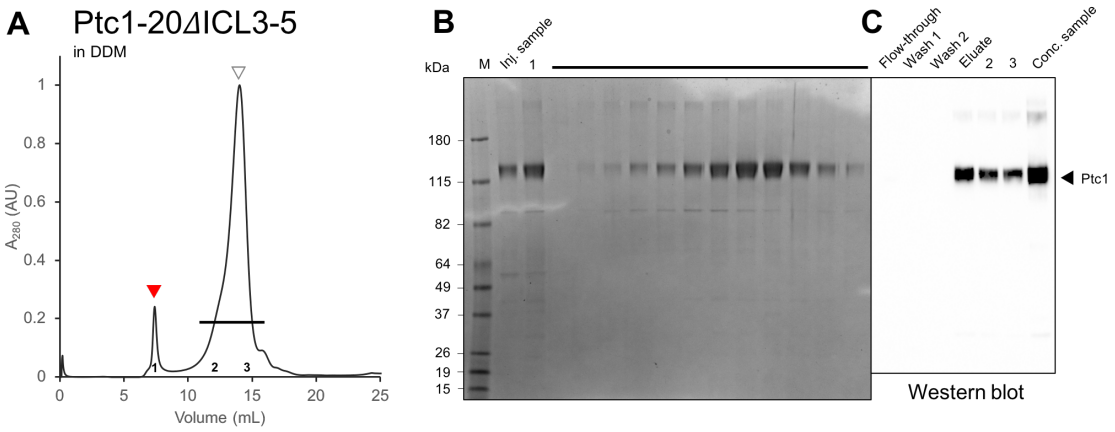


Figure 5.11: **Purification and analysis of Ptc1-20ΔICL3-5 in DDM. A-C**, SEC profile (A) with corresponding SDS-PAGE (B) and western blot (C) of Ptc1 ICL3 deletion mutant Ptc1-20ΔICL3-5 in DDM. **A**, Red arrowhead indicates the peak of aggregated protein species, black outlined arrowhead the peak of monomeric Ptc1. The minor shoulder on the left side of the peak originated from Ptc1 dimer. **B**, Lanes from left to right: BenchMark^{TM} ladder (M), sample subjected to SEC (Inj. sample), aggregated protein species (1). SEC fractions across the Ptc1 peak fraction are indicated by a black line. **C**, Purification steps analysed on western blot. Lanes 1-4: affinity purification, lanes 5-6: SEC fractions indicated in A (2, 3), lane 7: sample for crystallisation before EndoF1 treatment.

5.3.3 Nb64 binds to Ptc1 with high affinity

Crystallisation of membrane proteins is not trivial. Difficulties not only arise from their intrinsic properties, such as extensive hydrophobic surfaces and flexible regions, but also from the purification in detergent, which is known to destabilise the protein [42]. A frequently applied method to overcome these difficulties is to co-crystallise

membrane proteins with antibodies, antibody fragments or nanobodies. Binding of these so-called crystallisation chaperones can stabilise the proteins without perturbing their native conformation. Moreover, they introduce additional crystal contacts by expanding the existing protein surface and by reducing the hydrophobicity of the membrane protein.

In our study we aimed to explore the effects of nanobodies on Ptc1 crystallisation. Nanobodies were previously raised against the Ptc1 extracellular domain fusion construct (Ptc1 ECD1-ECD2) and tested by us for binding using bio-layer interferometry (BLI) (Chapter 3). In the following, we characterised these nanobodies for binding to Ptc1-20ΔICL3-5. To allow immobilisation of the protein to streptavidin-coated biosensors, biotinylated protein was required. The crystallisation construct already carried a C-terminal AviTag^{TM} for biotinylation and could be used for co-transfection of HEK293S GnTI^{-} cells with a soluble version of BirA biotin ligase. Biotinylated Ptc1-20ΔICL3-5 was extracted from the cell membranes using 1.5% DDM 10:1 CHS. After Rho1D4 antibody affinity chromatography, the protein was SEC purified using a Superose 6 10/30 column equilibrated in 10 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, and 0.03% DDM 10:1 CHS. As detergent might have an effect on nanobodies and their binding affinity, biotinylated Ptc1-20ΔICL3-5 was exchanged into Amphipols and 10 mM HEPES pH 7.5, 150 mM NaCl. The protein was concentrated to 0.1 mg mL^{-1}, and 0.005% Tween-20 was added to the sample before immobilisation to reduce unspecific binding to the biosensors. Immobilisation of Ptc1-20ΔICL3-5 occurred rapidly and with similar response values near 2.5 nm (Figure 5.12, A). In a BLI rapid screen assay, nanobodies were then presented to Ptc1-20ΔICL3-5 on biosensors at 0.1 μM concentration for 300 sec. Out of 27 nanobodies, 10 (Nb64, 65, 66, 67, 68, 790, 791, 793, 802, 803) were identified as Ptc1 binding partners (Figure 5.12, B-C). These nanobodies were subsequently characterised in affinity

measurements for kinetic rate constant (K_D) determination. Briefly, seven concentrations of nanobody, resulting from two-fold dilutions were tested in parallel and K_D values were calculated from the observed on- and off-rates. Data were analysed using ForteBio's data analysis software (Version 9.0) and a 2:1 heterogeneous binding model was applied to fit the data. The fact that the curves had a slightly biphasic character indicates that initial nanobody binding occurred rapidly, and remaining Ptc1 sites were slowly occupied afterwards. This phenomenon could have arisen from Ptc1 dimer being present in the sample. In this case, it might be more difficult for the nanobodies to access the binding interface of the dimer. To calculate affinities, response values near equilibrium were plotted against nanobody concentrations. Most nanobodies showed K_D values below 10 nM, all others had binding affinities in the range of 72- 360 nM (Figure 5.13, C). Nb64 was chosen as a representative example; the real-time binding chart and steady state analysis is shown in Figure 5.13, A-B. For comparison, binding of Nb64 was tested against detergent-solubilised protein. Interestingly, binding in Amphipols was almost 40 times stronger than in detergent, indicating that detergent interferes with binding of the nanobody (Figure 5.13, B-C).

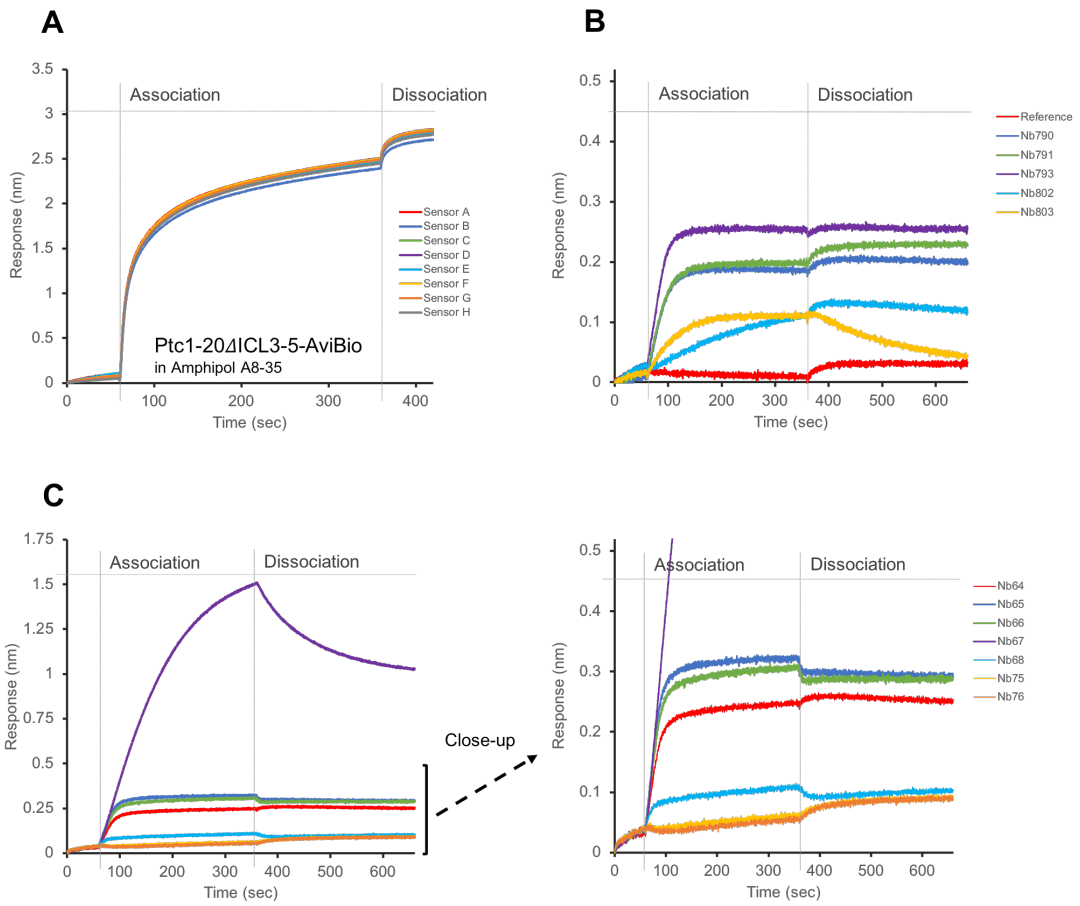
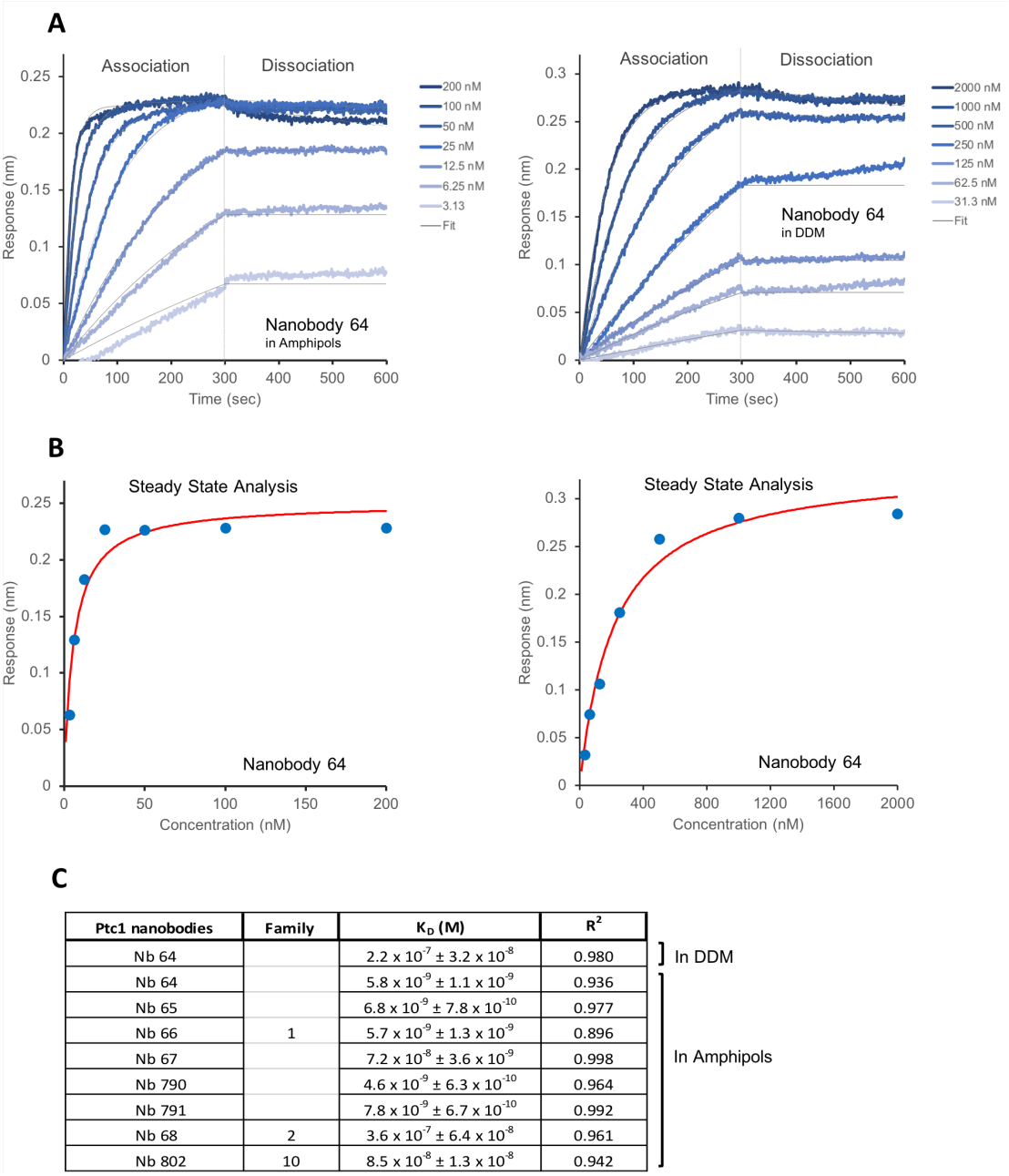


Figure 5.12: **Screening of nanobodies against Ptc1-20ΔICL3-5. A**, Immobilisation of biotinylated Ptc1-20ΔICL3-5 to optical biosensors for BLI analysis. **B-C**, Sensorgrams of BLI rapid screens to identify Ptc1-specific nanobodies. Most nanobodies showed response values between 0.1-0.3 nm. The low response of Nb75 and Nb76 indicated that those nanobodies are not interacting with Ptc1.



Ptc1 nanobodies	Family	K_D (M)	R^2	
Nb 64		$2.2 \times 10^{-7} \pm 3.2 \times 10^{-8}$	0.980	In DDM
Nb 64		$5.8 \times 10^{-9} \pm 1.1 \times 10^{-9}$	0.936	In Amphipols
Nb 65		$6.8 \times 10^{-9} \pm 7.8 \times 10^{-10}$	0.977	
Nb 66	1	$5.7 \times 10^{-9} \pm 1.3 \times 10^{-9}$	0.896	
Nb 67		$7.2 \times 10^{-8} \pm 3.6 \times 10^{-9}$	0.998	
Nb 790		$4.6 \times 10^{-9} \pm 6.3 \times 10^{-10}$	0.964	
Nb 791		$7.8 \times 10^{-9} \pm 6.7 \times 10^{-10}$	0.992	
Nb 68	2	$3.6 \times 10^{-7} \pm 6.4 \times 10^{-8}$	0.961	
Nb 802	10	$8.5 \times 10^{-8} \pm 1.3 \times 10^{-8}$	0.942	

Figure 5.13: **Kinetic analysis of nanobodies targeting Ptc1-20ΔICL3-5. A-B**, Characterisation of Ptc1-Nb64 interaction. **A**, Real-time sensorgrams for association and dissociation of Nb64 recorded with Ptc1 in Amphipols (left) and DDM (right). Statistical fitting of data is shown in grey. Association and dissociation steps are divided by a black dotted line. **B**, Steady state analysis to determine the equilibrium dissociation constant for binding of Nb64 to Ptc1 in Amphipols (left) and DDM (right). A K_D of $5.8 \times 10^{-9} \pm 1.1 \times 10^{-9}$ M and $2.2 \times 10^{-7} \pm 3.2 \times 10^{-8}$ M was determined for Ptc1 in Amphipols and DDM, respectively. Lower binding affinity of Nb64 to Ptc1 in DDM indicated that detergent interferes with binding. **C**, Table including dissociation constants of Ptc1-Nb complexes and corresponding coefficient of determination R^2.

Based on the BLI results of Ptc1-20ΔICL3-5 and previous studies on the Ptc1 extracellular domain, we chose to co-purify Ptc1 with Nb64. The protein was expressed and purified as described for Ptc1-20ΔICL3-5, but was prepared to co-elute from Rho1D4 antibody-coupled beads with molar excess of nanobody. The eluate was concentrated and subjected to SEC. The SEC profile revealed aggregation of the protein (∼7.5 mL retention volume), followed by two main protein peaks (Figure 5.14, A). The first peak (peak maximum at ∼12 mL retention volume) appeared to be slightly asymmetric, which suggested that Ptc1 dimer and monomer were not resolved. SDS-PAGE analysis later revealed that the peak contained the Ptc1-Nb64 complex, possibly formed as a 1:1, 2:1 or 2:2 complex (Figure 5.14, B). The fact that a band for dimeric Ptc1 was observed under reducing conditions indicated that the interaction between subunits was relatively strong. The second peak at ∼18.5 mL retention volume originated from Nb64. Western blot analysis revealed three main bands, one at approximately 120 kDa, one at about double the size, and one at a higher molecular weight (Figure 5.14, B). Although multi-angle light scattering (MALS) experiments (Chapter 5.3.4) only confirmed a dimeric and monomeric species in the sample, a trimeric form of Ptc1 might exist. Resistance-nodulation-division (RND) family transporters, which are bacterial homologues of Ptc1 have been reported to form homotrimers [195]. Less intense bands at lower molecular weights indicated degradation of the protein. As the protein was 1D4-tagged on its C-terminus, western blot analysis revealed only C-terminal fragments.

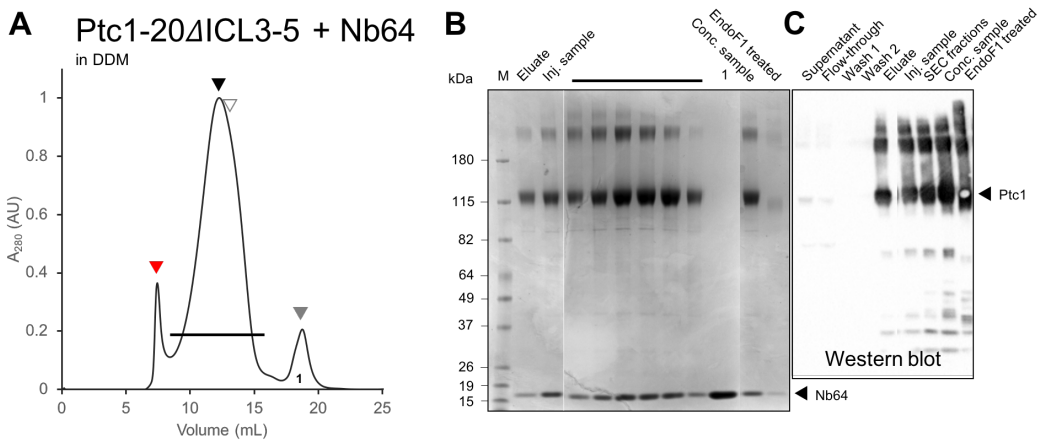


Figure 5.14: **Purification and analysis of Ptc1-20ΔICL3-5 in complex with Nb64. A-C**, SEC profile (A) with corresponding SDS-PAGE (B) and western blot (C) of Ptc1 ICL3 deletion mutant Ptc1-20ΔICL3-5 in complex with Nb64 in DDM. **A**, Red arrowhead indicates the peak of aggregated protein species, black and black outlined arrowhead the peak of dimeric and monomeric Ptc1, respectively. Grey arrowhead indicates the peak of Nb64. **B**, Lanes from left to right: BenchMark^{TM} ladder (M), eluate of affinity purification, sample subjected to SEC (Inj. sample), Nb64 (1), sample for crystallisation before and after EndoF1 treatment. SEC fractions across the Ptc1 peak fraction are indicated by a black line. Ptc1 co-eluted with Nb64. **C**, Purification steps analysed on western blot. Lanes 1-5: affinity purification, lane 6: sample subjected to SEC, lanes 7-8: sample for crystallisation before and after EndoF1 treatment. The sample consisted of Ptc1 dimer and monomer. Degradation products of lower molecular weight were observed.

In the presence of nanobody, protein yield could be improved substantially. One litre of cell culture yielded approximately 0.35 mg of protein; about six times more compared to before. The protein was concentrated to ∼54 mg mL^{-1}, followed by treatment with EndoF1. The protein was reconstituted in a mixture of 90% monoolein and 10% cholesterol in a 3:2 (w/v) lipid to protein ratio (see Appendix C for the chemical structures of the lipids). Despite the high protein concentration, crystallisation in LCP only yielded salt or detergent crystals. VD crystallisation of the Ptc1-Nb64 complex in DDM at a 1:1 protein:reservoir drop ratio, yielded, among other shapes, needle-like crystals at 6.5°C (Figure 5.15) and 20°C (Figure 5.16). None of them could be identified as protein crystals.

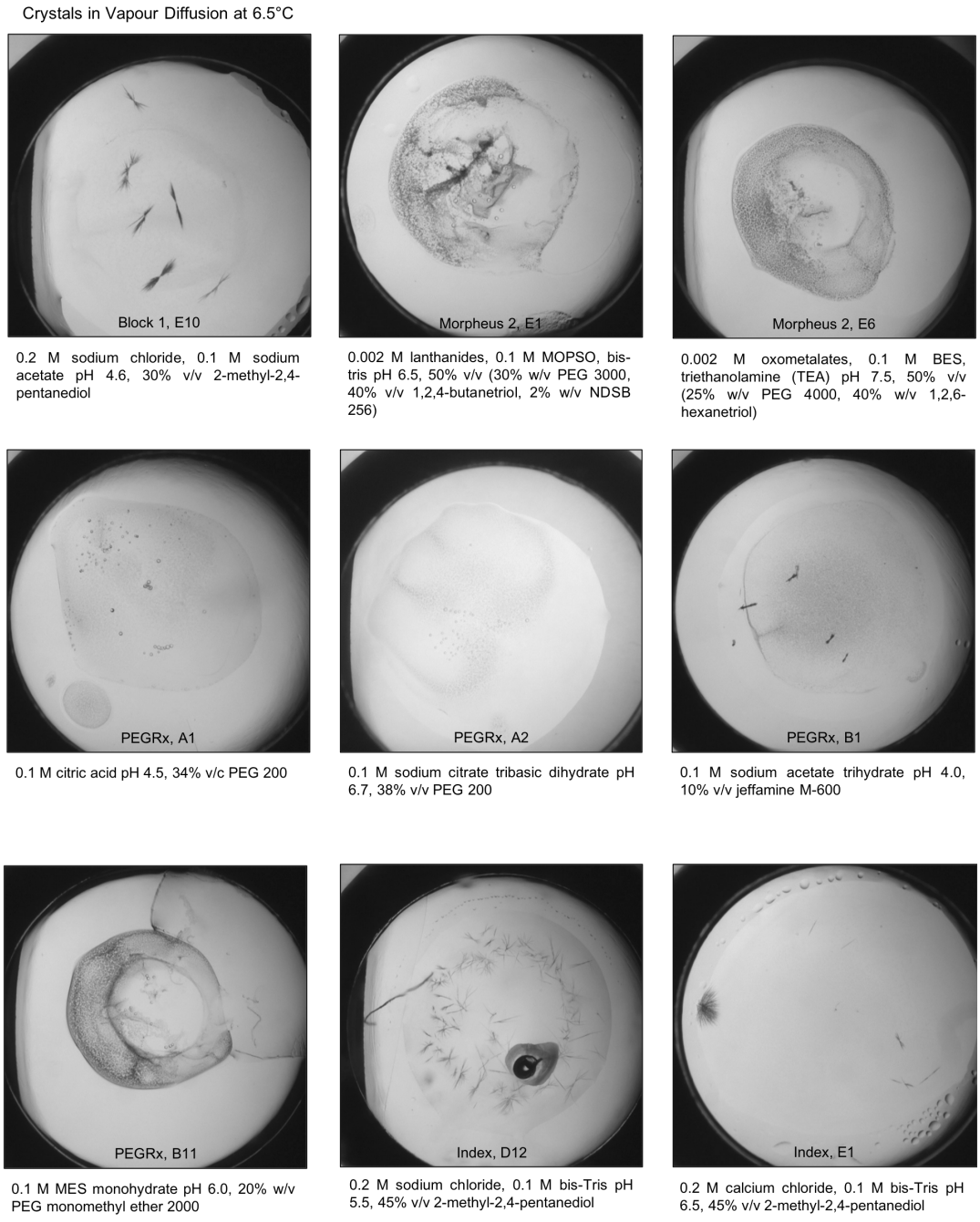


Figure 5.15: **Crystallisation trials of the Ptc1-Nb64 complex in DDM at 6.5°C.** Crystallisation was set up in vapour diffusion using a 1:1 drop ratio (protein:mother liqueur). Crystals were grown at 6.5°C and 20°C and appeared in needle-like or round shape. Detergent and salt crystals, but no protein crystals were obtained.

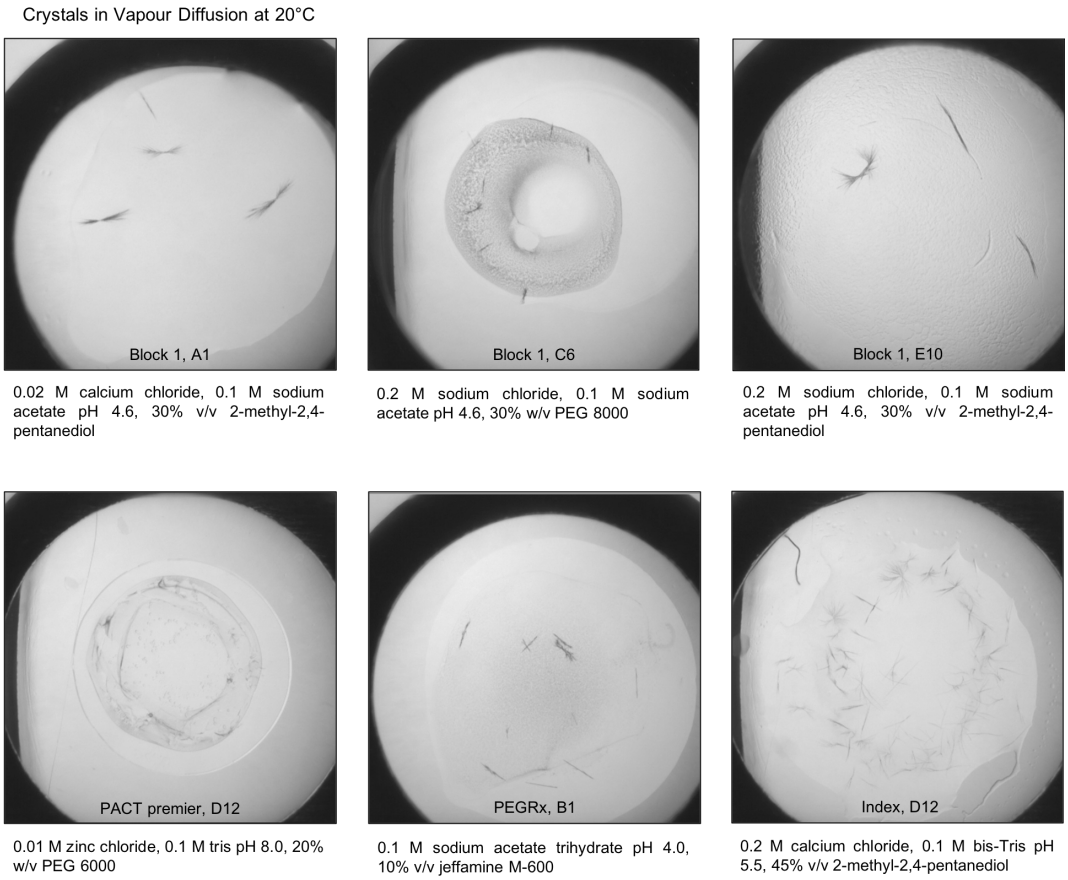


Figure 5.16: **Crystallisation trials of the Ptc1-Nb64 complex in DDM at 20°C.** Crystallisation was set up in vapour diffusion using a 1:1 drop ratio (protein:mother liqueur). Crystals were grown at 6.5°C and 20°C and appeared in a needle-like or round shape. Detergent and salt crystals, but no protein crystals were obtained.

For EM studies, the Ptc1-Nb64 complex was exchanged into Amphipols. Not only can Amphipols adsorb to the membrane protein, they also substitute its rather large detergent micelle. Just as detergent, these amphiphilic polymers can keep the membrane protein in solution. Due to lower sample concentration, the SEC profile revealed two peaks at ∼12.5 mL and ∼13.5 mL retention volume, however, the peaks could not be fully resolved (Figure 5.17, A). Almost no aggregation was observed and free Nb64 eluted at ∼16.5 mL retention volume. SDS-PAGE analysis confirmed the presence of dimer and monomer in the sample, co-eluting with Nb64

(Figure 5.17, B). SEC fractions marked with an orange or lilac dot were examined by electron microscopy. The sample was applied onto carbon-coated copper grids (CF300-Cu, EMS) and stained with uranyl formate stain (0.75% aqueous solution) Negative stain analysis using a Tecnai T12 (FEI) microscope displayed single, non-aggregated particles with an even distribution of either, dimer or monomer; an ideal sample for cryo-EM studies (Figure 5.18, A-B). Unfortunately, vitrification caused aggregation and denaturation of the protein, a phenomenon that is not uncommon in cryo-EM studies of biological specimens [13]. This problem might be circumvented by using a new strategy presented by S. Arnold and co-workers (2017). Sample preparation bypasses the blotting step which might prevent protein denaturation [13].

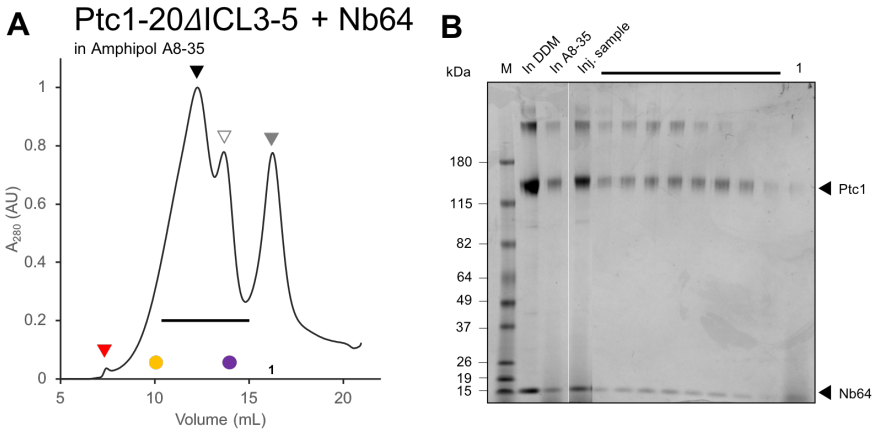


Figure 5.17: **Purification and analysis of Amphipol-exchanged Ptc1-20ΔICL3-5 in complex with Nb64. A-B**, SEC profile (A) with corresponding SDS-PAGE (B) analysis of Ptc1 ICL3 deletion mutant Ptc1-20ΔICL3-5 in complex with Nb64 in Amphipols. **A**, Red arrowhead indicates the peak of aggregated protein species, black and black outlined arrowhead the peak of dimeric and monomeric Ptc1. Grey arrowhead indicates the peak of Nb64. **B**, Lanes from left to right: BenchMark^{TM} ladder (M), sample in DDM, sample in Amphipols (A8-35), sample subjected to SEC (Inj. sample), Nb64 (1). SEC fractions across the Ptc1 peak fraction are indicated by a black line.

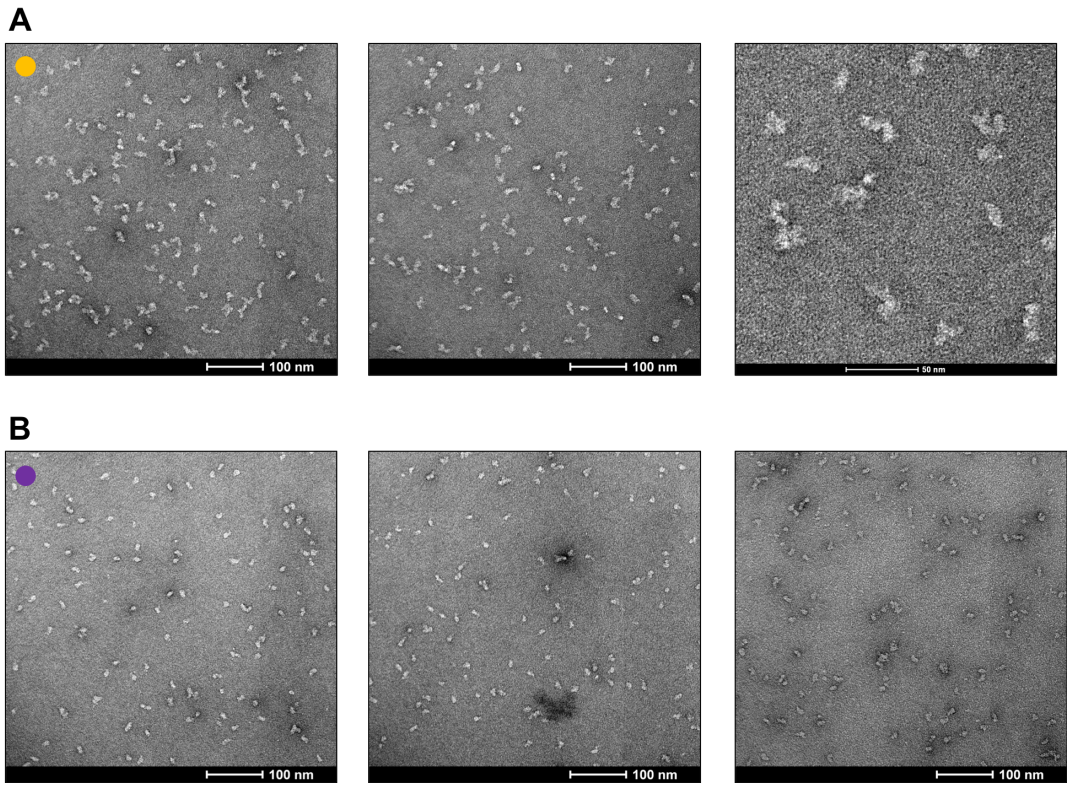


Figure 5.18: **Negative stain of Amphipol-exchanged Ptc1-20ΔICL3-5 in complex with Nb64. A-B**, Images of negative stained Ptc1 dimer (A) and monomer (B) in Amphipols originating from SEC fractions marked in Figure 5.17 with an orange and lilac dot, respectively. The sample was not aggregated, appeared homogeneous, with a good particle distribution.

5.3.4 BRIL, a fusion partner to improve crystallisability

Another method to improve protein crystallisability is to introduce a fusion partner to the protein of interest. For instance, T4 lysozyme and thermostabilised cytochrome b562 RIL (bRIL) have been studied extensively, especially for crystallisation of GPCRs. These proteins introduce hydrophilic regions to proteins of hydrophobic nature, and therefore help to form crystal contacts. We decided to introduce a bRIL domain into one of the Ptc1 intracellular loop regions ICL1, ICL2, ICL3, ICL4 or ICL5, either keeping the ICL3 (as in Ptc1-20) or deleting it (as in Ptc1-20ΔICL3-5). Details about the constructs can be found in Figure 5.19, C. All constructs were

synthesized by overlapping PCR and cloned into the pHLsec vector with C-terminal Rho1D4 antibody epitope tag. Constructs used for small-scale screening in FSEC included a mV-tag before the purification tag. Constructs derived from Ptc1-20, with a bRIL insertion in ICL1, ICL2, ICL4, or ICL5, displayed relatively low expression levels and two peaks in FSEC, corresponding to Ptc1 dimer and monomer (Figure 5.19, A). All constructs with the exception of Ptc1-20ΔICL5-bRIL showed lower expression than Ptc1-20 itself. Constructs derived from Ptc1-20ΔICL3-5 (Figure 5.19, B), eluted as a single, monodisperse peak; however, expression levels were substantially reduced, with Ptc1-20ΔICL3-5-ΔICL1-bRIL showing almost no expression. The low expression of this construct might be explained by the fact that four residues that were substituted by bRIL were predicted by the UniProt server to be part of the preceding transmembrane domain [18]. Interestingly, the control construct with ICL3 linker deletion and bRIL insertion (Ptc1-20ΔICL3-5-bRIL) showed a 30% reduction in expression level compared to the equivalent construct without bRIL (Ptc1-20ΔICL3-5). Despite lower expression, we decided to take the construct to large scale for the above mentioned reasons.

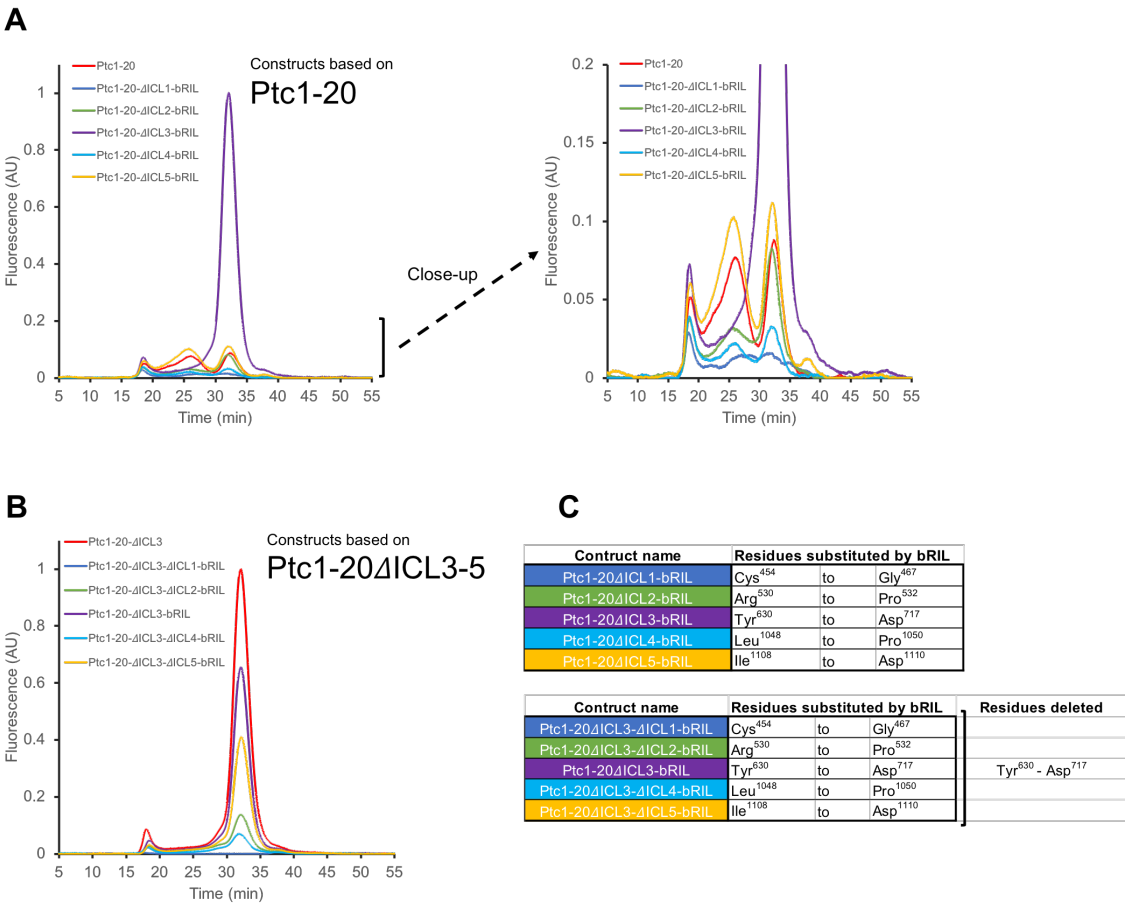


Figure 5.19: **FSEC screening of Ptc1 mutants with bRIL insertion. A-B**, FSEC analysis of Ptc1-20-based (A) and Ptc1-20ΔICL3-5-based (B) constructs with bRIL insertions into one of the ICLs. **A**, Ptc1-20 and mutants including bRIL in either ICL1, ICL2, ICL4, or ICL5, eluted as two, slightly overlapping peaks corresponding to Ptc1 dimer and monomer. FSEC of the Ptc1 mutant with bRIL insertion in ICL3 displayed one monodisperse peak at retention times typical for Ptc1 monomer. **B**, All Ptc1 mutants with ICL3 linker deletion and control construct Ptc1-20ΔICL3-5 eluted as single monodisperse peaks potentially resulting from a Ptc1 monomer. **C**, Overview of constructs tested in FSEC (A and B). Residues substituted by bRIL are indicated.

Ptc1-20ΔICL3-5-bRIL was expressed without a mV-tag at large scale in HEK293S GnTI$^-$ cells. Purification was carried out in DDM as described earlier. The SEC profile displayed two main but non-resolved peaks, and another peak that indicated aggregation of the protein (Figure 5.20, A). As expected, SDS-PAGE analysis revealed two bands corresponding to the size of Ptc1 dimer and monomer and some minor impurities of lower molecular weight (Figure 5.20, B).

To determine the molecular weight of Ptc1 in solution, multi-angle light scattering (MALS) experiments were carried out (Figure 5.20, E). As scattering of the detergent can be problematic, the experiment was carried out with Amphipol-exchanged sample. Purification and SDS-PAGE analysis suggested that both, Ptc1 dimer and monomer were present in the sample (Figure 5.20, C-D). For best light scattering results, the MALS detector was coupled to a SEC system. Protein-conjugate analysis of the MALS data using ASTRA chromatography software 6.1 (Wyatt Technology) revealed molecular weights of 325.5 kDa and 183.9 kDa (Figure 5.20, E). The molecular mass contribution of Amphipols is estimated to be 40- 54 kDa [298]; calculated values for the proteins resulted in a molecular weight range of 286- 272 kDa and 144- 130 kDa, respectively. The values are close to the theoretical molecular weight of the glycosylated Ptc1-20ΔICL3-5-bRIL monomer (∼143 kDa) and dimer (∼286 kDa).

For crystallisation experiments, Ptc1-20ΔICL3-5-bRIL was co-purified with Nb64 in DDM (Figure 5.21, A). SDS-PAGE analysis confirmed the presence of Ptc1 dimer and monomer, with fractions at later retention volumes mainly containing the monomeric Ptc1-Nb64 complex (Figure 5.21, B). Western blot analysis revealed minor degradation of the protein (Figure 5.21, B). The sample was concentrated and treated with EndoF1, followed by crystallisation in VD at 6.5- 8.0 mg mL^{-1} in 96-well 3-drop Swissci plates or LCP at 20- 54 mg mL^{-1} in conventional glass sandwich plates. VD crystallisation of the protein at a 1:1 protein:reservoir drop ratio resulted in precipitation across many conditions at both, 6.5°C and room temperature. Crystals were observed in two conditions, but did not yield protein crystal diffraction (Figure 5.22).

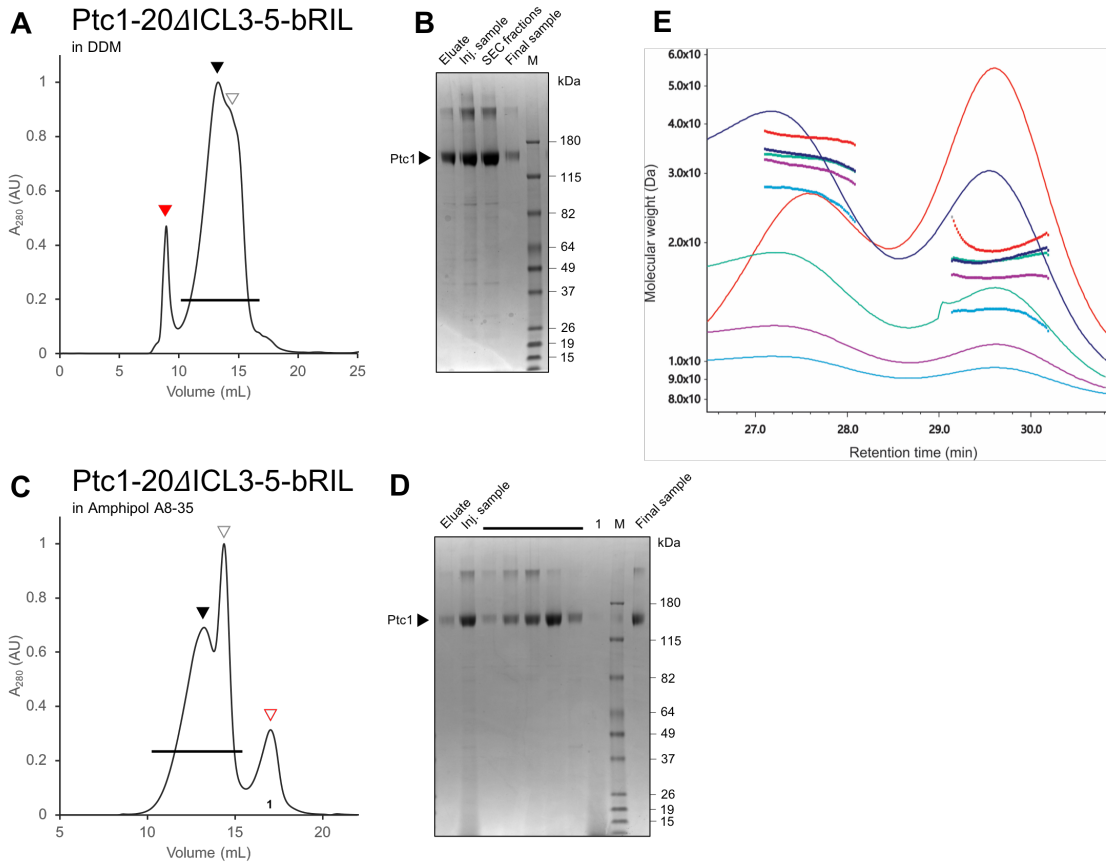


Figure 5.20: **MALS of Ptc1-20ΔICL3-5-bRIL in Amphipols. A-B**, SEC profile (A) with corresponding SDS-PAGE (B) of Ptc1-20ΔICL3-5-bRIL purified in DDM. **A**, Red arrowhead indicates the peak of aggregated protein species, black and black outlined arrowhead the peak of dimeric and monomeric Ptc1, respectively. **B**, Lanes from left to right: Eluate of affinity purification, sample subjected to SEC (Inj. sample), pooled sample from SEC fractions (indicated by black line in A), sample for Amphipol-exchange, BenchMark^{TM} ladder (M). **C-E**, SEC profile (C) with corresponding SDS-PAGE (D) and MALS analysis (F) of Ptc1-20ΔICL3-5-bRIL exchanged into Amphipols. **C**, Black and black outlined arrowhead indicate the peak of dimeric and monomeric Ptc1, respectively. Red outlined arrowhead indicates peak of degradation products. **D**, Lanes from left to right: Eluate of affinity purification, sample subjected to SEC (Inj. sample), SEC fractions indicated by black line in A, degradation products (1), BenchMark^{TM} ladder (M), sample for MALS experiment. **E**, MALS analysis of Ptc1-20ΔICL3-5-bRIL in Amphipols. A five point, two-fold dilution series was prepared and samples were loaded onto a Superose 6 Increase 10/30 column (GE Healthcare). SEC traces shown in blue, red, cyan, purple, and light blue correspond to a sample concentration of 1 mg mL^{-1}, 0.5 mg mL^{-1}, 0.25 mg mL^{-1}, 0.125 mg mL^{-1}, and 0.06 mg mL^{-1}, respectively. Protein-conjugate analysis revealed molecular weights of 286 kDa for the Ptc1 dimer and 144 kDa for the Ptc1 monomer. MALS experiments and analysis were performed in collaboration with D. Staunton.

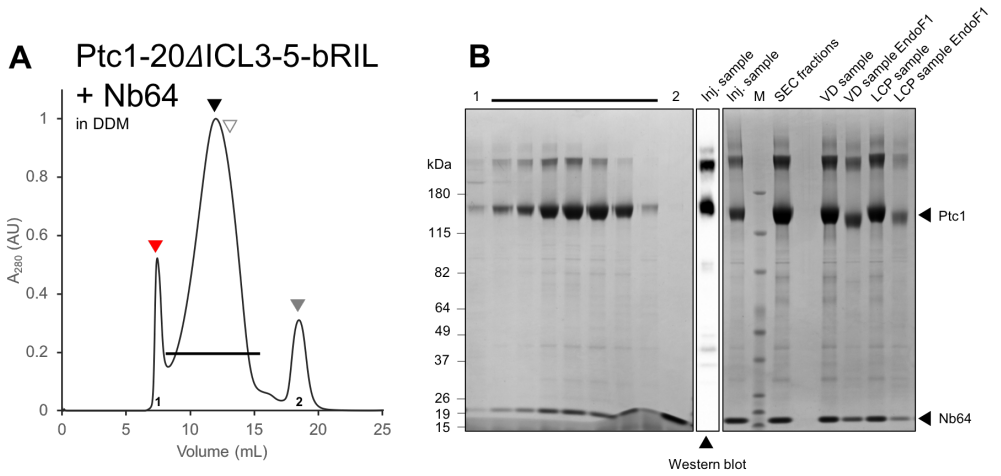


Figure 5.21: **Purification and analysis of Ptc1-20ΔICL3-5-bRIL in complex with Nb64 in DDM. A-B**, SEC profile (A) with corresponding SDS-PAGE analysis (B) of Ptc1-20ΔICL3-5-bRIL in complex with Nb64 in DDM. **A**, Red arrowhead indicates the peak of aggregated protein species, black and black outlined arrowhead the peak of dimeric and monomeric Ptc1-Nb64 complex, respectively. Grey arrowhead indicates the peak of Nb64. **B**, Lanes from left to right: Aggregate (1), SEC fractions across the Ptc1 peak fraction are indicated by a black line, Nb64 (2), BenchMark^{TM} ladder (M), sample subjected to SEC (Inj. sample), BenchMark^{TM} ladder (M), pooled SEC fractions (as indicated by a black line), sample for crystallisation in VD before and after EndoF1 treatment, sample for crystallisation in lipidic cubic phase (LCP) before and after EndoF1 treatment. The sample consisted of Ptc1 dimer and monomer in complex with Nb64.

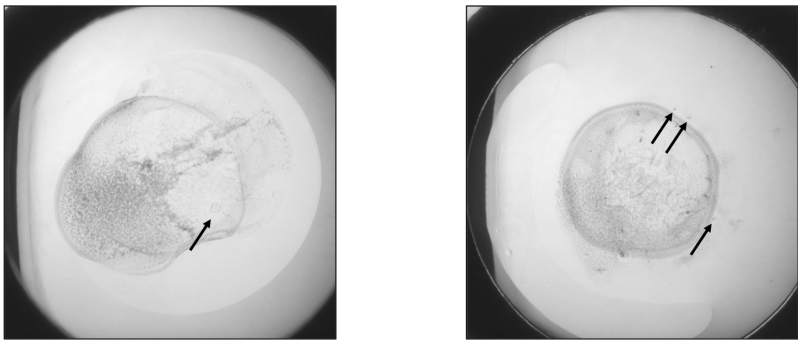


Figure 5.22: **Crystallisation of Ptc1-20ΔICL3-5-bRIL in complex with Nb64 in DDM.** Crystallisation was set up in vapour diffusion using a 1:1 (protein:mother liqueur) drop ratio. Crystals were obtained in two different conditions, however, they were highly birefringent, which is an indicator of salt crystals.

Besides crystallisation of Ptc1 in DMNG or DDM micelles, the protein was reconstituted in DMPC:CHAPSO bicelles before setting up VD crystallisation. Ptc1 at 10 mg mL^{-1} was mixed with 30% DMPC:CHAPSO (Anatrace) at a 4:1 protein:lipid (v/v) ratio until the samples regained a transparent appearance. The sample was incubated on ice for 30 min before crystallisation trials were set up at a 1:1 protein:reservoir drop ratio. Crystals were observed across a few conditions, however, they were proven to be of a lipidic nature (Figure 5.23, A). Crystallisation in LCP was carried out in either 9.9 MAG (monoolein), 7.8 MAG, or 7.7 MAG (Anatrace). The lipids differ in the length of their acyl carbon chain and the position of the *cis* double bond (see Appendix C for the chemical structures of the lipids). This influences the thickness of the lipid bilayer formed and the curvature thereof [40]. Monoolein is the most frequently used lipid for membrane protein crystallisation in LCP. Therefore, the Ptc1-Nb64 complex was reconstituted in monoolein at a 3:2 (w/v) lipid:protein ratio. Crystals were observed in one condition (30% PEG550), however, at the time the crystals were mounted, the cubic phase had already changed to sponge phase, which lowered the chance of yielding good protein diffraction (Figure 5.23, B). Diffraction data were inconclusive and crystallisation trials in this condition should be repeated. 7.8 MAG was the only lipid that allowed the LCP sample to clarify after mixing the two components at a 1:1 (w/v) lipid:protein ratio. Samples reconstituted in 7.7 MAG were prepared in a similar way, but remained turbid after mixing. Crystallisation trials in LCP were set up in all available crystallisation screen conditions (provided by S. Newstead, University of Oxford, UK) at 4.0°C and room temperature. Unfortunately, no crystals were observed using 7.8 MAG or 7.7 MAG.

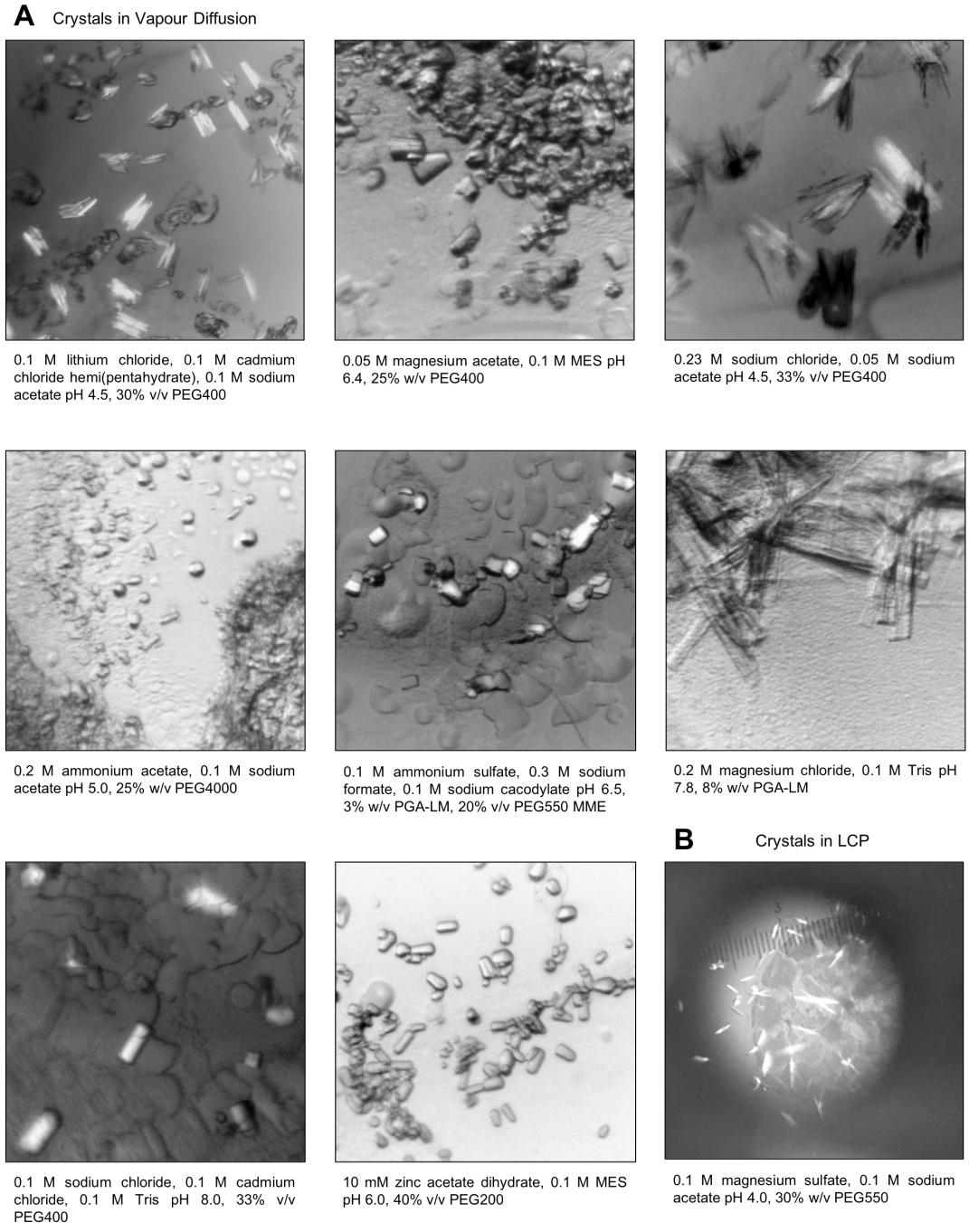


Figure 5.23: **Reconstitution of Ptc1-20ΔICL3-5-bRIL-Nb64 complex using lipidic bicelles and cubic phase. A**, Crystallisation in lipidic bicelles yielded crystals in different shapes across various crystallisation conditions. However, crystals were of lipidic nature. **B**, Crystallisation in lipidic cubic phase (LCP) using monoolein yielded in needle-like crystals in one conditions. Diffraction data were inconclusive as the phase had transitioned to sponge phase before the crystals were mounted.

In protein X-ray crystallography, protein engineering plays a vital role. One major focus is to improve protein stability as this is very likely to influence the protein's ability to crystallise [72]. To investigate the stability of the engineered Ptc1 construct Ptc1-20ΔICL3-5-bRIL in complex with Nb64 (in DDM), samples were incubated at 11 different temperatures for 1 hour, before being subjected to analytical SEC (Figure 5.24). At temperatures of 50°C and below, the sample eluted as two overlapping peaks, one formed by Ptc1 dimer and one by Ptc1 monomer. While at lower temperatures the dimer peak was more prominent, the amount of monomer increased with a rise in temperature. This is not unexpected, as contacts between residues involved in dimer interface formation start to weaken, causing the dimer to fall apart and the peak to shift. At 55°C and above, the protein gradually started to aggregate due to partial denaturation. However, even temperatures as high as 80°C did not result in complete protein precipitation or denaturation. As the sample contained Ptc1 dimer and monomer, it was difficult to determine the melting temperature (T_M) of the protein. Nevertheless, distinct protein peaks started to disappear at ∼55°C, suggesting that the protein is thermostable up to this temperature.

Similar results were obtained when studying the thermostability of the Ptc1-Nb64 complex using a NanoTemper (Prometheus) instrument, which uses intrinsic Trp and Tyr fluorescence to measure the protein's stability. Measurements revealed a T_M of 60.1°C for the Ptc1-20ΔICL3-5-bRIL-Nb64 complex in DDM and 51.9°C in Amphipols. The Ptc1-Nb64 complex of Ptc1-20ΔICL3-5 had a T_M of 61.0°C, suggesting that the construct without bRIL is slightly more stable.

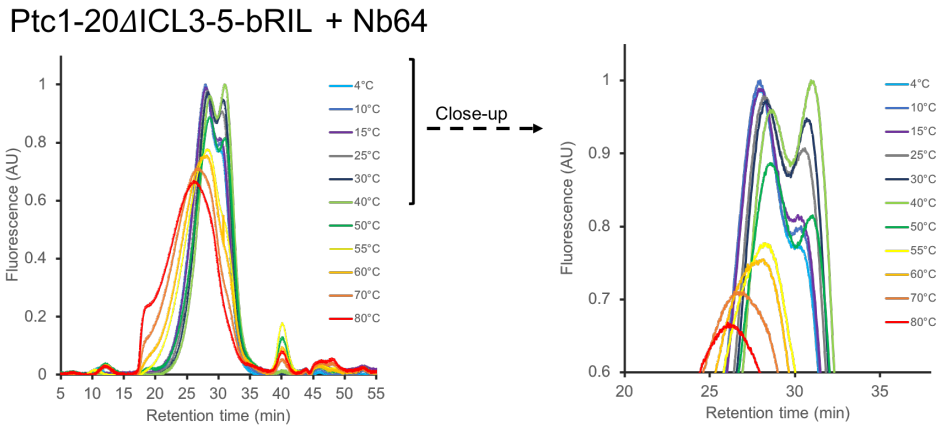


Figure 5.24: **Thermostability of Ptc1-20ΔICL3-5-bRIL in complex with Nb64.** Traces of analytical SEC after incubating the protein complex at different temperatures for 1 hour. The rise in temperature led to partial dissociation of the dimer into monomers. At 55°C, the protein started to denature as indicated by the disappearance of two distinct protein peaks. At 60°C, the protein started to aggregate. Results suggested that the protein complex is stable up to 55- 60°C.

5.3.5 Reconstitution of Ptc1 in nanodiscs for cryo-EM studies

Membrane scaffold proteins (MSPs) are amphipathic helical proteins that form a protein belt around a phospholipid bilayer by making contact with the hydrophobic alkyl chains of the lipids. Typically, two MSPs wrap around the phospholipid bilayer; one MSP per leaflet. Apolipoproteins are the prototypes of MSPs; they have been engineered to fit the purpose [74]. MSPs are available at different molecular weights and the size of the MSP controls the number of phospholipids integrated. In this way, the size of the nanodisc can be tailored to the size of the membrane protein of interest. The composition of the lipid bilayer not only depends on the choice of lipid, but can also be modulated by adding other molecules, such as cholesterol. Besides, nanodisc formation is temperature dependent, and most efficient near the phase transition temperature of the phospholipid [20].

Structure determination of membrane proteins is not a straightforward process; their hydrophobic nature and the presence of detergent makes them difficult to crystallise. However, membrane proteins embedded into nanodiscs represent an ideal sample for cryo-electron microscopy studies. Their size is suitable for microscopy, and the stability of the protein is considerably improved. In order to study Ptc1 in nanodiscs, two MSPs of different size were chosen, MSP1D1 and MSP2N2. The DNA sequences were acquired from Addgene in a pET28a backbone with C-terminal His_7-tag [74]. Proteins were expressed in *E.coli* BL21-Gold(DE3) cells (kind gift from E. Carpenter, University of Oxford, UK) and purified via IMAC, followed by SEC. Peak fractions from SEC were pooled and concentrated to 5 mg mL^{-1}. Proteins were flash-frozen in liquid nitrogen and stored at -80°C before being used for reconstitution experiments. Figure 5.25, A-B displays the SEC profile of MSP1D1 and MSP2N2 as analysed on analytical SEC using a Superose 6 3.2/300 column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl. First, both MSPs were studied in combination with three different phospholipids purchased from Avanti Polar Lipids: POPC, DMPC, and DPPC (see Appendix C for the chemical structures of the lipids). Samples were prepared by mixing the MSP with the lipid at a molar ratio of 1:65, 1:80, or 1:90 for MSP1D1 and 1:150, 1:160, or 1:190 for MSP2N2 with POPC, DMPC, or DPPC, respectively [5, 19]. In the presence of 12 mM sodium cholate, the samples were incubated for 45 min at the lipid's phase transition temperature: 4°C for POPC, 25°C for DMPC, and 37°C for DPPC. To initiate nanodisc self-assembly, sodium cholate was removed by adding polystyrene beads (BioBeads SM-2, Biorad). The samples were subsequently analysed on analytical SEC. MSP1D1:lipid nanodiscs displayed a single monodisperse peak at ∼22 min retention time (Figure 5.25, C). The SEC profile of MSP2N2:lipid nanodiscs displayed an aggregation peak, followed by multiple overlapping peaks (Figure 5.25, D). Indeed, it has been shown previously

that MSP2N2 can generate nanodiscs with a broader size distribution [106]. SDS-PAGE analysis revealed a protein band at the expected molecular weight of MSP1D1 (24.7 kDa) and MSP2N2 (43.1 kDa).

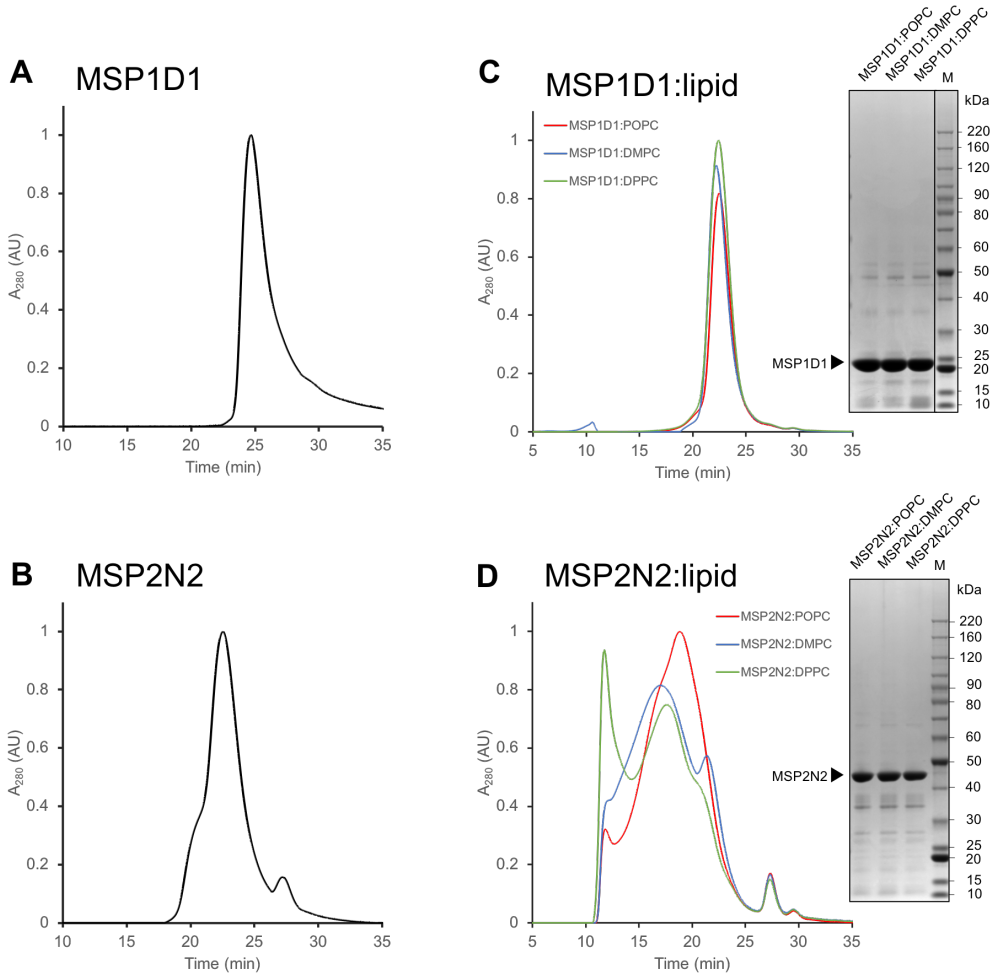


Figure 5.25: **Analysis of MSP nanodisc formation using analytical SEC. A-B**, MSP1D1 (A) and MSP2N2 (B) analysed on analytical SEC (Superose 6 3.2/300 column). **A**, MSP1D1 eluted as a single monodisperse, slightly asymetric peak. **B**, MSP2N2 displayed one main peak with a shoulder on the left side. The additional peaks left and right of the main peak might result from a dimeric form of MSP2N2 and/or impurities. **C-D**, Overlay of MSP1D1 (C) and MSP2N2 (D) after mixing with lipids POPC, DMPC, or DPPC. **C**, MSP1D1-based nanodiscs eluted as a single monodisperse peak. **D**, MSP2N2-based nanodiscs displayed several overlapping peaks on SEC, possibly resulting from MSP2N2, MSP2N2:lipid nanodiscs, and aggregated protein.

Next, we examined nanodisc formation in the presence of mV-tagged Ptc1. This elegant approach enabled us to locate Ptc1 in nanodiscs by recording protein absorption and fluorescence in parallel. Ptc1-20ΔICL3-5-mV was cloned into the pHLsec vector with an C-terminal mV-tag, followed by a Rho1D4 antibody epitope tag. HEK293S GnTI^{-} cells were transiently transfected with the construct and fluorescently-tagged Ptc1 was expressed in suspension tissue culture for 60 hours post-transfection. Proteins were solubilised using 1.5% DDM, purified via Rho1D4 antibody affinity chromatography, and exchanged into 10 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, and 0.03% DDM 10:1 CHS by SEC using a Superose 6 Increase 10/30 column (GE Healthcare). The SEC profile revealed relatively large amounts of protein aggregation and a broad peak for Ptc1 (Figure 5.26, A). As previously observed, the peak displayed a small shoulder on the left side of the main peak resulting from Ptc1 dimer. SDS-PAGE analysis confirmed purity and the presence of Ptc1 dimer and monomer in the sample (Figure 5.26, B). The SEC peak fractions were pooled and concentrated to 3.9 mg mL^{-1}.

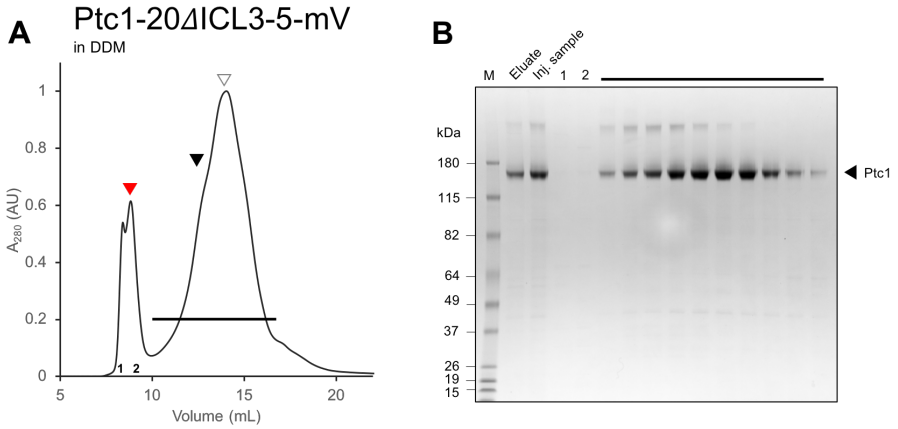


Figure 5.26: **Purification and analysis of Ptc1-20ΔICL3-5-mV in DDM. A-B**, SEC profile (A) with corresponding SDS-PAGE (B) analysis of Ptc1-20ΔICL3-5-mV in DDM. **A**, Red arrowhead indicates peaks of aggregated protein species, black and black outlined arrowhead the peak of dimeric and monomeric Ptc1. **B**, Lanes from left to right: BenchMark^{TM} ladder (M), eluate of affinity purification, sample subjected to SEC (Inj. sample), aggregated protein species (1, 2), SEC fractions across the Ptc1 peak fraction are indicated by a black line.

Nanodiscs were prepared by combining fluorescently-tagged Ptc1 in detergent buffer with the MSP and lipid of choice. After incubation at the phase transition temperature of the lipid, BioBeads MS-2 were added to remove the detergent from the sample. This step initiates spontaneous assembly of nanodiscs around Ptc1. The next day, samples were analysed using analytical SEC. Figure 5.27, A-C shows an overlay of MSP1D1, Ptc1, empty nanodiscs and Ptc1 reconstituted in nanodiscs. To monitor whether Ptc1 had been incorporated into nanodiscs, proteins were detected at different wavelength; fluorescently-tagged Ptc1 was detected at Ex./Em. 515/528 nm, whereas all other proteins were detected at the absorbance of 280 nm. As expected, MSP1D1 displayed a single monodisperse peak which shifted to shorter retention times (higher molecular weights) when mixed with lipids to form empty nanodiscs. Amphipol-exchanged Ptc1 eluted as a monodisperse peak (Ptc1 monomer) with a minor shoulder on the left side of the peak (Ptc1 dimer). Ptc1 reconstituted in nanodiscs displayed two overlapping peaks and an overlay with the absorption spectrum recorded at Ex./Em. 515/528 nm revealed that the peak at earlier retention times (∼17.5 mL) originated from nanodiscs that contained Ptc1, whereas the peak at a later retention time (∼22 mL) originated from empty nanodiscs present in the sample. Similar results were obtained for Ptc1 reconstituted in MSP2N2 nanodiscs (Figure 5.28, A-C). However, MSP2N2 eluted as a monodisperse peak with a shoulder apparent on its left side. This resulted in two overlapping peaks for the empty MSP2N2:lipid nanodiscs. In the presence of Ptc1, the peak at shorter retention times broadened. An overlay with the recorded Ptc1 SEC trace at Ex./Em. 515/528 nm indicated that this peak was formed by Ptc1-reconstituted, but also empty nanodiscs. Both nanodisc types showed aggregation, which was most pronounced when using DPPC.

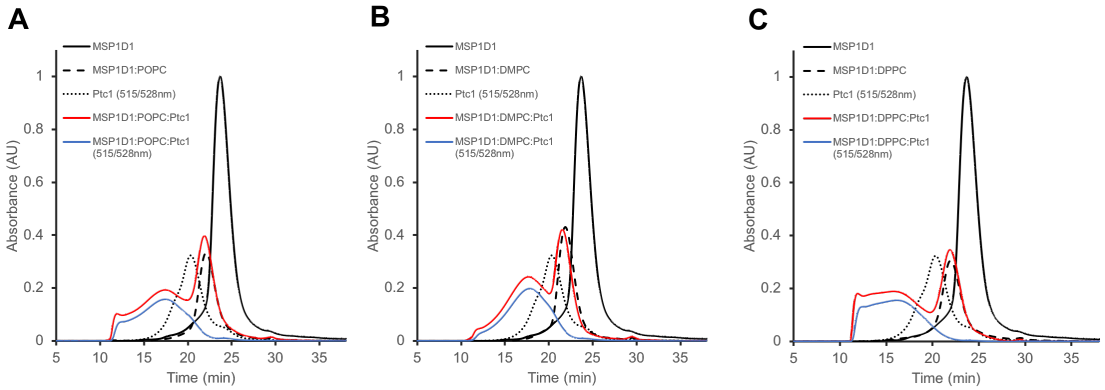


Figure 5.27: **SEC analysis of MSP1D1-based nanodiscs. A-C**, Analytical SEC of MSP1D1:POPC:Ptc1 (A), MSP1D1:DMPC:Ptc1 (B), and MSP1D1:DPPC:Ptc1 (C) nanodiscs. Overlay of SEC traces of MSP1D1 and Ptc1, with empty nanodiscs and Ptc1 reconstituted nanodiscs. Ptc1 and Ptc1 in nanodiscs were monitored by mV fluorescence detection. Ptc1 was successfully reconstituted in all types of nanodiscs.

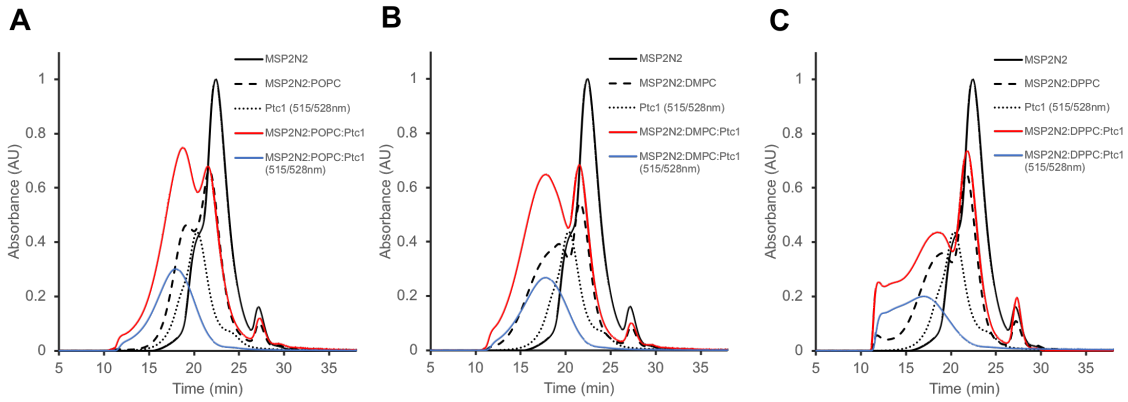


Figure 5.28: **SEC analysis of MSP2N2-based nanodiscs. A-C**, Analytical SEC of MSP2N2:POPC:Ptc1 (A), MSP2N2:DMPC:Ptc1 (B), and MSP2N2:DPPC:Ptc1 (C) nanodiscs. Overlay of SEC traces of MSP2N2 and Ptc1, with empty nanodiscs and Ptc1 reconstituted nanodiscs. Ptc1 and Ptc1 in nanodiscs were monitored by mV fluorescence detection. Ptc1 was successfully reconstituted in all types of nanodiscs.

Based on the screening results, we decided to prepare MSP2N2:DMPC:Ptc1 nanodiscs at a larger scale for cryo-EM studies. Ptc1:MSP2N2 was mixed at a 1:8 molar ratio, and lipid was added at a 1:180 MSP2N2:DMPC molar ratio. The reaction was carried out in detergent buffer for 90 min at room temperature, followed by

incubation with polystyrene beads overnight. The detergent-free eluate was purified via Rho1D4 antibody affinity chromatography to remove empty nanodiscs from the sample. The eluate containing Ptc1-reconstituted nanodiscs was collected and subjected to SEC using a Superose 6 Increase 10/30 column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl. The SEC trace revealed a broad peak with a small shoulder on its right side (Figure 5.29, A). Negatively stained protein samples of the SEC peak fraction were prepared using uranyl formate stain (0.75% aqueous solution) and carbon-coated copper grids (CF300-Cu, EMS). The grids were examined on a Tecnai T12 (FEI) microscope. Particles were evenly distributed and their adsorption onto the grids was not limited to a preferred orientation (Figure 5.29, C). Most particles displayed a size of approximately 20 nm, close to the predicted diameter of MSP2N2-based nanodiscs (17 nm) [106]. For cryo-EM studies, samples were applied onto three different types of grids (Ultra-thin lacey carbon on 400 mesh copper grids [Agar Scientific], C-Flat holey carbon on 300 mesh copper grids [Protochips], and Quantifoil® holey carbon R 1.2/1.3 on 400 mesh gold grids [EMS]) and vitrified in liquid ethane. Analysis of samples on C-Flat holey carbon did not yield promising results, and samples on Quantifoil® holey carbon displayed considerable protein aggregation. However, Ultra-thin lacey carbon grids displayed a good number of particles per grid square and a small dataset was collected using a Tecnai F30 (FEI) microscope equipped with a Gatan K2 detector (Figure 5.29, D). Unfortunately, the sample appeared to be slightly heterogeneous and only a small number of particles were subjected to 2D classification; no meaningful 2D classes were obtained. In addition, stacking of nanodiscs was observed, which lowered the number of particles that could be picked.

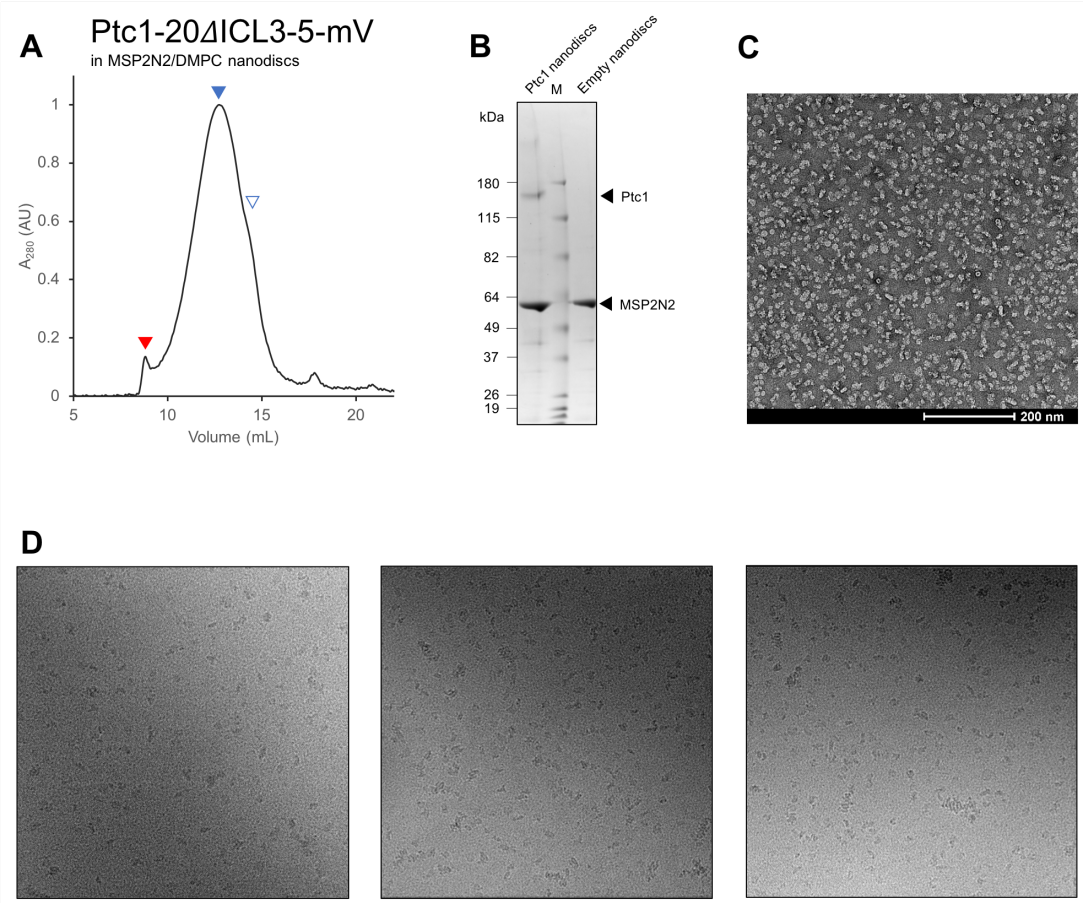


Figure 5.29: **Purification and analysis of Ptc1-20ΔICL3-5-mV in nanodiscs. A-B**, SEC profile (A) with corresponding SDS-PAGE (B) analysis of Ptc1 in MSP2N2:DMPC nanodiscs. **A**, Red arrowhead indicates peaks of aggregated protein species, blue and blue outlined arrowhead the peak of MSP2N2:DMPC:Ptc1 nanodiscs and MSP2N2:DMPC nanodiscs, respectively. **B**, Lanes from left to right: MSP2N2:DMPC:Ptc1 nanodiscs, BenchMark^{TM} ladder (M), MSP2N2:DMPC nanodiscs. **C**, Negative stain of MSP2N2:DMPC:Ptc1 nanodiscs. The sample was not aggregated, appeared homogeneous, with a good particle distribution and no preferred orientation. **D**, Cryo-EM micrographs of MSP2N2:DMPC:Ptc1 nanodiscs. The sample appeared slightly heterogeneous and nanodisc stacking was observed.

5.4 Discussion

In Hh signalling, Ptc1 plays a vital role. In the absence of the Hh ligand, Ptc1 constitutively represses the pathway. The exact mechanism by which Ptc1 inhibits the GPCR Smo is unknown, however, it has been proposed that inhibition is achieved via a small molecule such as cholesterol [39, 122, 173, 264, 272]. Once the Hh ligand binds to Ptc1, Hh signalling is activated, tightly regulating embryonic development and maintenance of adult stem cells. Mutations in Ptc1, which lead to aberrant signalling, have been implicated in a range of diseases, such as basal cell carcinoma (BCC), medulloblastoma and other types of cancer. Even though the role of Ptc1 in Hh signalling has been studied extensively, particularly in terms of signal transduction and its role in embryonic development and disease, no structural information of Ptc1 was available until recently [101, 227, 228]. The reason can be attributed to the challenges that come with membrane protein (MP) research. Extensive optimisation work needs to be carried out in order to study these proteins, and the process is often further complicated when working with MPs that consist of multiple subunits.

The aim of this study was to optimise the expression of human Ptc1, develop a purification strategy and to identify a suitable construct for crystallisation. We further envisaged determining the three-dimensional structure of Ptc1 in order to draw conclusions about its molecular mechanism of action. First, we optimised the expression of Ptc1; the wildtype construct showed only poor expression. Shortening the N- and C-termini, as well as introducing a deletion mutation in ICL3 led to a significant increase in protein yield. These deletions not only reduced flexible regions, but also removed two PPXY motifs, one located in the C-terminus and another one in the ICL3 of Ptc1. These motifs are targeted by E3 ubiquitin ligases, which mark the protein for degradation in lysosomes [294]. Even though the C-terminus of Ptc1 was

predicted to be involved in multimerisation of the protein, Ptc1 could still be isolated as a dimer. We further tested the localisation of the Ptc1 mutants in NIH/3T3 $Ptc^{-/-}$ fibroblasts to examine whether the engineered constructs are biologically active. Cellular assays revealed a potential cilia localisation motif located between residues Phe^{611} and Tyr^{630}; constructs missing this region were not found in the primary cilium. In addition, a signal sequence for internalisation might be located between residues Val^{692} and Asp^{717}; constructs lacking this region exhibited a significantly higher membrane localisation. Untransfected cells showed Smo localisation in cilia, indicating that the pathway is inactivated in the absence of Ptc1.

Inserting the fusion partner protein bRIL in ICL3 did not improve protein yield, however, it increased the intracellular surface area and the hydrophilic properties of Ptc1, which is thought to be beneficial for crystal contact formation. In parallel to construct design, we identified suitable detergents and buffer conditions, which aided extraction of the protein, besides keeping it soluble and stable throughout the elaborate purification process. In addition, we identified nanobodies that bind to Ptc1 with nanomolar affinity; co-purification of Ptc1 with Nb64 improved protein yield by approximately 40-fold. Furthermore, we tested the thermostability of the Ptc1-Nb64 complex by FSEC and determined a value of 55°C, indicating that the protein complex exhibits high stability.

In summary, milligram quantities of Ptc1 could be obtained, suitable for crystallisation trials in a traditional VD or LCP setup. Various constructs with and without Shh or Nb64 were screened against numerous crystallisation conditions at 4°C, 6.5°C and 20°C. To crystallise Ptc1, future work should focus on screening several higher protein concentrations, in particular for LCP. Unfortunately, there is no known rule of thumb which could guide the experimenter in choosing the appropriate

protein concentration for crystallisation in LCP. In fact, suitable concentrations are often protein rather than protein class dependent. For instance, LCP crystallisation of GPCRs has been set up over a wide concentration range of 20- 80 mg mL^{-1} [39, 252, 289]. Ptc1 was concentrated to 20- 54 mg mL^{-1}, concentrations that could possibly be increased further. In addition, other available lipids should be taken into consideration, in particular those that have been proven successful in crystallisation of larger MPs. 7.8 MAG, which was tested among 9.9 MAG and 7.7 MAG, was the only lipid able to clarify after mixing with Ptc1. A translucent sample is an indication for the formation of the lipidic phase and successful reconstitution of Ptc1 into the lipidic environment [156]. This experiment should be repeated, possibly with a higher protein or salt concentration. The presence of bRIL and preparation of the Ptc1-Nb64 or Shh-Ptc1 complex should have a positive effect on the crystallisability as these proteins introduce hydrophilic surfaces to the protein. Another reason that Ptc1 remained difficult to crystallise could lie in the potential monomer-dimer equilibrium. Further studies should pay attention to producing a sample of a purely monomeric or dimeric nature.

Exchange of the Ptc1-detergent complex into Amphipols paved the way for studying the protein by negative stain EM. An even protein distribution and a defined particle shape indicated that Ptc1 is suitable for cryo-EM studies. However, virtification of the sample introduced aggregation and denaturation of the protein. Future work could include analysis of a chemical crosslinked sample to reduce the risk of protein denaturation. Reconstitution of Ptc1 in nanodiscs is work currently in progress. Initial results indicated stacking of nanodiscs which might be due to incorporation of more than one copy of Ptc1 per nanodisc, possibly in opposite direction [66]. Hydrophobic interactions of the Ptc1 extracellular domain may lead to the formation of the observed nanodisc stacks. With the acquisition of a large enough

data set that accounts for sample heterogeneity, 2D classification should be able to separate the Ptc1 monomers and dimers, or in the case of nanodiscs separate particles with one or more Ptc1 molecules per disc [230]. A newer approach is to use megabodies, the next-generation nanobodies, for structural studies. Megabodies are four times the size of nanobodies and therefore introduce an even larger hydrophilic surface. In addition, they considerably increase the size of the protein, which is beneficial for EM studies. Due to the rather sparse intracellular landscape of Ptc1, a nanobody that targets bRIL would contribute to its intracellular surface and therefore help its identification in cryo-EM.

The structure of Ptc1 alone or in complex with Shh will give insights into the signal transduction mechanism and provide a platform for designing inhibitors to address diseases based on Hh signalling dysregulation.[5]

[5]During the preparation of this thesis, cryo-EM structures of Ptc1 have been published [101, 227, 228]. These findings will be discussed in Chapter 7.

Chapter 6

Structural studies of the Ptc1 extracellular domain

6.1 Abstract

The goal of this study was to shed light on the structural arrangement of the extracellular domains (ECDs) of the human Hedgehog (Hh) receptor Patched-1 (Ptc1). Firstly, we engineered a soluble ECD construct containing solely Ptc1 ECD1 and ECD2, connecting both domains via a flexible Gly/Ser (G/S)-rich linker (Ptc1 ECD1-ECD2). Functional relevance of this construct was demonstrated by its ability to bind its ligand Shh. Sample quality and quantity of Ptc1 ECD1-ECD2 both alone and in complex with Shh was suitable for crystallisation trials using the vapour diffusion method. The protein samples were further stabilised by nanobodies, small proteins that are known for their ability to increase the hydrophilic surface of their binding partner and therefore crystallisation probability. Secondly, we designed single ECD constructs and successfully co-crystallised Ptc1 ECD1 in complex with Nb64. *De novo* structure determination using sulphur SAD (S-SAD) phasing was carried out to a resolution of 2.8 Å, and further extended to 2.1 Å using the S-SAD structure for phasing of a new data set. The crystal structure presented is the first available structure of Ptc1 ECD1 at atomic resolution. The third complementarity determining region (CDR3) of Nb64 inserted into a protein cleft presented on the surface of Ptc1 ECD1. This interaction was further stabilised by non-bonded contacts contributed by the CDR1 of Nb64. In addition, a small molecule was identified in a hydrophobic core formed by three α-helices of Ptc1 ECD1. Co-crystallisation with the cholesterol derivative cholesteryl hemisuccinate (CHS) suggested that one of the natural Ptc1 ligands is a lipid. This ligand binding domain (LBD) might have a function in cholesterol transport and/or is responsible for Shh recognition as Shh is post-translationally modified by a palmitoyl- and cholesterol-moiety.

6.2 Introduction

6.2.1 Lipids and their implication in Hh signalling

First identified in biliary calculi (gallstones) by F. Poulletier de la Salle in the 18th century, cholesterol is a solid lipid involved in many biological processes. Cells either produce cholesterol on its own or extract it from external sources. Biosynthesis of cholesterol is a complicated process; many enzymes are involved in converting Acetyl-CoA into cholesterol. The synthesis of cholesterol is initiated either in the endoplasmatic reticulum (ER) or in peroxisomes and completed in the ER from where it is transported to the plasma membrane (PM) with the help of caveolin lipoprotein particles [29, 90]. At the PM, almost 90% of all cellular cholesterol accumulates, accounting for 20- 25% of the PM content [125]. There, it fulfills various functions, such as regulating fluidity, permeability, thickness, and curvature of the membrane. Cholesterol is not distributed equally in the PM; it is enriched in microdomains better known as lipid rafts. Receptors, ion channels and enzymes are associated with lipid rafts and cholesterol modulates their function by interacting directly with signalling domains or by regulating their microenvironment (membrane fluidity) [47, 99]. In this way, cholesterol and its derivatives (steroids, bile acids) participate in signal transduction and membrane trafficking. A prominent class of proteins that is found in cholesterol-rich microdomains is G protein-coupled receptors (GPCR) [130]. Smoothened (Smo), the GPCR involved in Hedgehog (Hh) signalling has been identified in lipid rafts and it has been shown that cholesterol is necessary for Smo activation [122, 173, 251]. Inhibition of cholesterol biosynthesis with inhibitors or trapping of membrane-localised cholesterol with cyclodextrin (Methyl-β-cyclodextrin, MβCD) impairs Smo signalling [57, 62]. By contrast, delivery of cholesterol in the form of MβCD:cholesterol is sufficient for localisation and activation of Smo in the PM; however, no accumulation of

Smo in the primary cilium was observed [122, 173]. For localisation of Smo in cilia, the presence of the Hh ligand (e.g. Shh) and its receptor Patched-1 (Ptc1) is required [73, 173]. In the absence of Shh, Ptc1 suppresses Smo activity in a sub-stoichiometrical manner; the inhibitory effect is most likely carried out by a small molecule [264]. Recent findings by E. Byrne *et al.* (2016) suggest that this small molecule is cholesterol and that it binds to the cysteine-rich domain (CRD) of Smo [39, 122]. Ptc1 has also been implicated in cholesterol binding and transport, and structural similarity of Ptc1 to Niemann-Pick C1 (NPC1), a protein involved in cholesterol homoeostasis, supports these results [25, 43, 170]. Furthermore, it has been found that mutations in the sterol-sensing domain (SSD) of Ptc1 cause accumulation of sterols in Ptc1 endosomes, an observation that suggests a role for Ptc1 in lipid trafficking [82, 261]. In this scenario, Ptc1 would target Smo for degradation or regulate its relocation to the PM by regulating the lipid content in (Ptc1-positive) endosomes. Once located in the PM, Smo is guided to the primary cilium by lateral transport [188]. The lipid is supplied by lipoprotein particles that are recruited by Ptc1 to Ptc1-positive endosomes [144]. Lipids also have another important function in Hh signalling; the Hh ligand requires post-translational modification with a cholesterol- and palmitoyl-moiety in order to be fully functional. After modification, lipoprotein particles regulate the movement of Hh through and between cells [209]. Furthermore, lipid modification of the Hh ligand permits integration into the PM and thus regulates its mobility [220, 232]. The precise mechanism by which Ptc1 regulates Smo and the exact function of cholesterol remains a black box.

Many cellular functions are dependent on a steady cholesterol supply [224]. Therefore, it does not come as a surprise that dysregulation of lipid metabolism has been implicated in a range of diseases. Lipid-associated disorders can be of genetic origin, or a result of irregular cholesterol supply from external sources [183]. Disorders

associated with aberrant cholesterol levels include Niemann-Pick disease (also known as lysosomal lipid storage disease), hypercholesterolaemia, Tangier disease, as well as neurodegenerative disease (dementia), cardiovascular disease, diabetes and cancer [124, 183]. Some cholesterol-related disorders have been directly linked to aberrant Hh signalling [62, 143]. For instance, Smith-Lemli-Opitz syndrome (SLOS) is an inborn developmental disorder caused by mutations in the 7-dehydrocholesterol reductase (*DHCR7*) gene [224]. As a result, 7-dehydrocholesterol can not be converted into cholesterol during the final stage of cholesterol biosynthesis (Figure 6.2, A). The decreasing cholesterol levels directly correlate with Hh signalling activity. The synthesis of dual-lipidated Shh as well as Smo activity and localisation are impaired (Figure 6.2, B) [28, 62, 109]. Diseases with similar implications are lathosterolosis and desmosterolosis (Figure 6.2, A). Dysfunction of enzymes that are responsible for the synthesis of lathosterol and 7-dehydro-cholesterol lead to lower cholesterol levels and manifest in cholesterol-related disorders.

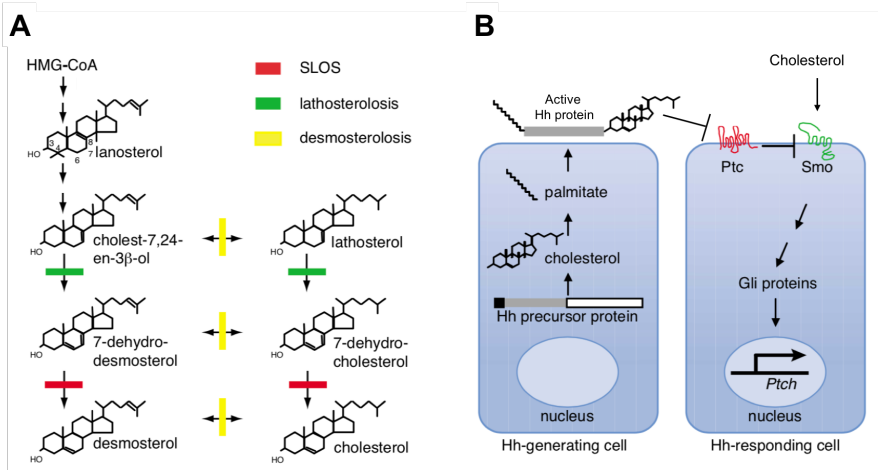


Figure 6.1: **The link between Hh signalling and cholesterol-derived disease. A**, Final stages of cholesterol biosynthesis. Mutations in enzymes involved in cholesterol precursor processing lead to diseases such as the Smith-Lemli-Opitz syndrome (SLOS), lathosterolosis, and desmosterolosis. **B**, Overview of Hh signalling. Hh-generating cells secrete dual-lipidated Hh protein, which binds to and inactivates Ptc1 on Hh-responding cells. In turn, the Ptc1 inhibition of Smo is released, leading to the activation of Hh target genes. Smo activity is further regulated by cholesterol and/or its derivatives. Figure adapted from [62].

6.3 Results

6.3.1 Ptc1 ECD1 & ECD2 can be expressed as a fusion protein

In 1996, V. Marigo and co-workers identified Ptc1 as the Hh signalling receptor. Their results revealed that Hh ligand-binding is achieved by the extracellular domains (ECDs) [179]. The goal of our study was to shed light on the structural arrangement of ECD1 and ECD2 by determining the structure by X-ray crystallography. Based on a previously published AcrB structure, a Ptc1 ECD construct was designed in which the ECDs form a structural unit. For this, the ECDs were interlinked via a flexible Gly/Ser (G/S)-rich linker (Ptc1 ECD1-ECD2).

Sequences were cloned into a pHLsec vector with a C-terminal AviTag^{TM}, followed by a Rho1D4 antibody epitope tag (1D4-tag) for antibody-affinity purification [12]. HEK293T cells were transiently transfected with the plasmid and the conditioned media was collected 60 hours post-transfection. The protein was purified by Rho1D4 antibody affinity chromatography and subsequently subjected to SEC using a Superdex 200 16/60 column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl. SEC profiles displayed two main protein peaks with peak maxima at ∼55 mL and 80 mL retention volume (Figure 6.2, A). Despite severe protein aggregation, an adequate amount of potential monomeric Ptc1 ECD1-ECD2 was obtained. The shoulder on the left side of the main peak might originate from a dimeric or multimeric form of Ptc1 ECD1-ECD2. In order to shed light on the oligomeric state of the protein, multi-angle light scattering (MALS) experiments were carried out. SDS-PAGE analysis revealed a band at ∼85 kDa, slightly higher than the expected molecular weight of 67.9 kDa (Figure 6.2, B). The presence of six asparagine (N)-linked glycosylation sites, incomplete protein denaturation, or simply repulsive

forces between acidic amino acid residues and SDS might have led to this observation [104]. The sample displayed no impurities or degradation and western blot analysis confirmed the stability of the sample (Figure 6.2, C). The analysis also revealed that most of the protein remained in the tissue culture medium and did not bind to the Rho1D4 antibody-coupled beads during the first step of affinity purification. Increasing the bead volume did not result in a higher protein yield, suggesting that availability of antibody binding sites was not the limiting factor. Alternatively, the protein might have formed higher molecular weight complexes in which the 1D4-tag was hidden and therefore not accessible for interaction with the Rho1D4 antibody. SEC peak fractions were pooled and concentrated to 5 mg mL^{-1} for setting up crystallisation using the vapour diffusion (VD) method. N-linked glycosylation sites were trimmed down to one N-acetylglycosamine residue by endoglycosidase F1 (EndoF1).

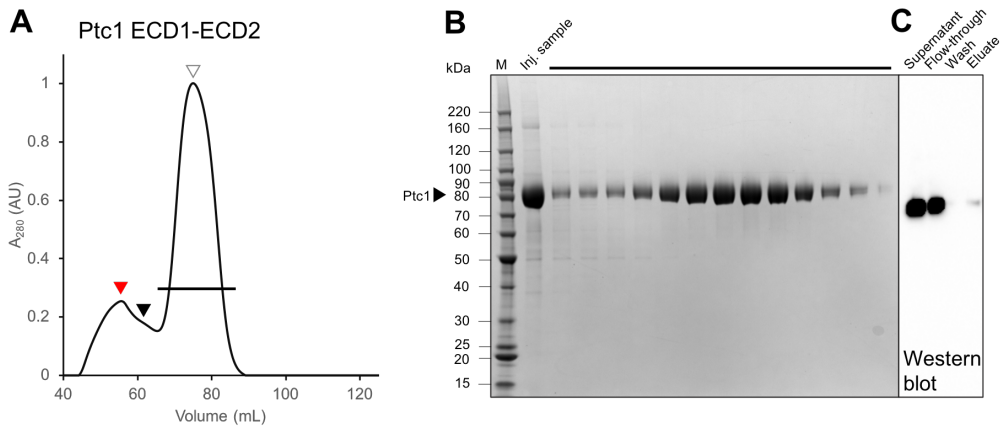


Figure 6.2: **Purification and analysis of Ptc1 ECD1-ECD2.** SEC profile (A) with corresponding SDS-PAGE (B) and western blot (C). **A**, Red arrowhead indicates the peak of aggregated protein species, black outlined arrowhead the peak of monomeric Ptc1. The peak in between most likely originated from Ptc1 dimer (black arrowhead). **B**, Lanes from left to right: BenchMark^{TM} ladder (M), sample subjected to SEC (Inj. sample). SEC fractions across the Ptc1 peak fraction are indicated by a black line. **C**, Purification steps analysed on western blot using Rho1D4 antibody for visualisation. Lanes 1-4: affinity purification steps of the protein.

The Ptc1 ECDs might exhibit highly flexible regions, which may be stabilised in the presence of the Hh ligand. Consequently, the protein was co-purified with Shh using the same strategy as presented in the previous study. The SEC profile revealed a slight shift in peak position to ∼78 mL retention volume, indicating that the protein complex was formed (Figure 6.3, A). Furthermore, this observation suggest that the designed Ptc1 ECD construct is able to provide a platform which is sufficient for Shh binding. As Shh was added in a two-fold molar excess, an additional peak was observed at ∼100 mL retention volume. Similar to the apo protein, aggregation and a potential dimeric version of Ptc1 ECD1-ECD2 were detected. The SEC peak fractions of monomeric Ptc1 ECD1-ECD2 were pooled, avoiding the fractions on the left side of the peak maximum to not contaminate the sample with aggregated protein. The complex was concentrated to 4.5 mg mL^{-1} and treated with EndoF1 before crystallisation using the VD method was set up.

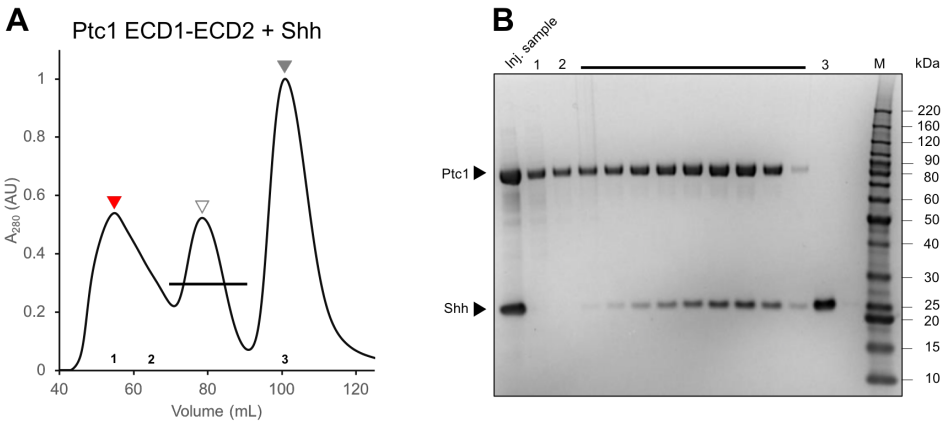


Figure 6.3: **Purification and analysis of the Shh-Ptc1 ECD1-ECD2 complex.** SEC profile (A) with corresponding SDS-PAGE (B). **A**, Red arrowhead indicates the peak of aggregated protein species, black outlined arrowhead the peak of monomeric Ptc1, and grey arrowhead indicates the peak of Shh. **B**, Lanes from left to right: Sample subjected to SEC (Inj. sample), SEC fractions 1-3 as indicated on the SEC profile (A), BenchMark^{TM} ladder (M). SEC fractions across the Ptc1 peak fraction are indicated by a black line. These experiments were performed in collaboration with A. Rudolf.

Nanobodies introduce additional hydrophilic properties to proteins, and have been found to enhance crystallisability of their binding partners. In most cases, nanobodies also stabilise their antigen by inserting (at least) one of their hypervariable loops into protein clefts. Therefore it was decided to raise nanobodies against the Shh-Ptc1 complex. For nanobody generation, camelids were immunised with the Shh-Ptc1 ECD1-ECD2 complex (Chapter 3). Nanobodies were expressed and purified from the periplasm of *E. coli* WK6 Su^- cells. Bio-layer interferometry (BLI) was used to identify Ptc1-specific nanobodies (Chapter 3). Before setting up crystallisation trials in VD, Ptc1-nanobody complexes were subjected to SEC using a Superdex 200 16/60 column (GE Healthcare) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl. Representative SEC traces of Ptc1 ECD1-ECD2 in complex with Nb64 or Nb65 are shown in Figure 6.4, A and Figure 6.5, A, respectively. Both protein complexes displayed three overlapping peaks, followed by a peak at ∼102 mL (Nb64) or ∼107 mL (Nb65) retention volume. Similar to studies with Shh, protein aggregation and a peak for potential multimeric Ptc1 ECD1-ECD2 was observed. SDS-PAGE analysis confirmed sample purity and the formation of a Ptc1-Nb complex in both cases (Figure 6.4, B and Figure 6.4, B). SEC peak fractions of the main protein peak were pooled and concentrated to 4.5- 5.5 mg mL^{-1}, followed by EndoF1 treatment. Nb64 and Nb65 complexes, as well as Nb67, Nb68, and Nb75 complexes were screened in this study.

In order to identify nanobodies that bind to the Shh-Ptc1 ECD-ECD2 complex, bio-layer interferometry (BLI) was carried (Chapter 3). The nanobodies were then used to prepare a trimeric complex, consisting of Ptc1 ECD1-ECD2, Shh and nanobody. After formation of the heterotrimer, the protein was purified via SEC as described for the Ptc1-Nb complexes. In comparison to the Ptc1-Nb complexes, less aggregation and a slight shift in peak position was observed, suggesting that

the trimeric complex was formed (Figure 6.6, A). Furthermore, SDS-PAGE analysis confirmed sample purity (Figure 6.6, B). The SEC peak fractions were pooled and concentrated to 4.6- 5 mg mL^{-1}. In addition to Ptc1-Nb66 complexes, Ptc1-Nb804 complexes were prepared and subjected to co-crystallisation using the VD method.

Ptc1 ECD1-ECD2 and its complexes were tested against 13 different crystallisation screens, each containing 96 unique conditions and covering almost 1250 conditions in total. A wide range of pH, PEGs, salts, and additives were screened. Crystallisation trials were set up at two different temperatures (6.5°C and 20°C) and at 100% and 50% precipitate concentrations. Drop ratios chosen were 1:1 (protein:precipitant solution). As of today, no crystals have been observed.

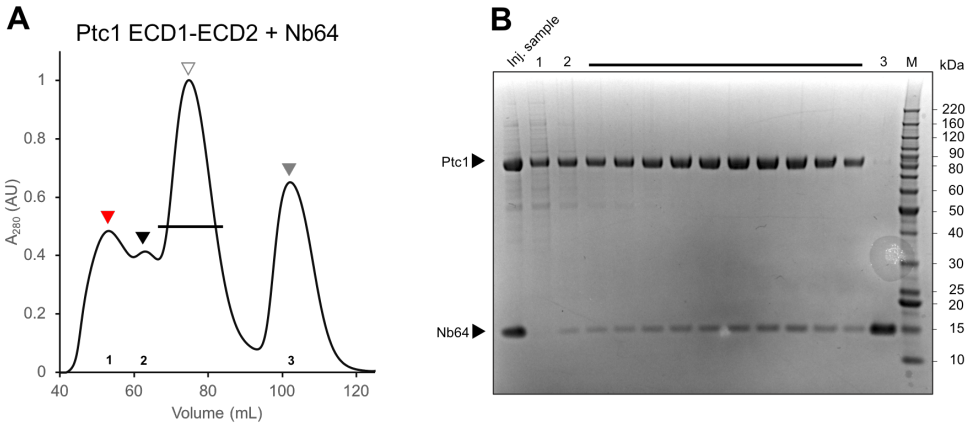


Figure 6.4: **Purification and analysis of Ptc1 ECD1-ECD2 in complex with Nb64. A-B**, SEC profile (A) and corresponding SDS-PAGE analysis (B) of Ptc1 ECD1-ECD2 in complex with Nb64. **A**, Red arrowhead indicates the peak of aggregates, black and black outlined arrowhead the peak of potential dimeric and monomeric Ptc1 species, respectively. Grey arrowhead indicates excess Nb64. SEC fractions analysed on SDS-PAGE are marked by numbers and a black line. **B**, SDS-PAGE analysis confirmed the presence of Ptc1 ECD1-ECD2 and Nb64; co-elution is an indicator for complex formation. Lanes from left to right: Sample subjected to SEC (Inj. sample), aggregated protein species (1), dimeric Ptc1 ECD1-ECD2 (2), Nb64 (3), BenchMark^{TM} ladder (M). SEC fractions across the monomeric peak fraction are indicated by a black line.

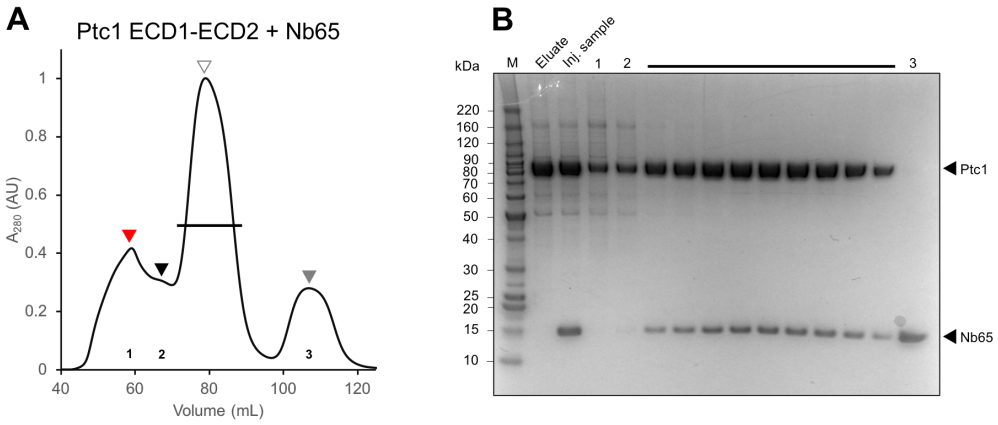


Figure 6.5: **Purification and analysis of Ptc1 ECD1-ECD2 in complex with Nb65. A-B**, SEC profile (A) and corresponding SDS-PAGE (B) of Ptc1 ECD1-ECD2 in complex with Nb65. **A**, Red arrowhead indicates the peak of aggregates, black and black outlined arrowhead the peak of potential dimeric and monomeric Ptc1, respectively. Grey arrowhead indicates excess Nb65. **B**, SDS-PAGE analysis confirmed successful formation of the complex. Lanes from left to right: BenchMark^{TM} ladder (M), eluate from affinity purification, sample subjected to SEC (Inj. sample), aggregated protein species (1), dimeric Ptc1 ECD1-ECD2 (2), Nb65 (3). SEC fractions across the monomeric peak fraction are indicated by a black line.

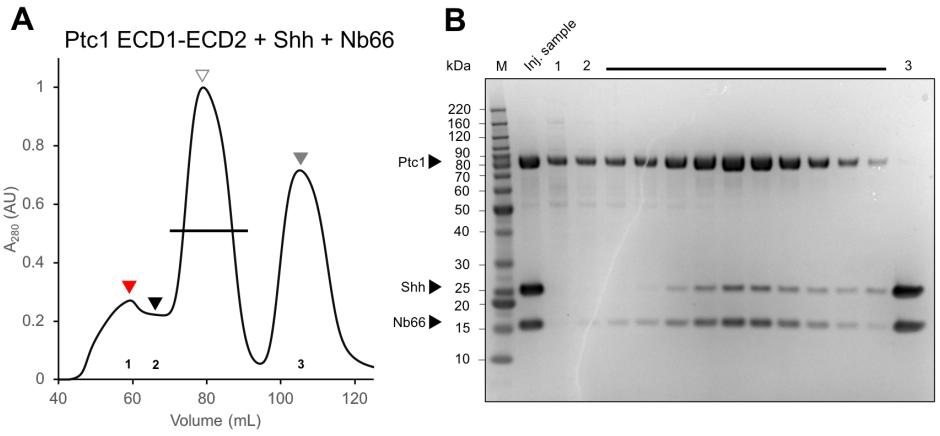


Figure 6.6: **Purification and analysis of Ptc1 ECD1-ECD2 in complex with Shh and Nb66. A-B**, SEC profile (A) and corresponding SDS-PAGE analysis (B). **A**, Red arrowhead indicates the peak of aggregates, black and black outlined arrowhead the peak of potential dimeric and monomeric Ptc1 species, respectively, grey arrowhead indicates excess Nb66. **B**, SDS-PAGE analysis confirmed the presence of Ptc1 ECD1-ECD2, Shh and Nb66 and the formation of a trimeric complex. Lanes from left to right: BenchMark^{TM} ladder (M), sample subjected to SEC (Inj. sample), aggregated protein species (1), dimeric Ptc1 ECD1-ECD2 (2), Shh and Nb66 (3). SEC fractions across the monomeric peak fraction are indicated by a black line.

6.3.2 Ptc1 ECD1 crystallises in complex with Nb64

Despite extensive screening of the Ptc1 ECD1-ECD2 construct with and without ligands, crystallisation has not proven successful yet. For this reason, it was decided to study the ECDs separately, with initial focus on ECD1. Ptc1 ECD1 was cloned into the pHLsec vector with C-terminal His_6-tag and expression was carried out in HEK293T cells. After metal affinity chromatography, Nb64 was added to the sample and the protein complex was subjected to SEC coupled to a Superdex 75 16/60 column equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl. The sample eluted as a single monodisperse peak and only minor protein aggregation was observed. SDS-PAGE revealed three bands, corresponding to Ptc1 ECD1, Nb64 and bovine serum albumin (BSA). Co-elution of Ptc1 ECD1 and Nb64 indicated that the complex was formed (Figure 6.7, A-B). The BSA contamination most likely originated from the expression medium, which contained a high concentration of serum. Unspecific binding of BSA to the Ni-NTA resin is a commonly observed phenomenon. However, this does not necessarily interfere with protein crystallisation. SEC peak fractions were pooled and crystallisation was set up at 6.7 mg mL^{-1} after EndoF1 treatment. Crystals were readily observed across various crystallisation conditions (Figure 6.7, C).

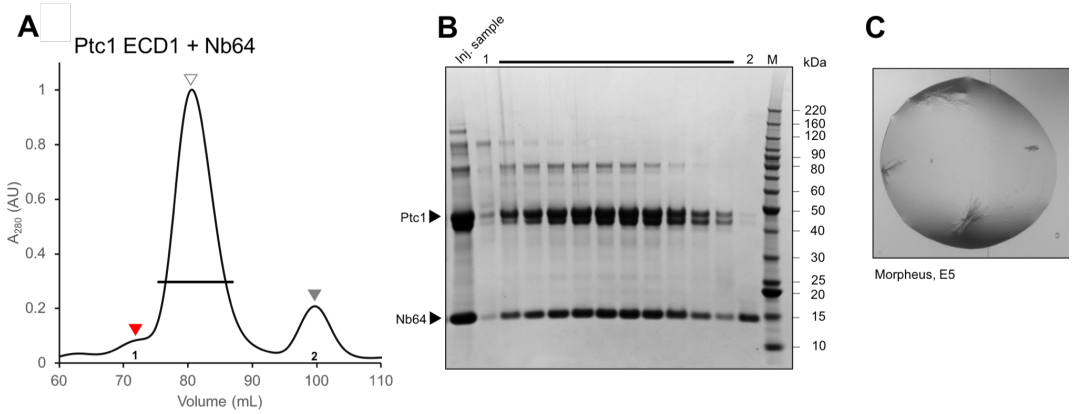


Figure 6.7: **Purification and analysis of the Ptc1 ECD1-Nb64 complex. A-B**, SEC profile (A) and corresponding SDS-PAGE (B). **A**, Ptc1 ECD1-Nb64 eluted as a single monodisperse peak. Red arrowhead indicates the peak of aggregates, black outlined arrowhead Ptc1 ECD1-Nb64, grey arrowhead indicates excess Nb64. **B**, Lanes from left to right: Sample subjected to SEC (Inj. sample), aggregated protein species (1), Nb64 (2), BenchMark^{TM} ladder (M). SEC fractions across the monomeric peak fraction are indicated by a black line. **C**, Crystals of the Ptc1 ECD1-Nb64 complex obtained in Morpheus (Molecular Dimensions) condition E5 after optimisation by microseeding. These experiments were performed in collaboration with A. Rudolf.

A powerful method for *de novo* structure determination of proteins is single-wavelength anomalous dispersion (SAD). SAD can be used in combination with crystals that have been treated with heavy atoms, selenomethionine-labelled protein, or protein with a sufficient number of intrinsic sulphur (S) atoms originating from cysteines (Cys, C) and methionine (Met, M). The signal obtained from these atoms can then be used for *de novo* experimental phasing. The Ptc1 ECD1-Nb64 complex contains a total of 21 S atoms (Figure 6.8, B); 17 S atoms are found in Ptc1 (6x Cys, 11x Met) and four in Nb64 (2x Cys, 2x Met). In order to determine the structure of the Ptc1 ECD1-Nb64 complex, anomalous signal of S atoms was acquired using the long-wavelength beamline I23 at DLS (Didcot, UK), operating at a wavelength of ∼2.78 Å. I23 beamline scientist K. El Omari assisted with the sulphur SAD (S-SAD) data collection and *de novo* structure determination. The anomalous signal of the S atoms extended to 4 Å and the positions of 14 S atoms could be determined. Initial electron density maps were calculated to 2.8 Å and iterative cycles of manual

model building combined with phase and density improvement was necessary to obtain an initial protein model. Figure 6.8, A displays the initial model overlaid with an anomalous difference map contoured at 3.7 σ. A native data set collected on beamline I03 at DLS (Didcot, UK) allowed phase extension to 2.1 Å resolution (Figure 6.9). Statistics of the S-SAD and high-resolution native data collection, as well as details on the structural refinement can be found in Table 6.1.

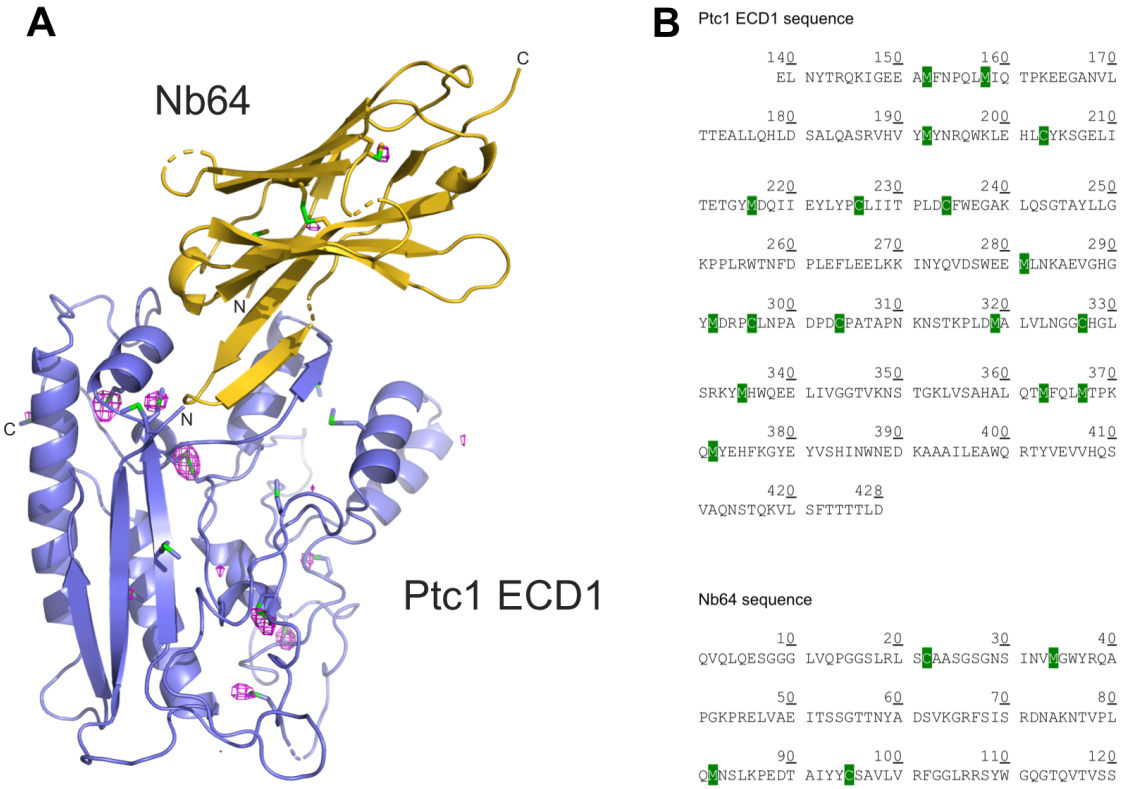


Figure 6.8: **S-SAD phased crystal structure of the Ptc1 ECD1-Nb64 complex. A**, Cartoon representation of Ptc1 ECD1 (slate) in complex with Nb64 (yellow) with anomalous difference map at 3.7σ (magenta chicken-wire). N- and C-terminus are indicated. Sulphur atoms are highlighted in green. S-SAD data collection and *de novo* structure determination was performed in collaboration with K. El Omari. **B**, Locations of cysteines and methionines marked on the sequence of Ptc1 ECD1 (top) and Nb64 (bottom).

The globular fold of Ptc1 ECD1 is formed by 7 α-helices and 8 β-strands, which are interconnected by a network of loops (Figure 6.9, A). Two motifs are apparent. Firstly, the N- and C-terminus of the protein assemble to form a sandwich-like scaffold; three β-strands (β1, β7, β8) fold against two α-helices (α1, α7). α6 lies on top of this cylindrical motif. Secondly, an α-helical domain consisting of three α-helices (α2, α4, α5) lies opposite of this motif. A β-strand (β4) and an extended loop region support this helical core. Interestingly, the Ptc1 homologue Niemann-Pick C1 (NPC1) protein does not exhibit a α-helical domain, suggesting that this particular fold is unique for Ptc1 and that it has an important function (Figure 6.9, B-C). Nb64 adopts a typical nanobody scaffold, comprised of nine β-strands. Two motifs of four and five β-strands, respectively, form two larger β-sheets, which lead to a classical β-sheet sandwich topology. The β-strands are interspersed by three complementary determining regions (CDRs), located between β2- and β3-strands (CDR1), β4- and β5-strands (CDR2), as well as the β8- and β9-strands (CDR3). The CDR3 is formed by a long β-hairpin, which inserts into a protein cleft exhibited on Ptc1 ECD1. In addition, a disulfide bond which is commonly observed among nanobodies, is located between Cys^{22} on β2 and Cys^{95} on β8 (Figure 6.8, A). The N-terminal tail of the nanobody packs against the loop of Ptc1 ECD1 following α6 (Figure 6.10).

A closer look into the interaction surface between Ptc1 and Nb64 revealed a total of ~190 non-covalent interactions (Figure 6.10). Whereas the CDR1 of Nb64 solely displayed hydrophobic interactions, the CDR3 additionally formed hydrogen bonds and salt bridges with Ptc1 ECD1. Six residues of the CDR3 (Arg^{101}, Phe^{102}, Gly^{104}, Ser^{105}, Arg^{106}, Arg^{107}) participated in the formation of 13 hydrogen bonds to Ptc1 ECD1 (Glu^{150}, Lys^{205}, Gly^{207}, Leu^{209}, Thr^{211}, Glu^{221}, Tyr^{224}, His^{375}). Salt bridges were observed between CDR3 residues Arg^{101}, Arg^{106}, Arg^{107} and Ptc1 ECD1 residues Glu^{150}, Glu^{208}, Asp^{217}, Glu^{221}. The formation of hydrophobic interactions

involved a total number of 16 and 21 residues of Nb64 (CDR1 and CDR3) and Ptc1 ECD1, respectively. Overall, the interface area of Ptc1 ECD1 covered 882 Å^2 whereas the nanobody used a slightly larger binding interface of 982 Å^2.

	Ptc1 ECD1-Nb64 (S-SAD)	**Ptc1 ECD1-Nb64**	**Ptc1 ECD1-Nb64:CHS**
DATA COLLECTION			
Beamline	DLS-I23	DLS-I03	DLS-I24
Wavelength [Å]	2.7500	0.97625	0.9686
Resolution [Å]	75.18-2.79 (2.94-2.79)	150.54-2.10 (2.14-2.10)	37.19-1.90 (1.95-1.90)
Space group	P212121	P212121	P212121
Cell dimensions	a = 48.1 Å b = 67.2 Å c = 150.4 Å α = 90° β = 90° γ = 90°	a = 47.7 Å b = 65.9 Å c = 150.5 Å α = 90° β = 90° γ = 90°	a = 47.3 Å b = 65.7 Å c = 151.4 Å α = 90° β = 90° γ = 90°
Unique reflections	11497 (1387)	27821 (1068)	37866 (2618)
Completeness [%]	90.9 (77.1)	97.5 (75.5)	99.4 (94.9)
R_{merge} [%]	11.6 (>100)	6.6 (>100)	11.1 (>100)
R_{pim} [%]	3.1 (81.7)	2.7 (69.6)	5.3 (86.7)
$CC_{1/2}$	0.999 (0.462)	0.998 (0.521)	0.998 (0.360)
I/σI	22.2 (1.1)	16.2 (1.7)	9.1 (0.8)
Multiplicity	25.9 (24.2)	11.8 (7.8)	5.2 (4.2)
REFINEMENT			
Resolution range [Å]		50.00-2.10 (2.17-2.10)	37.19-1.90 (1.95-1.90)
Number of reflections (test set)		25995 (1423)	37775 (2584)
R_{work} [%]		20.7 (37.1)	17.9 (35.1)
R_{free} [%]		25.1 (35.8)	22.1 (38.7)
Number of atoms Protein Ligand Water		 3150 42 84	 3159 77 212
B factors [Å^2] Protein Ligand Water		 63 104 53	 43 83 43
RMSD bonds [Å]		0.015	0.011
RMSD angles [deg]		1.329	1.127
Ramachandran statistics [%] Favoured Disallowed		 96.4 0	 98.2 0

Table 6.1: **Crystallographic table for Ptc1 ECD1-Nb64.** Overview of data collection and refinement statistics. Ptc1 ECD1-Nb64 complexes were solved to a resolution of 2.8 Å (S-SAD), 2.1 Å (native) and 1.9 Å (native, in complex with CHS). Numbers in parentheses refer to the highest resolution shell. R_{free} equals the R-factor against 5% of the data. RMSD, root-mean-square deviation from ideal geometry; R_{pim}, precision indicating merging R-factor [286]; $CC_{1/2}$, correlation coefficient between random half data sets [141]. Ramachandran statistics were calculated using MolProbity [51].

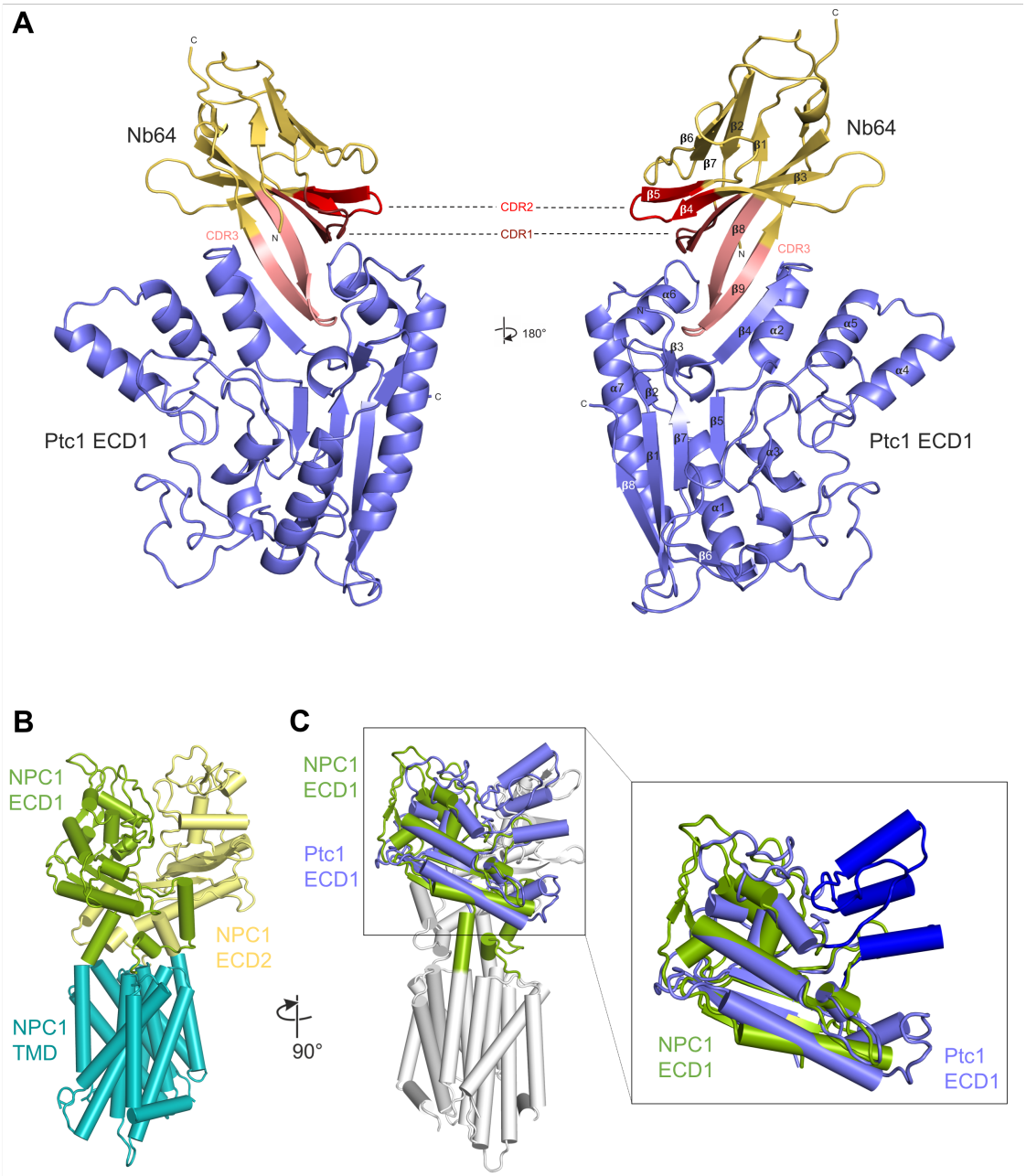


Figure 6.9: **Comparison of the Ptc1 ECD1 and NPC1 crystal structures.** **A**, Cartoon representation of the Ptc1 ECD1-Nb64 crystal structure, shown 180° rotated along the y-axis. Ptc1 ECD1 is represented in slate, Nb64 in yellow. CDRs are marked in ruby (CDR1), red (CDR2), and salmon (CDR3). **B**, Cartoon representation of the human Niemann-Pick C1 (NPC1) protein. NPC1 TMD (deepteal), NPC1 ECD1 (splitpea), NPC1 ECD2 (lightyellow). **C**, Superimposition of Ptc1 ECD1 on the NPC1 full-length structure, with close-up view of the NPC1 ECD1 and Ptc1 ECD1 domains. NPC1 TMD and NPC1 ECD2 are shown in white, NPC1 ECD1 in splitpea, and Ptc1 ECD1 in slate and blue.

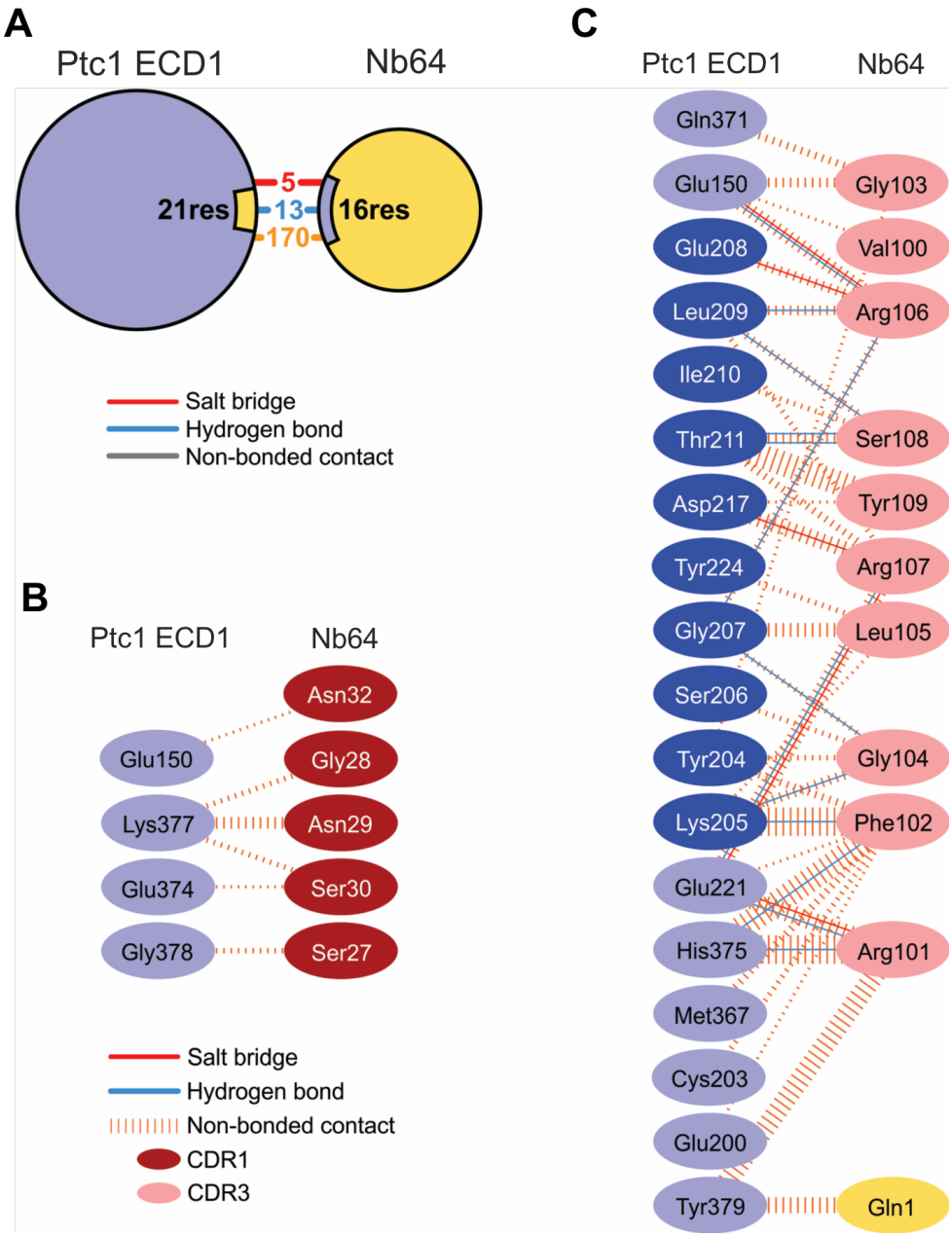


Figure 6.10: **Detailed overview of Ptc1 ECD1-Nb64 interface residues. A-C**, The interface between Ptc1 ECD1 and Nb64 is formed by 5 salt bridges (red line), 13 hydrogen bonds (blue line), and 170 non-bonded contacts (orange vertical dashed line). Figure adapted from PDBsum [157]. **A**, A total of 21 residues of Ptc1 and 16 residues on Nb64 are involved in interface formation. **B**, Contacts formed between Ptc1 ECD1 (slate) and Nb64 CDR1 (ruby). **C**, Interface residues of Ptc1 ECD1 (slate) with LBD residues (blue) and Nb64 CDR3 (salmon) and N-terminal residue Gln^1 (yellow). Figure adapted from [157].

6.3.3 Cholesterol binds to Ptc1 ECD1

During iterative refinement rounds on Ptc1 ECD1, additional electron density appeared in a hydrophobic core formed by three α-helices. The density could not be assigned to the protein itself, which indicated that a small molecule was bound to ECD1. Based on the theory that Ptc1 is involved in cholesterol transport, it was decided to co-crystallise Ptc1 ECD1 with cholesteryl hemisuccinate (CHS). CHS is structurally similar to cholesterol, but exhibits higher solubility due to a hydrophilic succinic acid modification.

X-ray data for the Ptc1 ECD1-Nb64:CHS complex were collected on the microfocus beamline I24 at DLS (Didcot, UK). The crystal diffracted to 1.9 Å and the structure was determined by molecular replacement using the Ptc1 ECD1-Nb64 apo structure (Chapter 6.3.2) as a template. Figure 6.11, A displays the structure of Ptc1 ECD1-Nb64 in complex with CHS (Ptc1 ECD1-Nb64:CHS). Data collection and refinement statistics are shown in Table 6.1. Overall, the structure adopted the same fold as described in Chapter 6.3.2. Extra electron density was observed immediately in the map, however this time, the density was more defined and the four hydrocarbon rings of CHS fitted the density nicely (Figure 6.11, C). Moreover, the density extended towards the inner core of the protein and very likely originates from the aliphatic carbon tail attached to C_{17} of CHS (Figure 6.11, A). Nevertheless, it cannot be entirely ruled out that CHS occupies the pocket the other way around. However, the succinate ester on the β-hydroxyl group exhibits higher hydrophilic properties, and thus might not occupy the inner part of the hydrophobic pocket. Figure 6.11, D illustrates the ligand binding domain (LBD) with the closest interacting residues.

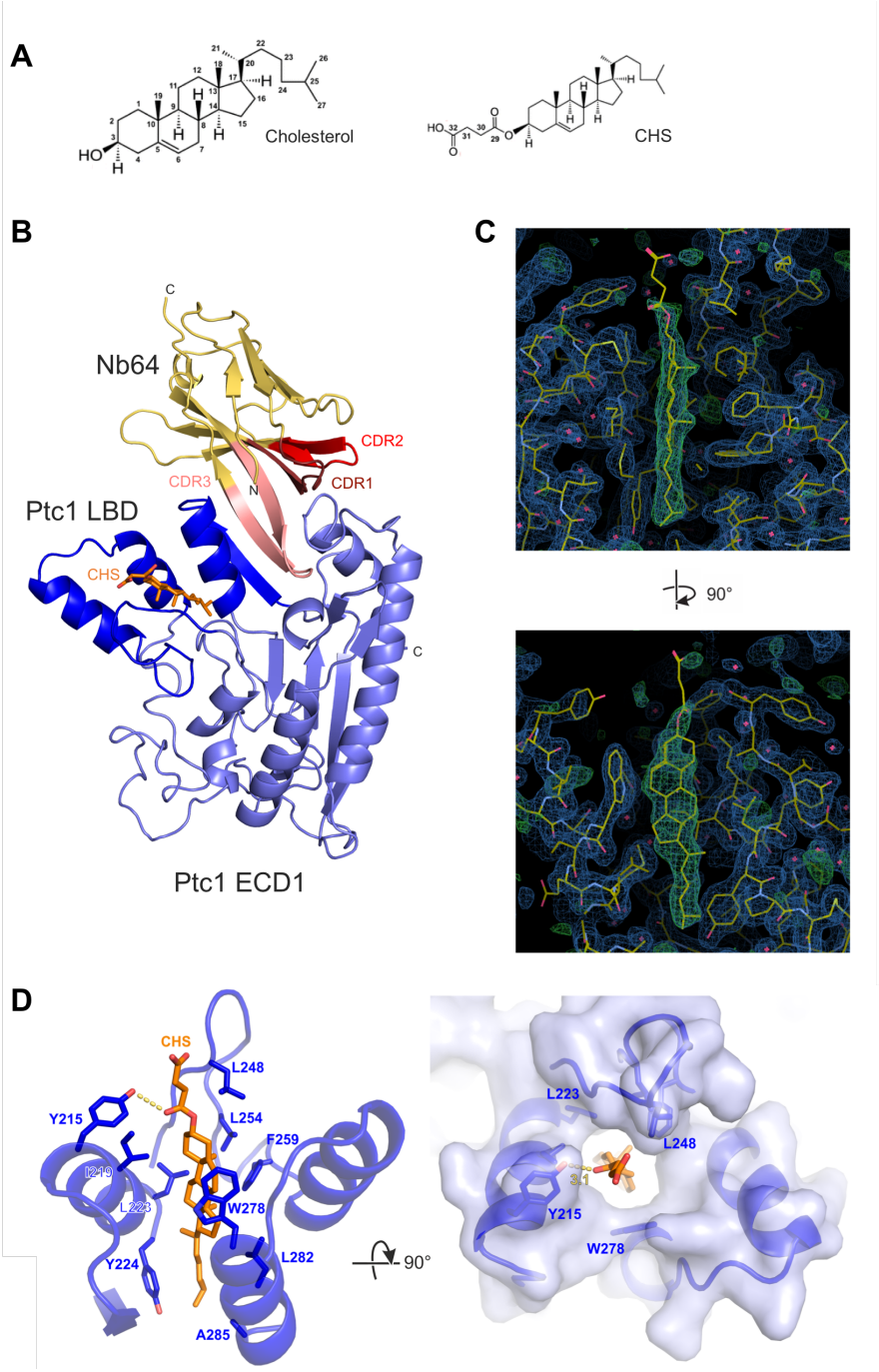


Figure 6.11: **Crystal structure of the Ptc1 ECD1-Nb64:CHS complex. A**, Chemical structures of cholesterol (left) and CHS (right). Carbon positions are indicated by numbers. Figure adapted from [153]. **B**, Cartoon representation of the Ptc1 ECD1-Nb64:CHS complex. Ptc1 ECD1 (slate), Nb64 (yellow, with CDR1 [ruby], CDR2 [red], CDR3 [salmon]), and CHS (orange). N- and C-terminus is indicated by N and C, respectively. CHS (sticks representation) is positioned in between the LBD (blue). **C**, Electron density of CHS and surrounding residues, shown rotated 90° along the y-axis. Overlay of the $2F_{obs}$-F_{calc} map at 1.0 σ and the F_{obs}-F_{calc} map at 3.0 σ shown in blue and green chicken-wire, respectively. Carbon atoms are coloured in yellow, oxygen atoms in red, and nitrogen atoms in blue. Water molecules are indicated by red crosses. **D**, Close-up view of the interface between Ptc1 LBD (blue) and CHS (orange) with interacting residues as sticks. Oxygen atoms are shown in red, the hydrogen bond as a yellow dotted line.

The LBD is formed by three α-helices, specifically $\alpha 2$, $\alpha 4$ and $\alpha 5$ (Figure 6.11, D). This motif was further stabilised by $\beta 4$ and an extended loop region. A total of 13 residues were involved in positioning CHS inside the pocket. The part resembling cholesterol formed 26 hydrophobic interactions with 11 residues of Ptc1 ECD1 (Figure 6.12). This arrangement was stabilised by two potential hydrophobic stacking interactions between the hydrophobic carbon rings of CHS and two aromatic residues of Ptc1 (F^{264} and W^{278}). Two hydrogen bonds were observed between the succinate ester and Ptc1 residues Tyr^{215} and Leu^{249}.

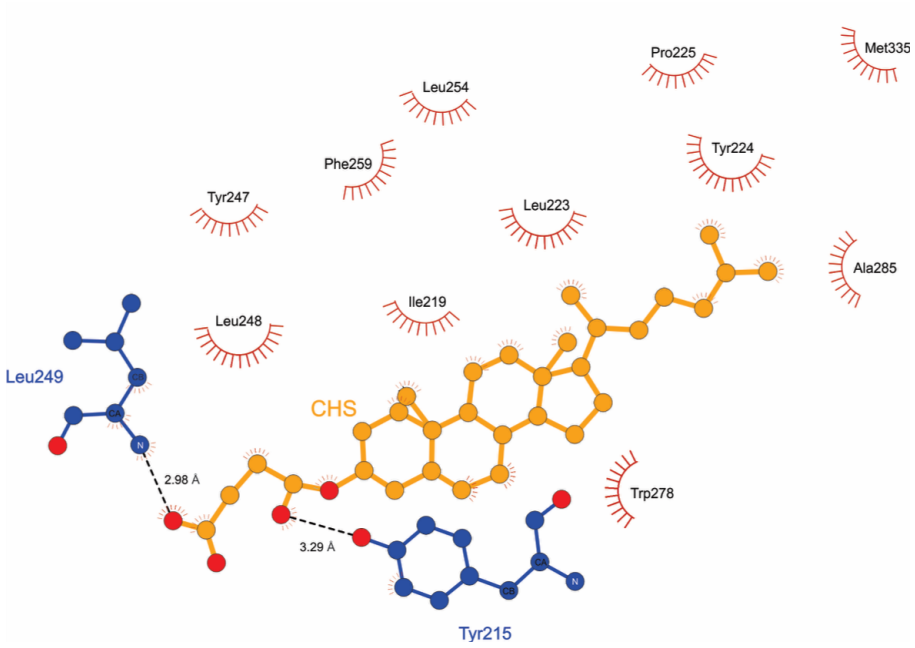


Figure 6.12: **Ptc1 ECD1 interactions involving CHS.** Schematic diagram of the interaction between Ptc1 ECD1 and CHS. Intermolecular interactions are formed between 13 residues on Ptc1 ECD1 and 16 atoms of CHS. Hydrogen bonds are shown as black dotted lines, non-bonded contacts are indicated by red spikes. Oxygen atoms are shown in red, Nitrogen atoms are indicated by N. Ligplot figure adapted from PDBsum [157, 284].

6.4 Discussion

Cell signalling pathways are formed by a complex network of interacting molecules. Signal transduction is activated once a ligand such as a morphogen, hormone, or neurotransmitter binds to its dedicated cell surface receptor. This leads to the activation of downstream signalling events which inevitably promote, or inhibit, cellular processes. In many cases, extracellular domains (ECDs) are responsible for ligand binding. In order to decipher the molecular mechanism of receptor-ligand binding by structural biology, ECDs are often studied separately. This fact is reflected in the growing number of crystal structures of receptor ECDs published each year. It is often less time consuming to work with soluble domains and it is often possible to achieve a higher resolution.

The aim of this study was to shed light on the structural arrangement of the Patched-1 (Ptc1) ECDs and the molecular mechanism of Hedgehog (Hh) signal reception. We successfully designed and produced a secreted version of the human Ptc1 ECDs as a fusion construct (Ptc1 ECD1-ECD2) in mammalian cells. From three litres of HEK293T cell expression, sufficient amounts of the protein were obtained to set up crystallisation trials using the vapour diffusion method. Despite extensive screening of the construct against a wide range of crystallisation conditions, the protein remained difficult to crystallise both by itself and in complex with Shh and/or nanobodies. Intrinsic flexibility in one or both domains might have impeded crystallisation of the construct. It is also unclear whether ECD1 and ECD2 have the ability to interact in the absence of Shh; insufficient or no contacts may have been formed between the two domains. Whilst ECD1 included almost all predicted residues to form the domain (Asn^{120}-Ser^{435} [predicted by the UniProt server: Glu^{122}-Asp^{436}]), ECD2 was ∼20 amino acids shorter (Asp^{789}-Gly^{1023} [Uniprot: Arg^{770}-Trp^{1027}]) [18].

The reason for this is that the Ptc1 ECD1-ECD2 construct was designed based on the structure of AcrB determined by S. Murakami *et al.* [195]. However, for Ptc1 those missing residues might be important for the interaction between the two ECDs and hence stabilisation of the construct. Future studies should therefore focus on the design of a Ptc1 ECD fusion construct that includes more residues of ECD2. In addition, co-crystallisation with mono- or dual-lipidated Shh could help facilitate crystallisation. For instance it was shown that palmitoylated Shh is ∼30 times more active in *in vitro* assays than its non-palmitoylated equivalent, and that the palmitate modification is essential for Hh signalling activation [217, 272]. ECD1 and/or ECD2 might form an anchor point for the cholesterol- or palmitoyl-moiety and therefore stabilise the protein fold.

We next focused on the single domains ECD1 and ECD2. Fortunately, ECD1 readily crystallised in complex with Nb64. This fact immediately adds to the hypothesis that flexibility of the fusion construct hampered crystallisation. However, the structure determination of Ptc1 ECD1 was not straightforward. Due to the lack of structural information of Ptc1 homologous and the small size of the nanobody, molecular replacement could not be employed. Therefore, single-wavelength anomalous dispersion of S atoms (S-SAD) phasing was carried out to determine the initial structural model of the Ptc1 ECD1-Nb64 complex. After crystal optimisation, this model could then be used to improve the resolution of the structure from 2.8 Å to 2.1 Å. This approach led to the first high resolution structure of Ptc1 ECD1. Nb64 formed an extensive interaction network on Ptc1 ECD1, mainly via the third hypervariable region; the β-hairpin inserted into a protein cleft exhibited on the Ptc1 ECD1. CDR1 formed additional contacts to stabilise the fold.

During the structural analysis, several interesting questions arose. For instance whether Shh and Nb64 use a similar binding interface on Ptc1 ECD1. Interestingly, Nb64 was not identified as an interaction partner of the Shh-Ptc1 complex in our BLI experiments (Chapter 3.3.3.3). This strongly suggests that Shh and Nb64 partly share the Ptc1 ECD1 binding interface. Notably, Nb65 as well as Nb66 and Nb67 were still able to bind to the Shh-Ptc1 complex in our binding studies, despite their CDR3 being identical to that of Nb64. Further sequence analysis of these nanobodies revealed that their CDR1 differed to a great extent, suggesting that their CDR1 does not overlap, at least not fully, with Shh binding. For Nb65, 66, and 67, this region most likely forms the main contacts to Ptc1 when its in complex with Shh. Therefore, these nanobodies might form additional hydrogen bonds and salt bridges between their CDR1 and Ptc1 ECD1, whereas Nb64 only displayed non-bonded contacts. In addition, the CDR2 of Nb66 and Nb67 is distinct from Nb64 and Nb65. This region might also be involved in contact formation between the nanobody and Ptc1 ECD1. This hypothesis is supported by the fact that Nb66 and Nb67 showed a higher binding affinity to the Shh-Ptc1 complex compared to Nb65. Overall, this analysis suggests that it is mainly the regions targeted by CDR3 and only partly the regions targeted by CDR1 which overlap with Shh binding on Ptc1 ECD1.

The structure revealed another striking feature close to the nanobody binding site. Three α-helices and one β-strand formed a hydrophobic core. This core is distinct from hydrophobic cholesterol pockets observed in Smo and the Ptc1 homologue NPC1 [36]. Our co-crystallisation efforts revealed that this hydrophobic pocket is able to accommodate the cholesterol derivative cholesteryl hemisuccinate (CHS). This experiment was based on the observation of weak electron density observed for the apo Ptc1-Nb64 complex that the protein chains itself did not account for. Interestingly, cholesterol or other lipids were not present during protein purification,

suggesting that a lipidic molecule had been taken up naturally by Ptc1 ECD1 during expression in HEK293T cells. To test whether this ligand binding domain (LBD) is solely responsible for the uptake of a lipid or could potentially be the anchor point for the cholesterol moiety of Shh, co-crystallisation with lipidated Shh should be carried out. Nevertheless, the observation of a binding pocket for lipids on the Ptc1 ECD1 reinforces the hypothesis of Ptc1 being a cholesterol transporter.

Structural studies of Ptc1 and Smo in the Siebold group have identified a cholesterol binding site in both, the ECD1 of Ptc1 (this study, unpublished) and the cysteine-rich domain (CRD) of Smo (E. Byrne *et al.*, [39]). These results confirm that cholesterol plays an important role in Hh signalling; however, the exact mechanism by which cholesterol is transported to these domains remains to be determined. The location of Smo in lipid rafts, which exhibit a high content of cholesterol and other lipids, has been found to be essential for Smo function [251]. However, extracting cholesterol from the outer membrane leaflet to the Smo CRD via the aqueous environment is energetically unfavourable [39, 108, 173, 291]. Hence, it is more likely that an external source such as lipoprotein particles regulate the delivery of cholesterol to Smo. In order to inhibit Smo function, Ptc1 might restrict the supply of cholesterol. What then could be the underlying mechanism? If Smo and Ptc1 are in close proximity, cholesterol could be directly taken over from the Smo CRD by the Ptc1 ECD1 and subsequently internalised via the sterol-sensing domain (SSD). In response, Smo could be removed from the membrane. This model is supported by previous studies which indicated that Ptc1 regulates Smo catalytically rather than stoichiometrically [264]. If this is the case, physical contact might not be established between the two proteins. Instead, Ptc1 may transport cholesterol away before it can be delivered to Smo. This mechanism could be disrupted once the Hh ligand binds to Ptc1. In response, cholesterol can be recruited by Smo which subsequently activates the path-

way. In summary, the detailed mechanism of cholesterol transport in the Hh pathway remains open for debate and additional data are required to support current models. It should also be noted that deriving a model is further complicated by the fact that Smo may change its subcellular localisation upon pathway activation [82, 264].

Chapter 7

Conclusion

7.1 The design of a novel Hh signalling inhibitor

The activity of cells is regulated by numerous signalling pathways. Their main function is to execute diverse cell fate decisions, such as cell growth, cell proliferation, and apoptosis. In vertebrates, five signalling pathways predominate in early development: the Hedgehog (Hh)-, Wnt-, Notch-, TGF-β-, and RTK signalling pathway [97]. Even though their tasks are remarkably diverse, their common goal is to activate specific target genes to direct cells towards their designated fate. Therefore, it is not surprising that proper signalling is an essential requirement throughout an organism's entire life cycle. Mutations at any stage frequently lead to several disease states and members of signalling pathways are therefore often the primary target for cancer therapy. The development and screening of cancer drugs, however, is inherently difficult and often a long process. A common approach to drug discovery is to isolate and test chemical compounds found in plants, fungi, or animals. Cyclopamine, a small molecule inhibitor of the Hh signalling pathway is naturally occurring in corn lilies [48, 274]. Vismodegib and sonidegib are small molecules that have been developed on the basis of cyclopamine and both drugs target the GPCR Smo. Initially cancer patients respond well to the treatment, however, they very rapidly develop resistance towards the drugs [243]. This phenomenon of drug resistance can be explained by the crystal structure of the Smo-vismodegib complex, which gives detailed insights into the binding interface [39]. When residues in the binding pocket are mutated, vismodegib fails to interact with Smo and thus renders the drug ineffective. This is also an excellent example of how structural characterisation of protein-ligand complexes can lead to the design of new compounds for cancer treatment. Besides small molecule inhibitors, pharmaceutical industries focus on the development of protein/peptide-based inhibitors. Recently, special emphasis has been placed on nanobodies, camelid

single-chain antibody fragments. Nanobodies often bind to their target with nanomolar affinity and exhibit high target specificity.

The first part of this thesis focused on the identification of Shh-specific nanobodies that show inhibitory effects on Hh signalling (Chapter 3 and 4). Nanobodies that target Shh have great potential for treating cancers resulting from Hh ligand-dependent over activation of Hh signalling. Cellular studies of the Shh-specific nanobodies unveiled that in particular Nb70 and Nb77 have a strong inhibitory effect on Hh signalling. The crystal structures of the Shh-Nb70 and Shh-Nb77 complexes revealed that these nanobodies bind to opposite sites of Shh. In order to explain how inhibition of the Hh signal is achieved, we compared the Shh-Nb binding interfaces to those of Cdo, a Hh signalling co-receptor, and Hhip, a negative regulator of Hh signalling. The interface of Nb70 on Shh overlaps partly with these proteins, suggesting that the nanobody prevents binding of Cdo to Shh and either supports the inhibitory function of Hhip or outcompetes the protein. In addition, comparison of our results to steric interference studies on Shh and Ptc1 suggested that parts of the Shh-Nb70 binding interface overlap with the Shh binding site of Ptc1 [216]. Even though these observations explained the inhibitory effects of Nb70, it did not explain them for Nb77 which uses a binding interface opposite to Nb70. However, these findings suggested that Ptc1 may either function as a homo- or heterodimer which is able to interact with two sites of Shh. During preparation of this thesis, three manuscripts on the cryo-EM structure of Ptc1 and the Shh-Ptc1 complex had been published which supported this hypothesis and helped to shed light on the inhibitory function of the nanobodies. X. Gong *et al.* reported that the interaction between Shh and Ptc1 is achieved by residues located near the Shh zinc-binding site, an observation that is consistent with previous literature findings [101, 186, 216]. On the contrary, the structures reported by X. Qi *et al.* revealed that Ptc1 interacts with residues on Shh

which lie opposite the zinc-binding site; an observation that had not been reported in the literature before [227, 228]. Although these structures were controversial at first, they were confirmed by our findings reported in Chapter 4: Nb77 can inhibit Hh signalling, even though its binding site lies opposite the Shh zinc-binding site. Our structures of the Shh-Nb70 and Shh-Nb77 complex therefore validated the structures of the Shh-Ptc1 complex reported by X. Qi and co-workers. Another publication by X. Qi *et al.* then reported the structure of a 1:2 Shh:Ptc1 complex, which was formed by the dual-lipidated N-terminal domain of Shh and full-length Ptc1 (Figure 7.1, A). In their structure, Shh contacts Ptc1 using two distinct binding sites. One interface is formed between the palmitoylated N-terminal tail of Shh, which inserts into a cavity formed by the two Ptc1 extracellular domains ECD1 and ECD2. Another interface is established between the residues near the zinc-binding site of Shh and the ECDs of the second Ptc1 molecule. A structural overlay of the Shh-Nb complexes with the 1:2 Shh:Ptc1 complex suggests that both nanobodies, Nb70 and Nb77, would interfere with Shh binding to the Ptc1 dimer (Figure 7.1, B-C). All CDRs of Nb70 as well as the CDR2 and a few framework residues on β6 and β7 of Nb77 conflict with residues located on the Ptc1 ECD1. Both Shh-Nb structures in the light of the 1:2 Shh:Ptc1 complex yield a structural explanation of how Nb70 and Nb77 inhibit Hh signalling. Before those structures had been published, it was decided to design bispecific nanobodies by linking Nb70 and Nb77 with a Gly/Ser (G/S)-rich linker of different lengths. In Gli1 transcription reporter assays it was verified that these bispecific nanobodies have an increased inhibitory potential. It will be interesting to continue the work on the bispecific nanobodies as they present novel Hh signalling inhibitors with the potential for the treatment of Hh ligand-dependent cancers, such as breast, lung, pancreas, prostate and stomach cancer [242].

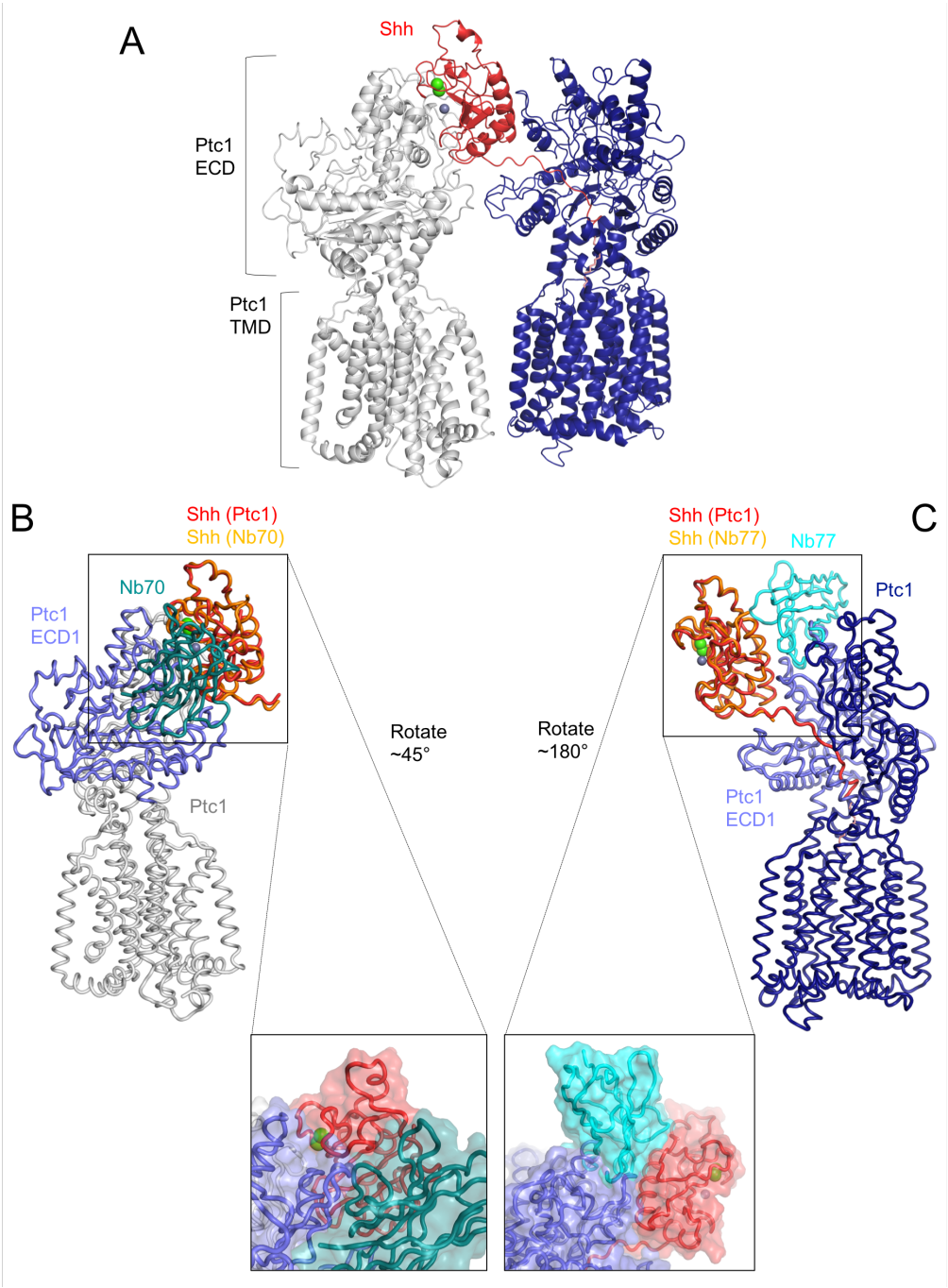


Figure 7.1: **A model to explain the inhibitory function of Shh-specific nanobodies. A**, Cryo-EM structure determined by X. Qi and co-workers (PDB code: 6E1H) as cartoon representation [227]. Ptc1 is coloured in grey and density, palmitoylated Shh is shown in red. Zn^{2+} ions and Ca^{2+} ions are shown as grey and green spheres, respectively. **B-C**, Structural alignment of the Shh-Ptc1 cryo-EM structure and the Shh-Nb crystal structures shown as tube representation. Both nanobodies (Nb70 in deepteal, Nb77 in cyan) collide with residues of the Ptc1 ECD1 (slate) of two different Ptc1 molecules (grey, density) in the 1:2 Shh:Ptc1 complex. The cryo-EM structure of Shh is coloured in red, the crystal structure in orange. Zn^{2+} ions and Ca^{2+} ions are shown as grey and green spheres, respectively. Close-up views show colliding areas between the nanobodies and Ptc1 ECD1.

7.2 Towards a structure of full-length Ptc1

Ptc1 is a major component of the Hh signalling pathway. In the absence of Shh, the pathway is silenced by Ptc1-mediated suppression of the GPCR Smo. Binding of Shh to Ptc1 activates the pathway. However, until today, the detailed mechanism remains to be determined. In order to contribute to our current understanding of Hh signalling, the second part of this thesis focused on the structural characterisation of Ptc1 (Chapter 5 and 6).

Ptc1 is a 12 TM-spanning protein with two large extracellular domains, which are responsible for Shh signal reception. For structural studies it is essential to design a stable construct with minimal flexibility. For this reason, we truncated the N- and C-termini as well as the third intracellular loop (ICL3) located in between TM-VI and TM-VII. Truncation of these regions also deleted two signal recognition sequences for ubiquitin ligases, which resulted in a higher protein yield as the construct was less prone to degradation in lysosomes. In cellular assays we further identified a region on Ptc1 ICL3 which might have a function in targeting Ptc1 to the primary cilium. This potential cilia localisation signal is located between residues Phe^{611} and Tyr^{630}. In order to enhance the protein's crystallisability even further, we introduced a bRIL domain to this region. In addition, we have identified a number of Ptc1-specific nanobodies that bind to the construct with nanomolar affinity and increase the hydrophilic surface of the membrane protein. Despite extensive protein engineering efforts and screening of a vast variety of crystallisation conditions and techniques, the protein remained difficult to crystallise. There are many factors that could have prevented crystallisation. MALS experiments for example suggested that the sample contained Ptc1 monomer and dimer; a monomer-dimer equilibrium might have interfered with crystal formation. The observation of an oligomeric form of Ptc1

is conform with the results presented by X. Gong *et al.* and X. Qi *et al.* [101, 227, 228]. In order for them to determine the cryo-EM structures of full-length Ptc1, a high number of particles had to be picked to generate a conclusive 3D model (∼0.6 million-3.3 million particles for the Ptc1 monomer and ∼0.28 million particles for the Ptc1 dimer). Interestingly, all three publications reported the use of Quantifoil® grids, which had not been tested in our cryo-EM studies with detergent-solubilised protein. However, promising EM studies on Ptc1 reconstituted in MSP2N2 nanodiscs were ongoing when the structures of full-length Ptc1 were published.

In parallel it was decided to study the extracellular domains separately from the TM domain. Based on the crystal structure of the bacterial RND permease AcrB, a Ptc1 homologue with 36% sequence similarity, a fusion construct of the Ptc1 ECD1 and ECD2 (Ptc1 ECD1-ECD2) was designed. Based on this construct, a total of 12 Ptc1-specific nanobodies were identified that could be used in co-crystallisation experiments. Sufficient yields and sample purity allowed extensive screening of the construct, however, neither Shh nor nanobodies could help facilitate crystallisation. Co-crystallisation with palmitoylated or dual-lipidated Shh might have helped in crystallising the Ptc1 ECD1-ECD2 construct. The cryo-EM structures of X. Qi and co-workers revealed that the palmitoyl moiety inserts in between the two ECDs and interacts with more than ten residues on ECD1 and five residues on ECD2. It is likely that these interactions stabilise the fold of the two domains thus reducing the flexibility of the construct. Furthermore, the cryo-EM structures indicated that the domain boundaries selected for the Ptc1 ECD1-ECD2 construct were sufficient to prevent folding artefacts. Future studies should focus on co-crystallisation experiments with dual-lipidated Shh. A high resolution crystal structure of the Shh-Ptc1 ECD1-ECD2 complex would add to our current understanding of the Shh-Ptc1 interaction.

7.3 Ptc1 ECD1 might be involved in cholesterol transport

Current literature findings suggest a role for Ptc1 as a sterol transporter [25, 126, 264]. It is therefore very likely that not only the TM domain, but also at least one of the two extracellular domains is involved in lipid transport. High resolution structures of the single domains could validate their involvement in transport and shed light on their structural arrangement. Therefore, it was decided to study Ptc1 ECD1 and Ptc1 ECD2 separately. Initial screening of the single domains indicated that Ptc1 ECD1 is more suitable for structural studies. Co-crystallisation of ECD1 with Ptc1-specific Nb64 readily yielded diffracting protein crystals. The crystal structure of the Ptc1 ECD1-Nb64 complex was determined by S-SAD phasing to a resolution of 2.1 Å. The structure revealed a novel fold with a putative hydrophobic subdomain formed by three α-helices. Interestingly, additional electron density appeared during iterative refinement rounds suggesting that the hydrophobic pocket is accommodating a small molecule. In order to confirm whether this molecule is a lipid it was decided to co-crystallise the Ptc1 ECD1-Nb64 complex with the cholesterol derivative cholesteryl hemisuccinate (CHS). The co-crystals diffracted to 1.9 Å and extra electron density was immediately apparent after molecular replacement, suggesting that this core is indeed a lipid-binding domain. The Ptc1 structures presented by X. Gong *et al.* confirm these results; a small molecule is located at a slightly different position in between the three α-helices of their structures [101]. In addition, the full-length structures displayed binding of a lipid in the Ptc1 TM domain [101, 227, 228]. An overlay of the Ptc1 ECD1-Nb64:CHS crystal structure with the 1:2 Shh:Ptc1 cryo-EM structure revealed two other interesting details (Figure 7.2) [227]. Firstly, Shh interacts with both Ptc1 molecules near their α-helical lipid-binding domains (LBD).

In theory, it is possible that the C-terminal cholesterol residue of dual-lipidated Shh inserts into the hydrophobic pocket (Figure 7.2, A). However, the ten most C-terminal residues as well as the covalently attached cholesterol moiety were not visible in the cryo-EM map. Therefore, future experiments will be required to test this hypothesis. Secondly, binding of Nb64, and in particular its CDR3 would most likely block the insertion of the palmitoylated Shh N-terminus in between ECD1 and ECD2 of full-length Ptc1 (Figure 7.2, B-C). Furthermore, in order for Nb64 to interact with the ECD1, it requires to introduce a positional shift of ECD2. This shift might also prevent Shh from binding to Ptc1 via residues near the zinc-binding site. These observations suggest that Nb64 could potentially act as a Hh signalling modulator by preventing Shh to insert its palmitoylated N-terminal tail and by disrupting the structural arrangement of ECD2. With the high resolution crystal structure of Ptc1 ECD1-Nb64:CHS and other information about Ptc1 available to date, a model on how Ptc1 facilitates lipid transport can be proposed. In the absence of Shh, Ptc1 operates as a cholesterol transporter and either enables or restricts cholesterol supply to Smo thus regulating the activity of the GPCR. The LBD on Ptc1 ECD1, the reported tunnel in between ECD1 and ECD2, as well as the Ptc1 sterol-sensing domain (SSD, formed by TM-II to TM-VI), all act together to transport cholesterol across the membrane. In the presence of dual-lipidated Shh, the transport activity is blocked. The palmitate inserts in between the ECD1 and ECD2 interface, locking the protein in its position (Figure 7.2, A). The Shh-Ptc1 interaction might be further stabilised by the C-terminal domain. The cholesterol moiety could insert into the LBD of ECD1. This model is supported by various other findings. Mutations in the SSD and in the tunnel have been shown to abolish the repressor function of Ptc1, whereas ECD1 with a $Leu^{282}Gln$ mutation fails to accommodate a lipid in its hydrophobic cavity [261, 227]. Further studies are required to test whether the mutation of Leu^{282} results

in constitutive activation of Hh signalling. The suggested model of lipid transport is further supported by a study from H. Tukachinsky and co-workers which reports that palmitoylation of Shh is necessary for Hh signalling activation [272]. Furthermore, a study presented by M. Bidet *et al.* reports that cholesterol transport by Ptc1 is prevented in the presence of Shh [25]. It remains to be determined whether the cholesterol moiety of Shh interacts with the lipid-binding domain on ECD1.

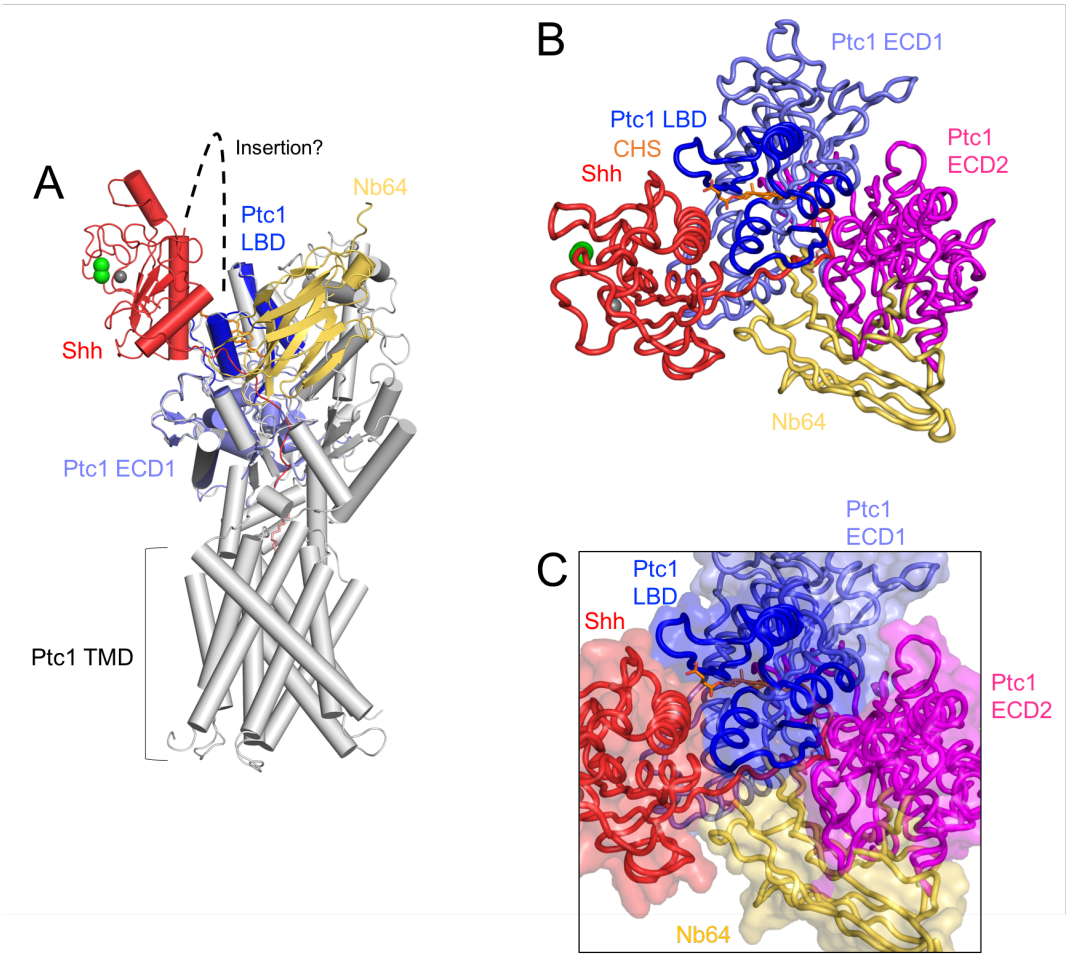


Figure 7.2: **Nb64 collides with residues located on Ptc1 ECD2. A-C**, Overlay of the Ptc1 ECD1-Nb64 crystal structure with the Ptc1 cryo-EM structure determined by X. Qi and co-workers (PDB code: 6E1H) [227]. For reasons of simplicity, only one Ptc1 domain is shown; helices shown as cylinders. **A**, The Ptc1 cryo-EM structure is coloured in grey, Shh in red. Zn^{2+} ions and Ca^{2+} ions are shown as grey and green spheres, respectively. The Ptc1 ECD1 crystal structure is shown in slate with the LBD highlighted in blue, Nb64 in yellow. CHS is shown as orange sticks. **B**, Top view of A shown as tube representation. Nb64 collides with residues of Ptc1 ECD2 and possibly prevents insertion of the Shh palmitoylated tail. Colour coding as in A, Ptc1 ECD2 is shown in magenta **C**, Surface representation of B.

7.4 Concluding remarks

In this study, a total of ten Shh-specific nanobodies were identified and their inhibitory potential was tested by using a combination of X-ray crystallography and cellular studies. Based on the structures of eight Shh-Nb complexes, their binding affinities and their inhibitory potential in cellular assays, novel bispecific nanobodies were designed. Compared to monospecific nanobodies, the bispecific nanobodies showed enhanced inhibitory effects on Hh signalling activity. We further validated their inhibitory potential on the basis of the Shh-Ptc1 cryo-EM structures presented by X. Gong *et al.* and X. Qi *et al.* [101, 227, 228]. This study may lay the groundwork for structure-based drug design and the development of a nanobody-based treatment of Hh ligand-dependent cancers. This thesis further reports the design of a stable and high yield producing full-length Ptc1 construct for structural studies. The construct can be concentrated to at least 54 mg mL^{-1} for crystallisation trials in LCP and reconstituted in MSP-based nanodiscs for EM studies. In order to study the Ptc1 extracellular domains on its own, a functional, secreted Ptc1 ECD construct was designed. Based on this construct a total of 12 Ptc1-specific nanobodies were identified for co-crystallisation studies of Ptc1 constructs. This work then paved the way for determining the crystal structure of Ptc1 ECD1 in complex with Nb64 and CHS at high resolution (<2 Å). The structure revealed a novel lipid-binding domain (LBD) located on Ptc1 ECD1 suggesting a role for this domain in lipid transport. On the basis of the presented Ptc1 ECD1-Nb64:CHS crystal structure and the 1:2 Shh:Ptc1 cryo-EM structure determined by X. Qi *et al.*, we further suggest that Nb64 could function as a Hh signalling modulator by impeding the interaction of Shh with Ptc1 thus locking it in its repressor state [101]. In addition, Nb64 may have potential in preventing Ptc1-induced chemotherapeutic resistance by blocking chemotherapeutic

drug transport which might be carried out by Ptc1 [26, 263]. Such a function for Nb64 should be explored in future studies as it may have potential in cancer therapy.

Appendix A

Plasmid maps



Figure A.1: **Plasmid maps for Shh and nanobody constructs.** Map of the pET22b(+) plasmid (top) for expression of Shh, and the pMESy4 plasmid (bottom) for expression of nanobodies, showing key regions. Expression host: *E.coli*.

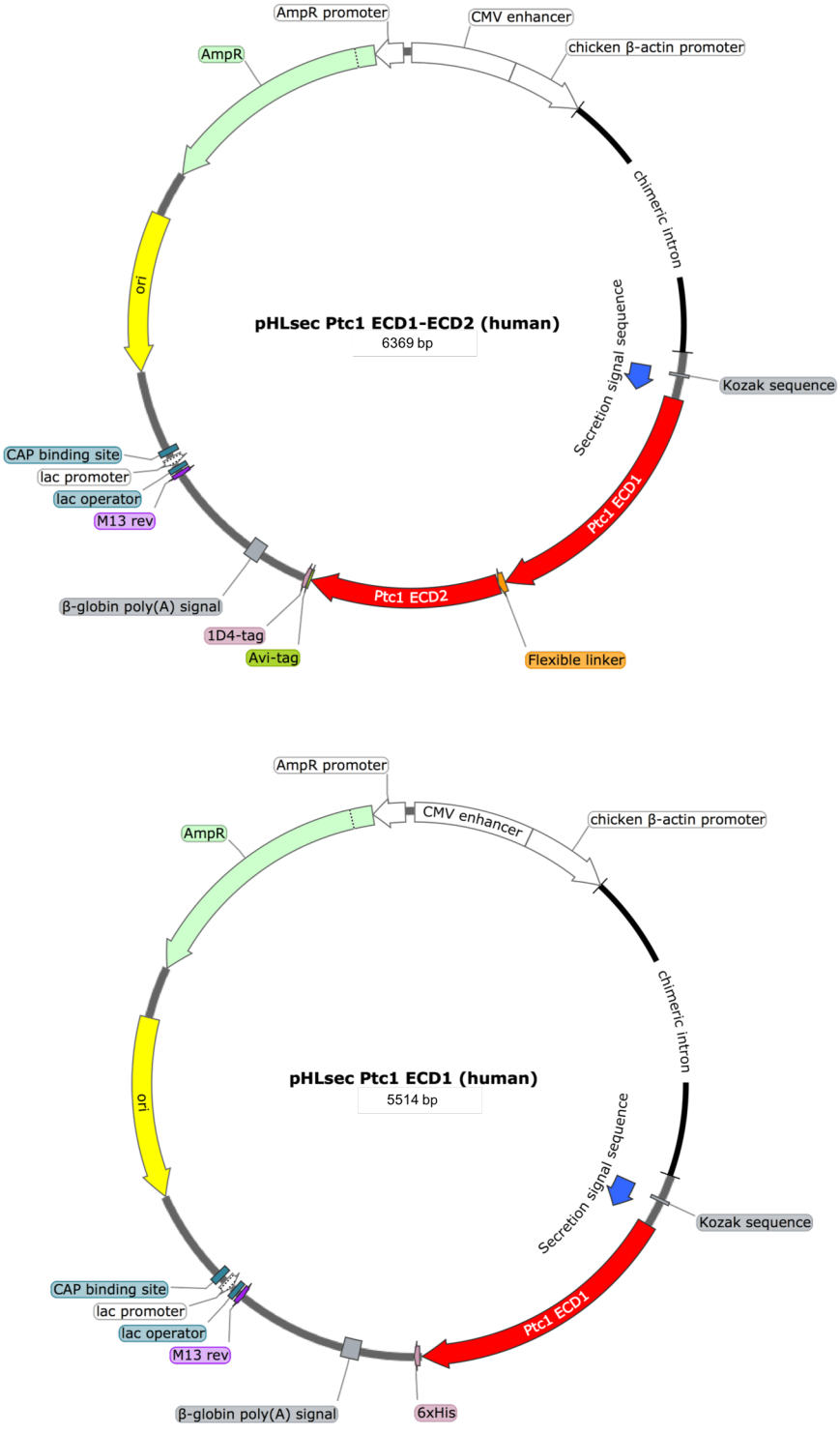


Figure A.2: **Plasmid maps for Ptc1 ECD constructs.** Map of the pHLsec plasmid for the expression of Ptc1 ECD domains, showing key regions. A plasmid with the same features as pHLsec Ptc1 ECD1-ECD2 was used for the expression of Nb70/Nb77 and Nb77/Nb70 linker constructs. The 1D4-tag was substituted with a His_6-tag. Expression host: HEK293 cells.

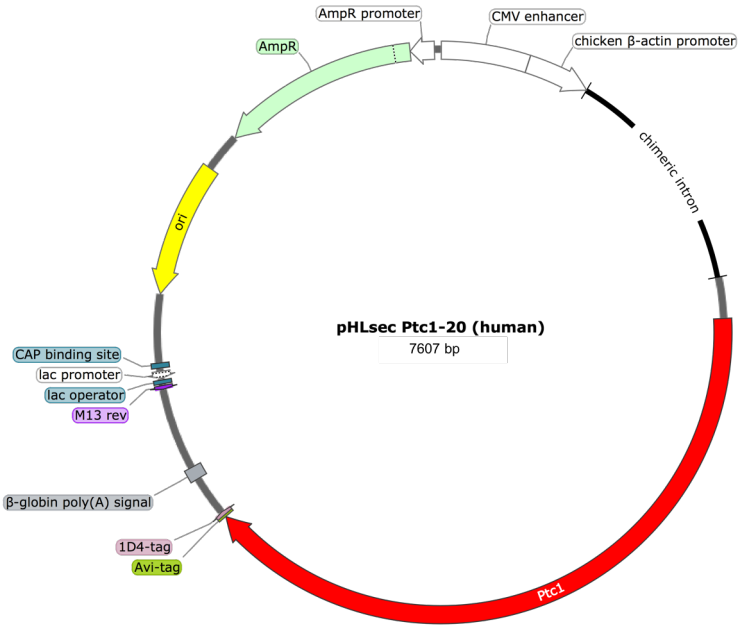


Figure A.3: **Plasmid map for Ptc1 constructs.** Map of the pHLsec plasmid as used for the expression of Ptc1 membrane protein constructs, showing key regions. The same vector backbone was used for the expression of Ptc1-20ΔICL3-5 and Ptc1-20ΔICL3-5-bRIL constructs. Expression host: HEK293 cells.

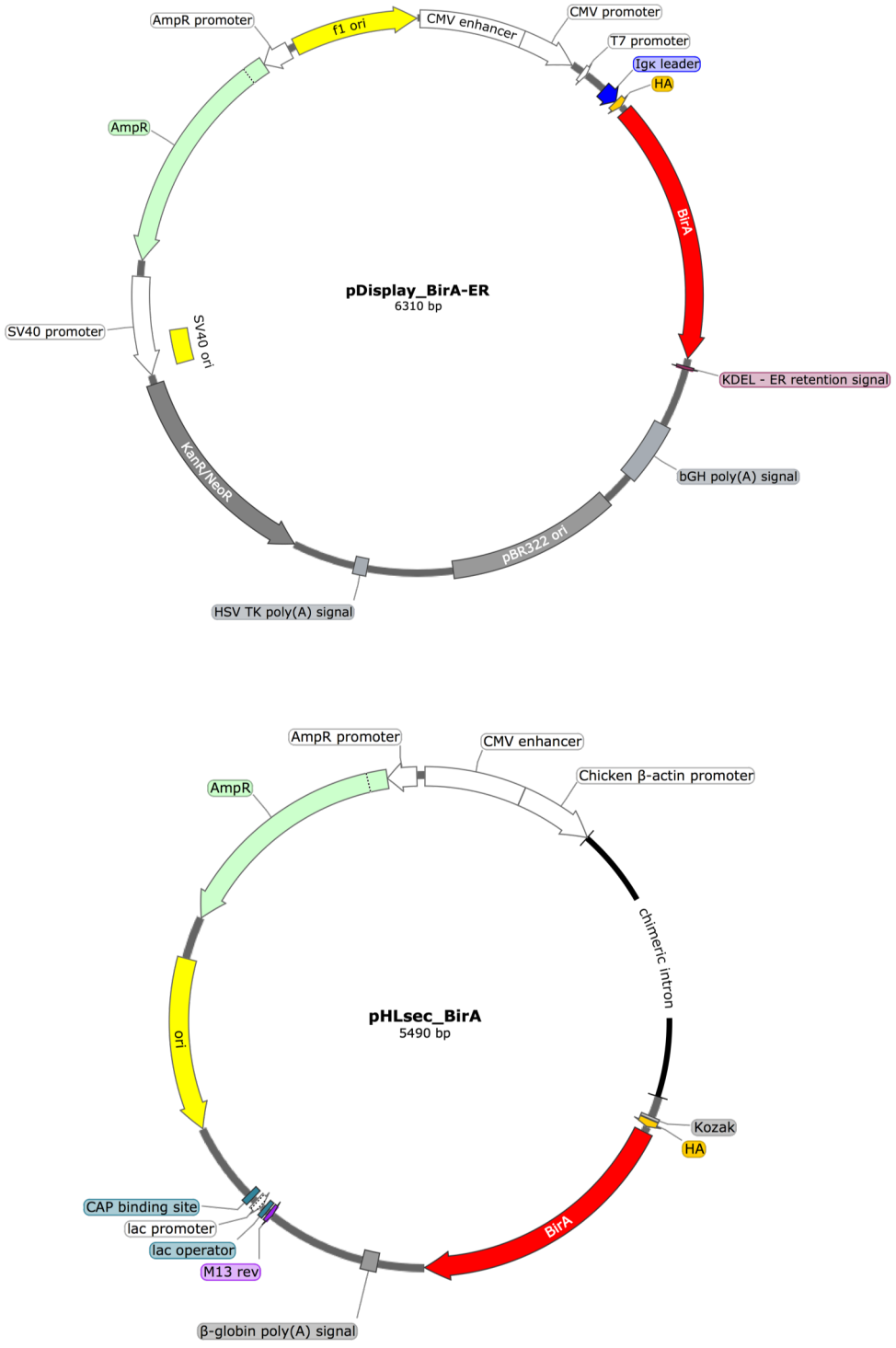


Figure A.4: **Plasmid maps for BirA constructs.** Map of the pDisplay plasmid (top) for the expression of BirA biotin ligase in the endoplasmatic reticulum (ER), and the pHLsec plasmid (bottom) for the expression of soluble BirA, showing key regions. Expression host: HEK293 cells.

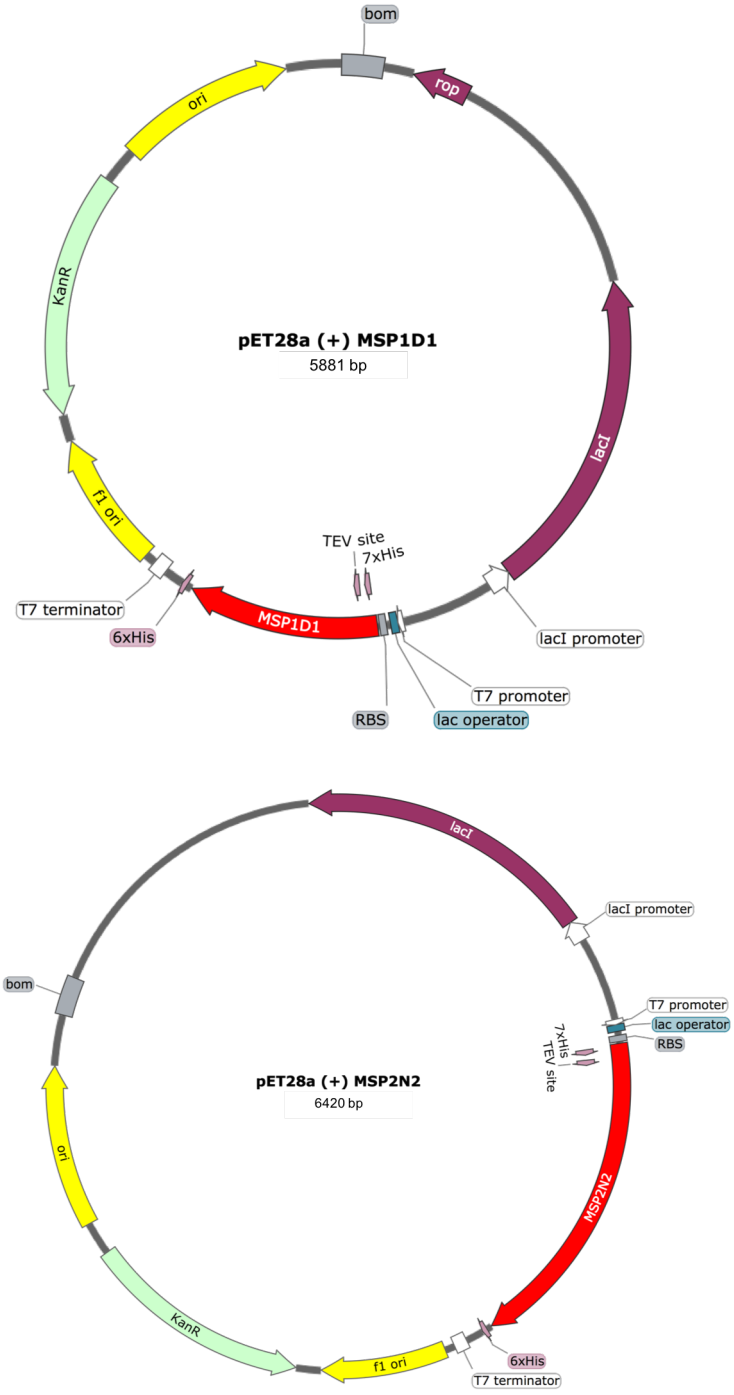


Figure A.5: **Plasmid maps for MSP constructs.** Map of pET28a(+) plasmid for expression of MSP1D1 (top) and MSP2N2 (bottom), showing key regions. Expression host: *E.coli*.

Appendix B

Constructs

Construct name	Plasmid	Tag	Protein	Amino Acids Start-End	Linker (L)
Shh	pHLsec	Avi, 6x His	human Shh	/24-197	none
Shh	pET22b	Avi, 6x His	human Shh	/24-197	none
Ptc1 ECD1-ECD2	pHLsec	Avi, 1D4	human Patched-1	/120-435 + L + /789-1023	GGSNGSGG
Ptc1 ECD1	pHLsec	6 x His	human Patched-1	/139-428	none
Ptc1-20	pHLsec	Avi, 1D4	human Patched-1	/75-1185	none
Ptc1-1-mV	pHLsec	mVenus, 1D4	human Patched-1	/1-1179	none
Ptc1-2-mV	pHLsec	mVenus, 1D4	human Patched-1	/1-1185	none
Ptc1-3-mV	pHLsec	mVenus, 1D4	human Patched-1	/1-1235	none
Ptc1-4-mV	pHLsec	mVenus, 1D4	human Patched-1	/1-1265	none
Ptc1-5-mV	pHLsec	mVenus, 1D4	human Patched-1	/1-1381	none
Ptc1-6-mV	pHLsec	mVenus, 1D4	human Patched-1	/1-1447	none
Ptc1-7-mV	pHLsec	mVenus, 1D4	human Patched-1	/48-1179	none
Ptc1-8-mV	pHLsec	mVenus, 1D4	human Patched-1	/48-1185	none
Ptc1-9-mV	pHLsec	mVenus, 1D4	human Patched-1	/48-1235	none
Ptc1-10-mV	pHLsec	mVenus, 1D4	human Patched-1	/48-1265	none
Ptc1-11-mV	pHLsec	mVenus, 1D4	human Patched-1	/48-1381	none
Ptc1-12-mV	pHLsec	mVenus, 1D4	human Patched-1	/48-1447	none
Ptc1-13-mV	pHLsec	mVenus, 1D4	human Patched-1	/58-1179	none
Ptc1-14-mV	pHLsec	mVenus, 1D4	human Patched-1	/58-1185	none
Ptc1-15-mV	pHLsec	mVenus, 1D4	human Patched-1	/58-1235	none
Ptc1-16-mV	pHLsec	mVenus, 1D4	human Patched-1	/58-1265	none
Ptc1-17-mV	pHLsec	mVenus, 1D4	human Patched-1	/58-1381	none
Ptc1-18-mV	pHLsec	mVenus, 1D4	human Patched-1	/58-1447	none
Ptc1-19-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-1179	none
Ptc1-20-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-1185	none
Ptc1-21-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-1235	none
Ptc1-22-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-1265	none
Ptc1-23-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-1381	none
Ptc1-24-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-1447	none
Ptc1-25-mV	pHLsec	mVenus, 1D4	human Patched-1	/99-1179	none
Ptc1-26-mV	pHLsec	mVenus, 1D4	human Patched-1	/99-1185	none
Ptc1-27-mV	pHLsec	mVenus, 1D4	human Patched-1	/99-1235	none
Ptc1-28-mV	pHLsec	mVenus, 1D4	human Patched-1	/99-1265	none
Ptc1-29-mV	pHLsec	mVenus, 1D4	human Patched-1	/99-1381	none
Ptc1-30-mV	pHLsec	mVenus, 1D4	human Patched-1	/99-1447	none
Ptc1-20ΔICL3-1-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-610 + L + /693-1185	GSGGSSGG
Ptc1-20ΔICL3-2-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-610 + L + /718-1185	GSGGSSGG
Ptc1-20ΔICL3-3-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-629 + L + /650-1185	GSGGSSGG
Ptc1-20ΔICL3-4-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-629 + L + /693-1185	GSGGSSGG
Ptc1-20ΔICL3-5-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-629 + L + /718-1185	GSGGSSGG
Ptc1-20ΔICL3-5	pHLsec	Avi, 1D4	human Patched-1	/75-629 + L + /718-1185	GSGGSSGG
Ptc1-20ΔICL3-ΔICL1-bRIL-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-453 + bRIL + /468-629 + L + /718-1185	GSGGSSGG
Ptc1-20ΔICL3-ΔICL2-bRIL-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-529 + bRIL + /533-629 + L + /718-1185	GSGGSSGG
Ptc1-20ΔICL3-bRIL-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-629 + bRIL + /718-1185	GSGGSSGG
Ptc1-20ΔICL3-ΔICL4-bRIL-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-629 + L + /718-1047 + bRIL + /1051-1185	GSGGSSGG
Ptc1-20ΔICL3-ΔICL5-bRIL-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-629 + L + /718-1107 + bRIL + /1111-1185	GSGGSSGG
Ptc1-20ΔICL3-bRIL	pHLsec	Avi, 1D4	human Patched-1	/75-629 + bRIL + /718-1185	none
Ptc1-20ΔICL1-bRIL-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-453 + bRIL + /468-1185	none
Ptc1-20ΔICL2-bRIL-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-529 + bRIL + /533-1185	none
Ptc1-20ΔICL4-bRIL-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-1047 + bRIL + /1051-1185	none
Ptc1-20ΔICL5-bRIL-mV	pHLsec	mVenus, 1D4	human Patched-1	/75-1107 + bRIL + /1111-1185	none

Figure B.1: **Overview of constructs used in this thesis (1/2).**

Construct name	Plasmid	Tag	Protein	Amino Acids Start-End	Linker (L)
Nanobody 64 (Nb64)	pMESy4	6 x His	camelid Antibody	/1-120	none
Nanobody 65 (Nb65)	pMESy4	6 x His	camelid Antibody	/1-120	none
Nanobody 66 (Nb66)	pMESy4	6 x His	camelid Antibody	/1-120	none
Nanobody 67 (Nb67)	pMESy4	6 x His	camelid Antibody	/1-120	none
Nanobody 68 (Nb68)	pMESy4	6 x His	camelid Antibody	/1-124	none
Nanobody 69 (Nb69)	pMESy4	6 x His	camelid Antibody	/1-124	none
Nanobody 70 (Nb70)	pMESy4	6 x His	camelid Antibody	/1-126	none
Nanobody 71 (Nb71)	pMESy4	6 x His	camelid Antibody	/1-124	none
Nanobody 72 (Nb72)	pMESy4	6 x His	camelid Antibody	/1-123	none
Nanobody 75 (Nb75)	pMESy4	6 x His	camelid Antibody	/1-124	none
Nanobody 76 (Nb76)	pMESy4	6 x His	camelid Antibody	/1-122	none
Nanobody 77 (Nb77)	pMESy4	6 x His	camelid Antibody	/1-115	none
Nanobody 78 (Nb79)	pMESy4	6 x His	camelid Antibody	/1-126	none
Nanobody 790 (Nb790)	pMESy4	6 x His	camelid Antibody	/1-120	none
Nanobody 791 (Nb791)	pMESy4	6 x His	camelid Antibody	/1-120	none
Nanobody 793 (Nb793)	pMESy4	6 x His	camelid Antibody	/1-120	none
Nanobody 794 (Nb794)	pMESy4	6 x His	camelid Antibody	/1-124	none
Nanobody 795 (Nb795)	pMESy4	6 x His	camelid Antibody	/1-126	none
Nanobody 796 (Nb796)	pMESy4	6 x His	camelid Antibody	/1-124	none
Nanobody 797 (Nb797)	pMESy4	6 x His	camelid Antibody	/1-123	none
Nanobody 798 (Nb798)	pMESy4	6 x His	camelid Antibody	/1-127	none
Nanobody 799 (Nb799)	pMESy4	6 x His	camelid Antibody	/1-126	none
Nanobody 800 (Nb800)	pMESy4	6 x His	camelid Antibody	/1-124	none
Nanobody 802 (Nb802)	pMESy4	6 x His	camelid Antibody	/1-123	none
Nanobody 803 (Nb803)	pMESy4	6 x His	camelid Antibody	/1-126	none
Nanobody 804 (Nb804)	pMESy4	6 x His	camelid Antibody	/1-126	none
Nb70/Nb77:10-res	pHLsec	6 x His	camelid Antibody	/1-126 + L10 + /1-115	see below
Nb70/Nb77:20-res	pHLsec	6 x His	camelid Antibody	/1-126 + L20 + /1-115	see below
Nb70/Nb77:30-res	pHLsec	6 x His	camelid Antibody	/1-126 + L30 + /1-115	see below
Nb77/Nb70:10-res	pHLsec	6 x His	camelid Antibody	/1-115 + L10 + /1-126	see below
Nb77/Nb70:20-res	pHLsec	6 x His	camelid Antibody	/1-115 + L20 + /1-126	see below
Nb77/Nb70:30-res	pHLsec	6 x His	camelid Antibody	/1-115 + L30 + /1-126	see below
MSP1D1	pET28a	7 x His	Apolipoprotein	/1-189	none
MSP2N2	pET28a	7 x His	Apolipoprotein	/1-254	none
L10 = GGSGGSGGSG					
L20 = GGSGGSGGSGGSGGSGGSGG					
L30 = GGSGGSGGSGGSGGSGGSGGSGGSGGSGGS					

Figure B.2: **Overview of constructs used in this thesis (2/2).**

Appendix C

Chemical structures of detergents and lipids

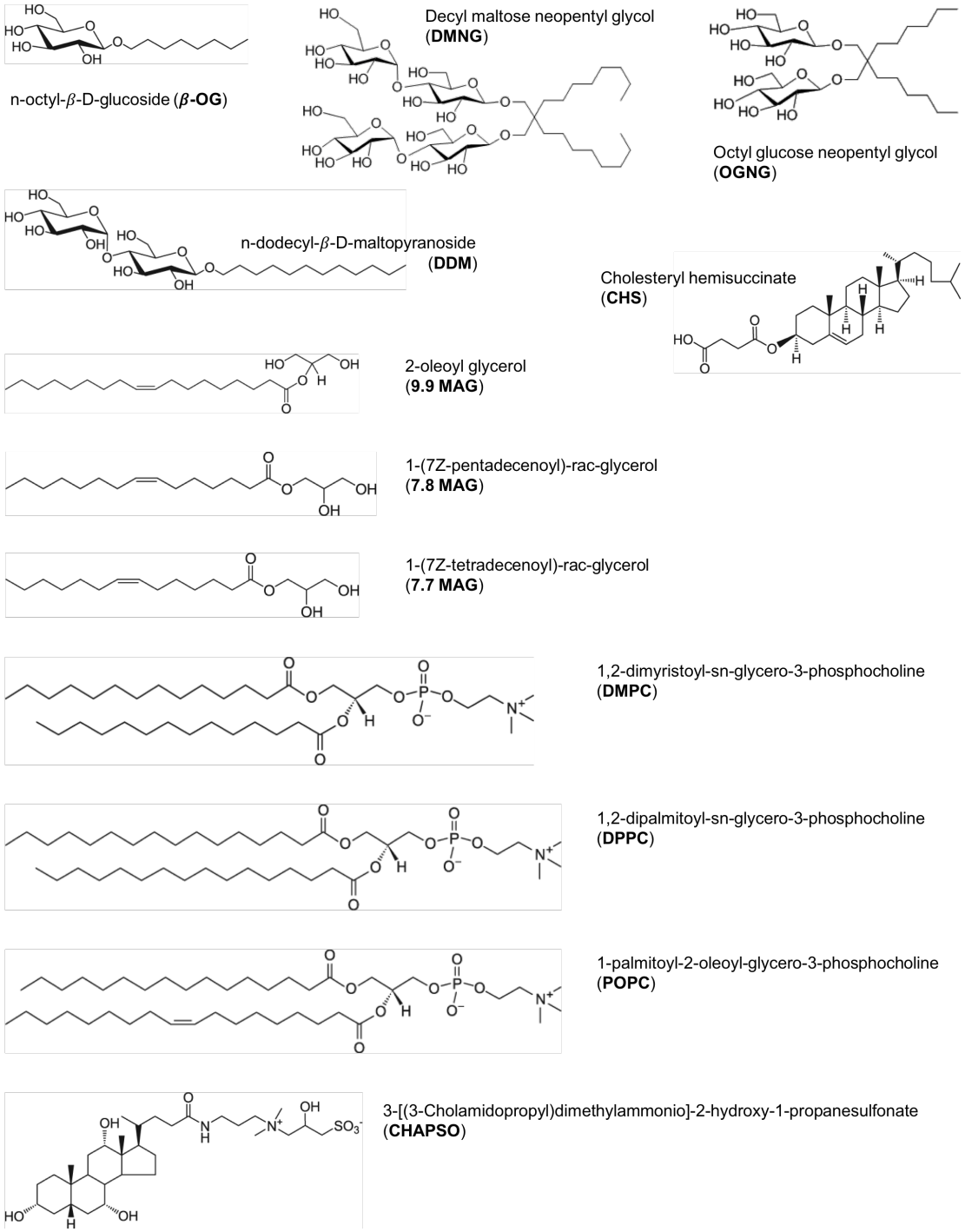


Figure C.1: **Detergents and lipids used for Ptc1 work.** Detergents used for purification of membrane proteins: β-OG, DMNG, OGNG, DDM, and CHS. Lipids used for LCP: CHS, 9.9 MAG, 7.8 MAG, and 7.7 MAG. Lipids used for nanodisc preparation: DMPC, DPPC, and POPC. Lipids used for crystallisation in bicelles: DMPC and CHAPSO.

References

[1] Cablivi^{TM} (caplacizumab) approved in Europe for adults with acquired thrombotic thrombocytopenic purpura (aTTP) (Status 2018) [Online] www.news.sanofi.us. Available at: http://www.news.sanofi.us/2018-09-03-Cablivi-TM-caplacizumab-approved-in-Europe-for-adults-with-acquired-thrombotic-thrombocytopenic-purpura-aTTP, Accessed 2019-02-02.

[2] Exclusive: Amgen's new leukemia drug to carry $178,000 price tag (Status 2014) [Online] www.reuters.com. Available at: https://www.reuters.com/article/us-amgen-cancer-exclusive-idUSKBN0JV1YU20141217, Accessed 2018-08-20.

[3] Membrane Proteins of Known 3D Structure (Status 2018) [Online] http://blanco.biomol.uci.edu/mpstruc/. Available at: http://blanco.biomol.uci.edu/mpstruc/, Accessed on 2018-08-31.

[4] Monoclonal Antibodies Approved by the EMA and FDA for Therapeutic Use (Status 2017) [Online] www.actip.org. Available at: http://www.actip.org/products/monoclonal-antibodies-approved-by-the-ema-and-fda-for-therapeutic-use/, Accessed on 2018-06-25.

[5] Nanodisc Technology: Protocols for preparation of nanodiscs (Status 2008) [Online] www.sligarlab.life.uiuc.edu. Available at: http://sligarlab.life.uiuc.edu/nanodisc/protocols.html, Accessed 2018-08-30.

[6] R. N. Abskharon, G. Giachin, A. Wohlkonig, S. H. Soror, E. Pardon, G. Legname, and J. Steyaert. Probing the N-terminal β-sheet conversion in the crystal structure of the human prion protein bound to a nanobody. *Journal of the American Chemical Society*, 136(3):937–944, 2014.

[7] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica. Section D, Biological Crystallography*, 66(Pt 2):213–21, 2 2010.

[8] M. Akam. The molecular basis for metameric pattern in the Drosophila embryo. *Development*, 101(1):1–22, 9 1987.

[9] A. Allegra, V. Innao, D. Gerace, D. Vaddinelli, A. G. Allegra, and C. Musolino. Nanobodies and cancer: Current status and new perspectives. *Cancer Investigation*, 36(4):221–237, 4 2018.

[10] B. L. Allen, J. Y. Song, L. Izzi, I. W. Althaus, J.-S. Kang, F. Charron, R. S. Krauss, and A. P. McMahon. Overlapping roles and collective requirement for the coreceptors GAS1, CDO, and BOC in SHH pathway function. *Developmental cell*, 20(6):775–87, 6 2011.

[11] D. Amakye, Z. Jagani, and M. Dorsch. Unraveling the therapeutic potential of the Hedgehog pathway in cancer. *Nature Medicine*, 19(11):1410–1422, 11 2013.

[12] A. R. Aricescu, W. Lu, and E. Y. Jones. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallographica. Section D, Biological Crystallography*, 62(10):1243–1250, 10 2006.

[13] S. A. Arnold, S. Albiez, A. Bieri, A. Syntychaki, R. Adaixo, R. A. McLeod, K. N. Goldie, H. Stahlberg, and T. Braun. Blotting-free and lossless cryo-electron microscopy grid preparation from nanoliter-sized protein samples and single-cell extracts. *Journal of Structural Biology*, 197(3):220–226, 3 2017.

[14] X.-C. Bai, C. Yan, G. Yang, P. Lu, D. Ma, L. Sun, R. Zhou, S. H. W. Scheres, and Y. Shi. An atomic structure of human γ-secretase. *Nature*, 525(7568):212–217, 9 2015.

[15] P. Bannas, J. Hambach, and F. Koch-Nolte. Nanobodies and nanobody-based human heavy chain antibodies as antitumor therapeutics. *Frontiers in Immunology*, 8:1603, 11 2017.

[16] X. Barros-Álvarez, K. M. Kerchner, C. Y. Koh, S. Turley, E. Pardon, J. Steyaert, R. M. Ranade, J. R. Gillespie, Z. Zhang, C. L. Verlinde, E. Fan, F. S. Buckner, and W. G. Hol. Leishmania donovani tyrosyl-tRNA synthetase structure in complex with a tyrosyl adenylate analog and comparisons with human and protozoan counterparts. *Biochimie*, 138:124–136, 7 2017.

[17] K. Basler and G. Struhl. Compartment boundaries and the control of Drosopfiffa limb pattern by Hedgehog protein. *Nature*, 368(6468):208–214, 3 1994.

[18] A. Bateman, M. Martin, C. O'Donovan, M. Magrane, E. Alpi, R. Antunes, B. Bely, M. Bingley, C. Bonilla, R. Britto, B. Bursteinas, H. Bye-A-Jee, A. Cowley, A. D. Silva, M. D. Giorgi, T. Dogan, F. Fazzini, L. G. Castro, L. Figueira, P. Garmiri, G. Georghiou, D. Gonzalez, E. Hatton-Ellis, W. Li, W. Liu, R. Lopez, J. Luo, Y. Lussi, A. MacDougall, A. Nightingale, B. Palka, K. Pichler, D. Poggioli, S. Pundir, L. Pureza, G. Qi, A. Renaux, S. Rosanoff, R. Saidi, T. Sawford, A. Shypitsyna, E. Speretta, E. Turner, N. Tyagi, V. Volynkin, T. Wardell, K. Warner, X. Watkins, R. Zaru, H. Zellner, I. Xenarios, L. Bougueleret, A. Bridge, S. Poux, N. Redaschi, L. Aimo, G. Argoud-Puy, A. Auchincloss, K. Axelsen, P. Bansal, D. Baratin, M.-C. Blatter, B. Boeckmann, J. Bolleman, E. Boutet, L. Breuza, C. Casal-Casas, E. d. Castro, E. Coudert, B. Cuche, M. Doche, D. Dornevil, S. Duvaud, A. Estreicher, L. Famiglietti, M. Feuermann, E. Gasteiger, S. Gehant, V. Gerritsen, A. Gos, N. Gruaz-Gumowski, U. Hinz, C. Hulo, F. Jungo, G. Keller, V. Lara, P. Lemercier,

D. Lieberherr, T. Lombardot, X. Martin, P. Masson, A. Morgat, T. Neto, N. Nouspikel, S. Paesano, I. Pedruzzi, S. Pilbout, M. Pozzato, M. Pruess, C. Rivoire, B. Roechert, M. Schneider, C. Sigrist, K. Sonesson, S. Staehli, A. Stutz, S. Sundaram, M. Tognolli, L. Verbregue, A.-L. Veuthey, C. H. Wu, C. N. Arighi, L. Arminski, C. Chen, Y. Chen, J. S. Garavelli, H. Huang, K. Laiho, P. McGarvey, D. A. Natale, K. Ross, C. R. Vinayaka, Q. Wang, Y. Wang, L.-S. Yeh, and J. Zhang. UniProt: the universal protein knowledgebase. *Nucleic Acids Research*, 45(D1):D158–D169, 1 2017.

[19] T. H. Bayburt, Y. V. Grinkova, and S. G. Sligar. Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Letters*, 2(8):853–856, 8 2002.

[20] T. H. Bayburt and S. G. Sligar. Membrane protein assembly into Nanodiscs. *FEBS Letters*, 584(9):1721–7, 5 2010.

[21] P. A. Beachy, S. G. Hymowitz, R. A. Lazarus, D. J. Leahy, and C. Siebold. Interactions between Hedgehog proteins and their binding partners come into view. *Genes & Development*, 24(18):2001–12, 9 2010.

[22] E. M. Beauchamp, L. Ringer, G. Bulut, K. P. Sajwan, M. D. Hall, Y.-C. Lee, D. Peaceman, M. Ozdemirli, O. Rodriguez, T. J. Macdonald, C. Albanese, J. A. Toretsky, and A. Uren. Arsenic trioxide inhibits human cancer cell growth and tumor development in mice by blocking Hedgehog/GLI pathway. *The Journal of Clinical Investigation*, 121(1):148–60, 1 2011.

[23] J. Bethuyne, S. De Gieter, O. Zwaenepoel, A. Garcia-Pino, K. Durinck, A. Verhelle, G. Hassanzadeh-Ghassabeh, F. Speleman, R. Loris, and J. Gettemans. A nanobody modulates the p53 transcriptional program without perturbing its functional architecture. *Nucleic Acids Research*, 42(20):12928–12938, 10 2014.

[24] C. Bich, S. Maedler, K. Chiesa, F. Degiacomo, N. Bogliotti, and R. Zenobi. Reactivity and applications of new amine reactive cross-linkers for mass spectrometric detection of protein-protein complexes. *Analytical Chemistry*, 82(1):172–179, 1 2010.

[25] M. Bidet, O. Joubert, B. Lacombe, M. Ciantar, R. Nehmé, P. Mollat, L. Brétillon, H. Faure, R. Bittman, M. Ruat, and I. Mus-Veteau. The Hedgehog receptor Patched is involved in cholesterol transport. *PLoS One*, 6(9):e23834, 2011.

[26] M. Bidet, A. Tomico, P. Martin, H. Guizouarn, P. Mollat, and I. Mus-Veteau. The Hedgehog receptor Patched functions in multidrug transport and chemotherapy resistance. *Molecular Cancer Research: MCR*, 10(11):1496–508, 11 2012.

[27] B. Bishop, A. R. Aricescu, K. Harlos, C. A. O'Callaghan, E. Y. Jones, and C. Siebold. Structural insights into Hedgehog ligand sequestration by the human Hedgehog-interacting protein HHIP. *Nature Structural & Molecular Biology*, 16(7):698–703, 7 2009.

[28] R. Blassberg, J. I. Macrae, J. Briscoe, and J. Jacob. Reduced cholesterol levels impair Smoothened activation in Smith–Lemli–Opitz syndrome. *Human Molecular Genetics*, 25(4):693–705, 2 2016.

[29] K. Bloch. The biological synthesis of cholesterol. *Science*, 150(3692):19–28, 10 1965.

[30] I. Bosanac, H. R. Maun, S. J. Scales, X. Wen, A. Lingel, J. F. Bazan, F. J. de Sauvage, S. G. Hymowitz, and R. A. Lazarus. The structure of SHH in complex with HHIP reveals a recognition role for the Shh pseudo active site in signaling. *Nature Structural & Molecular Biology*, 16(7):691–697, 7 2009.

[31] J. Briscoe, Y. Chen, T. M. Jessell, and G. Struhl. A Hedgehog-insensitive form of Patched provides evidence for direct long-range morphogen activity of Sonic hedgehog in the neural tube. *Molecular Cell*, 7(6):1279–1291, 6 2001.

[32] J. Briscoe and P. P. Thérond. The mechanisms of Hedgehog signalling and its roles in development and disease. *Nature Reviews Molecular Cell Biology*, 14(7):416–429, 7 2013.

[33] J. Broecker, B. T. Eger, and O. P. Ernst. Crystallogenesis of membrane proteins mediated by polymer-bounded lipid nanodiscs. *Structure*, 25(2):384–392, 2 2017.

[34] J. A. Buglino and M. D. Resh. Hhat is a palmitoylacyltransferase with specificity for N-palmitoylation of Sonic hedgehog. *The Journal of Biological Chemistry*, 283(32):22076–88, 8 2008.

[35] J. A. Buglino and M. D. Resh. Palmitoylation of Hedgehog proteins. *Vitamins and Hormones*, 88:229–52, 2012.

[36] A. N. Bukiya and A. M. Dopico. Common structural features of cholesterol binding sites in crystallized soluble proteins. *Journal of Lipid Research*, 58(6):1044–1054, 6 2017.

[37] J. S. Burg, J. R. Ingram, A. J. Venkatakrishnan, K. M. Jude, A. Dukkipati, E. N. Feinberg, A. Angelini, D. Waghray, R. O. Dror, H. L. Ploegh, and K. C. Garcia. Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. *Science*, 347(6226):1113–1117, 3 2015.

[38] R. Burke, D. Nellen, M. Bellotto, E. Hafen, K.-A. Senti, B. J. Dickson, and K. Basler. Dispatched, a novel sterol-sensing domain protein dedicated to the release of cholesterol-modified Hedgehog from signaling cells. *Cell*, 99(7):803–815, 12 1999.

[39] E. F. X. Byrne, R. Sircar, P. S. Miller, G. Hedger, G. Luchetti, S. Nachtergaele, M. D. Tully, L. Mydock-McGrane, D. F. Covey, R. P. Rambo, M. S. P. Sansom, S. Newstead, R. Rohatgi, and C. Siebold. Structural basis of Smoothened regulation by its extracellular domains. *Nature*, 535(7613):517–522, 2016.

[40] M. Caffrey. A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes. *Acta Crystallographica. Section F, Structural Biology Communications*, 71(Pt 1):3–18, 1 2015.

[41] D. Carpenter, D. M. Stone, J. Brush, A. Ryan, M. Armanini, G. Frantz, A. Rosenthal, and F. J. de Sauvage. Characterization of two Patched receptors for the vertebrate Hedgehog protein family. *Proceedings of the National Academy of Sciences*, 95(23):13630–4, 11 1998.

[42] E. P. Carpenter, K. Beis, A. D. Cameron, and S. Iwata. Overcoming the challenges of membrane protein crystallography. *Current Opinion in Structural Biology*, 18(5):581–6, 10 2008.

[43] E. D. Carstea, J. A. Morris, K. G. Coleman, S. K. Loftus, D. Zhang, C. Cummings, J. Gu, M. A. Rosenfeld, W. J. Pavan, D. B. Krizman, J. Nagle, M. H. Polymeropoulos, S. L. Sturley, Y. A. Ioannou, M. E. Higgins, M. Comly, A. Cooney, A. Brown, C. R. Kaneski, E. J. Blanchette-Mackie, N. K. Dwyer, E. B. Neufeld, T. Y. Chang, L. Liscum, J. F. Strauss, K. Ohno, M. Zeigler, R. Carmi, J. Sokol, D. Markie, R. R. O'Neill, O. P. van Diggelen, M. Elleder, M. C. Patterson, R. O. Brady, M. T. Vanier, P. G. Pentchev, and D. A. Tagle. Niemann-Pick C1 disease gene: homology to mediators of cholesterol homeostasis. *Science*, 277(5323):228–31, 7 1997.

[44] P. S. Chae, K. Gotfryd, J. Pacyna, L. J. W. Miercke, S. G. F. Rasmussen, R. A. Robbins, R. R. Rana, C. J. Loland, B. Kobilka, R. Stroud, B. Byrne, U. Gether, and S. H. Gellman. Tandem facial amphiphiles for membrane protein stabilization. *Journal of the American Chemical Society*, 132(47):16750–2, 12 2010.

[45] P. S. Chae, S. G. F. Rasmussen, R. R. Rana, K. Gotfryd, R. Chandra, M. A. Goren, A. C. Kruse, S. Nurva, C. J. Loland, Y. Pierre, D. Drew, J.-L. Popot, D. Picot, B. G. Fox, L. Guan, U. Gether, B. Byrne, B. Kobilka, and S. H. Gellman. Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nature Methods*, 7(12):1003–8, 12 2010.

[46] Z. Chamoun, R. K. Mann, D. Nellen, D. P. von Kessler, M. Bellotto, P. A. Beachy, and K. Basler. Skinny hedgehog, an acyltransferase required for palmitoylation and activity of the Hedgehog signal. *Science*, 293(5537):2080–4, 9 2001.

[47] S. Chatterjee and S. Mayor. The GPI-anchor and protein sorting. *Cellular and Molecular Life Sciences*, 58(14):1969–1987, 12 2001.

[48] J. K. Chen, J. Taipale, M. K. Cooper, and P. A. Beachy. Inhibition of Hedgehog signaling by direct binding of cyclopamine to Smoothened. *Genes & Development*, 16(21):2743–8, 11 2002.

[49] J. K. Chen, J. Taipale, K. E. Young, T. Maiti, and P. A. Beachy. Small molecule modulation of Smoothened activity. *Proceedings of the National Academy of Sciences*, 99(22):14071–6, 10 2002.

[50] M.-H. Chen, Y.-J. Li, T. Kawakami, S.-M. Xu, and P.-T. Chuang. Palmitoylation is required for the production of a soluble multimeric Hedgehog protein complex and long-range signaling in vertebrates. *Genes & Development*, 18(6):641–59, 3 2004.

[51] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, and D. C. Richardson. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica. Section D, Biological Crystallography*, 66(Pt 1):12–21, 1 2010.

[52] X. Chen, H. Tukachinsky, C.-H. Huang, C. Jao, Y.-R. Chu, H.-Y. Tang, B. Mueller, S. Schulman, T. A. Rapoport, and A. Salic. Processing and turnover of the Hedgehog protein in the endoplasmic reticulum. *The Journal of Cell Biology*, 192(5):825–38, 3 2011.

[53] X. L. Chen, P. Chinchilla, J. Fombonne, L. Ho, C. Guix, J. H. Keen, P. Mehlen, and N. A. Riobo. Patched-1 proapoptotic activity is downregulated by modification of K1413 by the E3 ubiquitin-protein ligase Itchy homolog. *Molecular and Cellular Biology*, 34(20):3855–66, 10 2014.

[54] Y. Chen, N. Sasai, G. Ma, T. Yue, J. Jia, J. Briscoe, and J. Jiang. Sonic hedgehog dependent phosphorylation by CK1α and GRK2 is required for ciliary accumulation and activation of Smoothened. *PLoS Biology*, 9(6):e1001083, 6 2011.

[55] V. Cherezov, J. Clogston, M. Z. Papiz, and M. Caffrey. Room to move: crystallizing membrane proteins in swollen lipidic mesophases. *Journal of Molecular Biology*, 357(5):1605–1618, 4 2006.

[56] J. H. Chill, J. M. Louis, C. Miller, and A. Bax. NMR study of the tetrameric KcsA potassium channel in detergent micelles. *Protein Science*, 15(4):684–698, 3 2006.

[57] A. E. Christian, M. P. Haynes, M. C. Phillips, and G. H. Rothblat. Use of cyclodextrins for manipulating cellular cholesterol content. *Journal of Lipid Research*, 38(11):2264–72, 11 1997.

[58] P.-T. Chuang and A. P. McMahon. Vertebrate Hedgehog signalling modulated by induction of a Hedgehog-binding protein. *Nature*, 397(6720):617–621, 2 1999.

[59] H. Chug, S. Trakhanov, B. Hülsmann, T. Pleiner, and D. Görlich. Crystal structure of the metazoan Nup62.Nup58.Nup54 nucleoporin complex. *Science*, 350(6256):106–110, 8 2015.

[60] K. Y. Chung, T. H. Kim, A. Manglik, R. Alvares, B. K. Kobilka, and R. S. Prosser. Role of detergents in conformational exchange of a G protein-coupled receptor. *The Journal of Biological Chemistry*, 287(43):36305–11, 10 2012.

[61] M. J. Coloma and S. L. Morrison. Design and production of novel tetravalent bispecific antibodies. *Nature Biotechnology*, 15(2):159–163, 2 1997.

[62] M. K. Cooper, C. A. Wassif, P. A. Krakowiak, J. Taipale, R. Gong, R. I. Kelley, F. D. Porter, and P. A. Beachy. A defective response to Hedgehog signaling in disorders of cholesterol biosynthesis. *Nature Genetics*, 33(4):508–513, 4 2003.

[63] F. Corpet. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research*, 16(22):10881–10890, 11 1988.

[64] A. Creanga, T. D. Glenn, R. K. Mann, A. M. Saunders, W. S. Talbot, and P. A. Beachy. Scube/You activity mediates release of dually lipid-modified Hedgehog signal in soluble form. *Genes & Development*, 26(12):1312–25, 6 2012.

[65] P. Dai, H. Akimaru, Y. Tanaka, T. Maekawa, M. Nakafuku, and S. Ishii. Sonic hedgehog-induced activation of the Gli1 promoter is mediated by GLI3. *The Journal of Biological Chemistry*, 274(12):8143–52, 3 1999.

[66] L. Daury, J.-C. Taveau, D. Salvador, M. Glavier, and O. Lambert. Reconstitution of membrane proteins into nanodiscs for single-particle electron microscopy. In *Methods in Molecular Biology*, pages 317–327. Springer New York, 2017.

[67] J. P. Davies, F. W. Chen, and Y. A. Ioannou. Transmembrane molecular pump activity of Niemann-Pick C1 protein. *Science*, 290(5500):2295–8, 12 2000.

[68] I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall, J. Snoeyink, J. S. Richardson, D. C. Richardson, and D. C. Richardson. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research*, 35(Web Server):W375–83, 7 2007.

[69] E. De Genst, K. Silence, K. Decanniere, K. Conrath, R. Loris, J. Kinne, S. Muyldermans, and L. Wyns. Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proceedings of the National Academy of Sciences*, 103(12):4586–4591, 3 2006.

[70] T. De Meyer, S. Muyldermans, and A. Depicker. Nanobody-based products as research and diagnostic tools. In *Trends in Biotechnology*, volume 32, pages 263–270. Elsevier BV, 5 2014.

[71] J. Deisenhofer, O. Epp, K. Miki, R. Huber, and H. Michel. Structure of the protein subunits in the photosynthetic reaction centre of Rhodopseudomonas viridis at 3Å resolution. *Nature*, 318(6047):618–624, 12 1985.

[72] M. C. Deller, L. Kong, and B. Rupp. Protein stability: a crystallographer's perspective. *Acta Crystallographica. Section F, Structural Biology Communications*, 72(Pt 2):72–95, 2 2016.

[73] N. Denef, D. Neubüser, L. Perez, and S. M. Cohen. Hedgehog induces opposite changes in turnover and subcellular localization of Patched and Smoothened. *Cell*, 102(4):521–531, 8 2000.

[74] I. G. Denisov, Y. V. Grinkova, A. A. Lazarides, and S. G. Sligar. Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size. *Journal of the American Chemical Society*, 126(11):3477–3487, 3 2004.

[75] A. Desmyter, S. Spinelli, C. Boutton, M. Saunders, C. Blachetot, H. de Haard, G. Denecker, M. Van Roy, C. Cambillau, and H. Rommelaere. Neutralization of human interleukin 23 by multivalent nanobodies explained by the structure of cytokine–nanobody complex. *Frontiers in Immunology*, 8:884, 8 2017.

[76] E. Dessaud, A. P. McMahon, J. Briscoe, R. Mo, H. Sasaki, J. Rossant, and C. C. Hui. Pattern formation in the vertebrate neural tube: a Sonic hedgehog morphogen-regulated transcriptional network. *Development*, 135(15):2489–503, 8 2008.

[77] A. DiGiandomenico, A. E. Keller, C. Gao, G. J. Rainey, P. Warrener, M. M. Camara, J. Bonnell, R. Fleming, B. Bezabeh, N. Dimasi, B. R. Sellman, J. Hilliard, C. M. Guenther, V. Datta, W. Zhao, C. Gao, X. Q. Yu, J. A. A. Suzich, and C. K. Stover. A multifunctional bispecific antibody protects against Pseudomonas aeruginosa. *Science Translational Medicine*, 6(262):262ra155, 11 2014.

[78] O. Y. Dmitriev, S. Lutsenko, and S. Muyldermans. Nanobodies as probes for protein dynamics in vitro and in cells. *Journal of Biological Chemistry*, 291(8):3767–3775, 2 2016.

[79] W. Driever and C. Nüsslein-Volhard. A gradient of bicoid protein in Drosophila embryos. *Cell*, 54(1):83–93, 7 1988.

[80] W. Driever and C. Nüsslein-Volhard. The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. *Cell*, 54(1):95–104, 7 1988.

[81] S. Eaton. Release and trafficking of lipid-linked morphogens. *Current Opinion in Genetics & Development*, 16(1):17–22, 2 2006.

[82] S. Eaton. Multiple roles for lipids in the Hedgehog signalling pathway. *Nature Reviews Molecular Cell Biology*, 9(6):437–445, 6 2008.

[83] Y. Echelard, D. J. Epstein, B. St-Jacques, L. Shen, J. Mohler, J. A. McMahon, and A. P. McMahon. Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell*, 75(7):1417–1430, 12 1993.

[84] P. Emsley and K. Cowtan. Coot: model-building tools for molecular graphics. *Acta Crystallographica. Section D, Biological Crystallography*, 60(12):2126–2132, 12 2004.

[85] P. Evans. Scaling and assessment of data quality. *Acta Crystallographica. Section D, Biological Crystallography*, 62(1):72–82, 1 2006.

[86] S. Faham and J. U. Bowie. Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *Journal of Molecular Biology*, 316(1):1–6, 2 2002.

[87] M. Fairhead and M. Howarth. Site-specific biotinylation of purified proteins using BirA. *Methods in Molecular Biology*, 1266:171–84, 2015.

[88] G. Fan, Z. Wang, M. Hao, and J. Li. Bispecific antibodies and their applications. *Journal of Hematology & Oncology*, 8:130, 12 2015.

[89] G. K. Farrington, N. Caram-Salas, A. S. Haqqani, E. Brunette, J. Eldredge, B. Pepinsky, G. Antognetti, E. Baumann, W. Ding, E. Garber, S. Jiang, C. Delaney, E. Boileau, W. P. Sisk, and D. B. Stanimirovic. A novel platform for engineering blood-brain barrier-crossing bispecific biologics. *FASEB Journal*, 28(11):4764–4778, 11 2014.

[90] P. L. Faust and W. J. Kovacs. Cholesterol biosynthesis and ER stress in peroxisome deficiency. *Biochimie*, 98:75–85, 3 2014.

[91] J. Feng, B. White, O. V. Tyurina, B. Guner, T. Larson, H. Y. Lee, R. O. Karlstrom, and J. D. Kohtz. Synergistic and antagonistic roles of the Sonic hedgehog N- and C-terminal lipids. *Development*, 131(17):4357–70, 9 2004.

[92] L. Fiorini, M.-A. Tribalat, L. Sauvard, J. Cazareth, E. Lalli, I. Broutin, O. P. Thomas, and I. Mus-Veteau. Natural paniceins from mediterranean sponge inhibit the multidrug resistance activity of Patched and increase chemotherapy efficiency on melanoma cells. *Oncotarget*, 6(26):22282–97, 9 2015.

[93] J. B. Fitzgerald, B. W. Johnson, J. Baum, S. Adams, S. Iadevaia, J. Tang, V. Rimkunas, L. Xu, N. Kohli, R. Rennard, M. Razlog, Y. Jiao, B. D. Harms, K. J. Olivier,

B. Schoeberl, U. B. Nielsen, and A. A. Lugovskoy. MM-141, an IGF-IR- and ErbB3-directed bispecific antibody, overcomes network adaptations that limit activity of IGF-IR inhibitors. *Molecular Cancer Therapeutics*, 13(2):410–425, 11 2013.

[94] A. Fjose, W. J. McGinnis, and W. J. Gehring. Isolation of a homoeo box-containing gene from the engrailed region of Drosophila and the spatial distribution of its transcripts. *Nature*, 313(6000):284–289, 1 1985.

[95] N. Fuse, T. Maiti, B. Wang, J. A. Porter, T. M. T. Hall, D. J. Leahy, and P. A. Beachy. Sonic hedgehog protein signals not as a hydrolytic enzyme but as an apparent ligand for Patched. *Proceedings of the National Academy of Sciences*, 96(20):10992–10999, 9 1999.

[96] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, and A. Bairoch. Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*, pages 571–607. Humana Press, 2005.

[97] J. Gerhart. 1998 Warkany lecture: Signaling pathways in development. *Teratology*, 60(4):226–239, 10 1999.

[98] G. Gil, J. R. Faust, D. J. Chin, J. L. Goldstein, and M. S. Brown. Membrane-bound domain of HMG CoA reductase is required for sterol-enhanced degradation of the enzyme. *Cell*, 41(1):249–258, 5 1985.

[99] G. Gimpl, K. Burger, and F. Fahrenholz. Cholesterol as modulator of receptor function. *Biochemistry*, 36(36):10959–10974, 9 1997.

[100] S. C. Goetz and K. V. Anderson. The primary cilium: a signalling centre during vertebrate development. *Nature Reviews Genetics*, 11(5):331–44, 5 2010.

[101] X. Gong, H. Qian, P. Cao, X. Zhao, Q. Zhou, J. Lei, and N. Yan. Structural basis for the recognition of Sonic hedgehog by human Patched1. *Science*, 361(6402):eaas8935, 6 2018.

[102] X. Gong, H. Qian, X. Zhou, J. Wu, T. Wan, P. Cao, W. Huang, X. Zhao, X. Wang, P. Wang, Y. Shi, G. F. Gao, Q. Zhou, and N. Yan. Structural insights into the Niemann-Pick C1 (NPC1)-mediated cholesterol transfer and ebola infection. *Cell*, 165(6):1467–1478, 6 2016.

[103] L. V. Goodrich, L. Milenković, K. M. Higgins, and M. P. Scott. Altered neural cell fates and medulloblastoma in mouse Patched mutants. *Science*, 277(5329):1109–13, 8 1997.

[104] P. Graceffa, A. Jancsó, and K. Mabuchi. Modification of acidic residues normalizes sodium dodecyl sulfate-polyacrylamide gel electrophoresis of caldesmon and other proteins that migrate anomalously. *Archives of Biochemistry and Biophysics*, 297(1):46–51, 8 1992.

[105] A. S. Greenberg, A. L. Hughes, J. Guo, D. Avila, E. C. McKinney, and M. F. Flajnik. A novel "chimeric" antibody class in cartilaginous fish: IgM may not be the primordial immunoglobulin. *European Journal of Immunology*, 26(5):1123–1129, 5 1996.

[106] Y. V. Grinkova, I. G. Denisov, and S. G. Sligar. Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers. *Protein Engineering, Design & Selection*, 23(11):843–8, 11 2010.

[107] L. Guan and T. Nakae. Identification of essential charged residues in transmembrane segments of the multidrug transporter MexB of Pseudomonas aeruginosa. *Journal of Bacteriology*, 183(5):1734–9, 3 2001.

[108] R. Guixà-González, J. L. Albasanz, I. Rodriguez-Espigares, M. Pastor, F. Sanz, M. Martí-Solano, M. Manna, H. Martinez-Seara, P. W. Hildebrand, M. Martín, and J. Selent. Membrane cholesterol access into a G-protein-coupled receptor. *Nature Communications*, 8:14505, 2 2017.

[109] R. K. Guy. Inhibition of Sonic hedgehog autoprocessing in cultured mammalian cells by sterol deprivation. *Proceedings of the National Academy of Sciences*, 97(13):7307–12, 6 2000.

[110] I. Habib, D. Smolarek, C. Hattab, M. Grodecka, G. Hassanzadeh-Ghassabeh, S. Muyldermans, S. Sagan, C. Gutiérrez, S. Laperche, C. Le-Van-Kim, Y. C. Aronovicz, K. Wasniowska, S. Gangnard, and O. Bertrand. VHH (nanobody) directed against human glycophorin A: A tool for autologous red cell agglutination assays. *Analytical Biochemistry*, 438(1):82–89, 7 2013.

[111] F. Hagn, M. L. Nasr, and G. Wagner. Assembly of phospholipid nanodiscs of controlled size for structural studies of membrane proteins by NMR. *Nature Protocols*, 13(1):79–98, 12 2017.

[112] H. Hahn, C. Wicking, P. G. Zaphiropoulos, M. R. Gailani, S. Shanley, A. Chidambaram, I. Vorechovsky, E. Holmberg, A. B. Unden, S. Gillies, K. Negus, I. Smyth, C. Pressman, D. J. Leffell, B. Gerrard, A. M. Goldstein, M. Dean, R. Toftgard, G. Chenevix-Trench, B. Wainwright, and A. E. Bale. Mutations of the human homolog of Drosophila Patched in the nevoid basal cell carcinoma syndrome. *Cell*, 85(6):841–851, 6 1996.

[113] T. M. Hall, J. A. Porter, K. E. Young, E. V. Koonin, P. A. Beachy, and D. J. Leahy. Crystal structure of a Hedgehog autoprocessing domain: Homology between Hedgehog and self-splicing proteins. *Cell*, 91(1):85–97, 10 1997.

[114] C. Hamers-Casterman, T. Atarhouch, S. Muyldermans, G. Robinson, C. Hammers, E. B. Songa, N. Bendahman, and R. Hammers. Naturally occurring antibodies devoid of light chains. *Nature*, 363(6428):446–448, 6 1993.

[115] G. Hassanzadeh-Ghassabeh, N. Devoogdt, P. De Pauw, C. Vincke, and S. Muyldermans. Nanobodies and their potential applications. *Nanomedicine*, 8(6):1013–1026, 6 2013.

[116] F. Hauer, C. Gerle, N. Fischer, A. Oshima, K. Shinzawa-Itoh, S. Shimada, K. Yokoyama, Y. Fujiyoshi, and H. Stark. GraDeR: Membrane protein complex preparation for single-particle cryo-EM. *Structure*, 23(9):1769–1775, 9 2015.

[117] G. Hausmann, C. von Mering, and K. Basler. The Hedgehog signaling pathway: where did it come from? *PLoS Biology*, 7(6):e1000146, 6 2009.

[118] C. J. Haycraft, B. Banizs, Y. Aydin-Son, Q. Zhang, E. J. Michaud, and B. K. Yoder. Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function. *PLoS Genetics*, 1(4):e53, 10 2005.

[119] L. Hicke and R. Dunn. Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. *Annual Review of Cell and Developmental Biology*, 19(1):141–172, 11 2003.

[120] J. E. Hooper and M. P. Scott. The Drosophila Patched gene encodes a putative membrane protein required for segmental patterning. *Cell*, 59(4):751–765, 11 1989.

[121] X. Hua, A. Nohturfft, J. L. Goldstein, and M. S. Brown. Sterol resistance in CHO cells traced to point mutation in SREBP cleavage-activating protein. *Cell*, 87(3):415–26, 11 1996.

[122] P. Huang, D. Nedelcu, M. Watanabe, C. Jao, Y. Kim, J. Liu, and A. Salic. Cellular cholesterol directly activates Smoothened in Hedgehog signaling. *Cell*, 166(5):1176–1187, 8 2016.

[123] W. Huang, A. Manglik, A. J. Venkatakrishnan, T. Laeremans, E. N. Feinberg, A. L. Sanborn, H. E. Kato, K. E. Livingston, T. S. Thorsen, R. Kling, S. Granier, P. Gmeiner, S. M. Husbands, J. R. Traynor, W. I. Weis, J. Steyaert, R. O. Dror, and B. K. Kobilka. Structural insights into μ-opioid receptor activation. *Nature*, 524(7565):315–321, 8 2015.

[124] E. Ikonen. Mechanisms for cellular cholesterol transport: defects and human disease. *Physiological Reviews*, 86(4):1237–1261, 10 2006.

[125] E. Ikonen. Cellular cholesterol trafficking and compartmentalization. *Nature Reviews Molecular Cell Biology*, 9(2):125–138, 2 2008.

[126] J. P. Incardona and S. Eaton. Cholesterol in signal transduction. *Current Opinion in Cell Biology*, 12(2):193–203, 4 2000.

[127] R. E. Infante, M. L. Wang, A. Radhakrishnan, H. J. Kwon, M. S. Brown, and J. L. Goldstein. NPC2 facilitates bidirectional transfer of cholesterol between NPC1 and lipid bilayers, a step in cholesterol egress from lysosomes. *Proceedings of the National Academy of Sciences*, 105(40):15287–92, 10 2008.

[128] P. W. Ingham and A. P. McMahon. Hedgehog signaling in animal development: paradigms and principles. *Genes & Development*, 15(23):3059–87, 12 2001.

[129] P. W. Ingham and A. P. McMahon. Hedgehog signalling: Kif7 is not that fishy after all. *Current Biology*, 19(17):R729–R731, 9 2009.

[130] P. A. Insel, B. P. Head, H. H. Patel, D. M. Roth, R. A. Bundey, and J. S. Swaney. Compartmentation of G-protein-coupled receptors and their signalling components in lipid rafts and caveolae. *Biochemical Society Transactions*, 33(Pt 5):1131–4, 11 2005.

[131] A. Ishchenko, E. E. Abola, and V. Cherezov. Crystallization of membrane proteins: An overview. In *Protein Crystallography. Methods in Molecular Biology*, volume 1607, pages 117–141. Humana Press, New York, 2017.

[132] S. Iwata, C. Ostermeier, B. Ludwig, and H. Michel. Structure at 2.8 Å resolution of cytochrome c oxidase from Paracoccus denitrificans. *Nature*, 376(6542):660–669, 8 1995.

[133] L. Izzi, M. Lévesque, S. Morin, D. Laniel, B. C. Wilkes, F. Mille, R. S. Krauss, A. P. McMahon, B. L. Allen, and F. Charron. Boc and Gas1 each form distinct Shh receptor complexes with Ptch1 and are required for Shh-mediated cell proliferation. *Developmental Cell*, 20(6):788–801, 6 2011.

[134] X. Jiang, I. Smirnova, V. Kasho, J. Wu, K. Hirata, M. Ke, E. Pardon, J. Steyaert, N. Yan, and H. R. Kaback. Crystal structure of a LacY–nanobody complex in a periplasmic-open conformation. *Proceedings of the National Academy of Sciences*, 113(44):12420–12425, 10 2016.

[135] E. A. Jones, M. I. Sajid, A. Shenton, and D. G. Evans. Basal Cell Carcinomas in Gorlin Syndrome: A Review of 202 Patients. *Journal of Skin Cancer*, 2011:217378, 2011.

[136] W. Kabsch. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography*, 26(6):795–800, 12 1993.

[137] W. Kabsch. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica. Section D, Biological Crystallography*, 66(Pt 2):133–44, 2 2010.

[138] H. J. Kang, C. Lee, and D. Drew. Breaking the barriers in membrane protein crystallography. *The International Journal of Biochemistry & Cell Biology*, 45(3):636–644, 3 2013.

[139] R. Kaptein and G. Wagner. NMR studies of membrane proteins. *Journal of Biomolecular NMR*, 61(3-4):181–4, 4 2015.

[140] H. E. Karpen, J. T. Bukowski, T. Hughes, J. P. Gratton, W. C. Sessa, and M. R. Gailani. The Sonic hedgehog receptor Patched associates with caveolin-1 in cholesterol-rich microdomains of the plasma membrane. *The Journal of Biological Chemistry*, 276(22):19503–11, 6 2001.

[141] P. A. Karplus and K. Diederichs. Linking crystallographic model and data quality. *Science*, 336(6084):1030–3, 5 2012.

[142] S. Kawamura, K. Hervold, F.-A. Ramirez-Weber, and T. B. Kornberg. Two Patched protein subtypes and a conserved domain of group I proteins that regulates turnover. *The Journal of Biological Chemistry*, 283(45):30964–9, 11 2008.

[143] R. I. Kelley and G. E. Herman. Inborn errors of sterol biosynthesis. *Annual Review of Genomics and Human Genetics*, 2(1):299–341, 9 2001.

[144] H. Khaliullina, D. Panáková, C. Eugster, F. Riedel, M. Carvalho, and S. Eaton. Patched regulates Smoothened trafficking using lipoprotein-derived lipids. *Development*, 136(24):4111–21, 12 2009.

[145] M. Kijanka, B. Dorresteijn, S. Oliveira, and P. M. van Bergen en Henegouwen. Nanobody-based cancer therapy of solid tumors. *Nanomedicine*, 10(1):161–174, 1 2015.

[146] S. Koide. Engineering of recombinant crystallization chaperones. *Current Opinion in Structural Biology*, 19(4):449–457, 8 2009.

[147] R. E. Kontermann and U. Brinkmann. Bispecific antibodies. *Drug Discovery Today*, 20(7):838–847, 7 2015.

[148] A. Kopp, M. A. Muskavitch, and I. Duncan. The roles of Hedgehog and Engrailed in patterning adult abdominal segments of Drosophila. *Development*, 124(19):3703–14, 10 1997.

[149] L. C. Kovari, C. Momany, and M. G. Rossmann. The use of antibody fragments for crystallization and structure determinations. *Structure*, 3(12):1291–1293, 12 1995.

[150] S. Krauss, J.-P. Concordet, and P. Ingham. A functionally conserved homolog of the Drosophila segment polarity gene hh is expressed in tissues with polarizing activity in zebrafish embryos. *Cell*, 75(7):1431–1444, 12 1993.

[151] T. Kromann-Hansen, E. Louise Lange, H. Peter Sørensen, G. Hassanzadeh-Ghassabeh, M. Huang, J. K. Jensen, S. Muyldermans, P. J. Declerck, E. A. Komives, and P. A. Andreasen. Discovery of a novel conformational equilibrium in urokinase-type plasminogen activator. *Scientific Reports*, 7(1):3385, 6 2017.

[152] A. C. Kruse, A. M. Ring, A. Manglik, J. Hu, K. Hu, K. Eitel, H. Hübner, E. Pardon, C. Valant, P. M. Sexton, A. Christopoulos, C. C. Felder, P. Gmeiner, J. Steyaert, W. I. Weis, K. C. Garcia, J. Wess, and B. K. Kobilka. Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature*, 504(7478):101–106, 12 2013.

[153] W. Kulig, J. Tynkkynen, M. Javanainen, M. Manna, T. Rog, I. Vattulainen, and P. Jungwirth. How well does cholesteryl hemisuccinate mimic cholesterol in saturated phospholipid bilayers? *Journal of Molecular Modeling*, 20(2):2121, 2 2014.

[154] P. E. Kuwabara and M. Labouesse. The sterol-sensing domain: multiple families, a unique role? *Trends in Genetics*, 18(4):193–201, 4 2002.

[155] H. J. Kwon, L. Abi-Mosleh, M. L. Wang, J. Deisenhofer, J. L. Goldstein, M. S. Brown, and R. E. Infante. Structure of N-terminal domain of NPC1 reveals distinct subdomains for binding and transfer of cholesterol. *Cell*, 137(7):1213–24, 6 2009.

[156] E. M. Landau and J. P. Rosenbusch. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proceedings of the National Academy of Sciences*, 93(25):14532–5, 12 1996.

[157] R. A. Laskowski. PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Research*, 29(1):221–2, 1 2001.

[158] M. Lauth, A. Bergström, T. Shimokawa, and R. Toftgård. Inhibition of GLI-mediated transcription and tumor cell growth by small-molecule antagonists. *Proceedings of the National Academy of Sciences*, 104(20):8455–60, 5 2007.

[159] J. J. Lee, S. C. Ekker, D. P. von Kessler, J. A. Porter, B. I. Sun, and P. A. Beachy. Autoproteolysis in Hedgehog protein biogenesis. *Science*, 266(5190):1528–37, 12 1994.

[160] R. T. H. Lee, Z. Zhao, and P. W. Ingham. Hedgehog signalling. *Development*, 143(3):367–72, 2 2016.

[161] S. T. Lee, K. D. Welch, K. E. Panter, D. R. Gardner, M. Garrossian, and C.-W. T. Chang. Cyclopamine: from cyclops lambs to cancer treatment. *Journal of Agricultural and Food Chemistry*, 62(30):7355–7362, 7 2014.

[162] R. Lehmann and C. Nüsslein-Volhard. The maternal gene nanos has a central role in posterior pattern formation of the Drosophila embryo. *Development*, 112(3):679–91, 7 1991.

[163] A. G. W. Leslie. The integration of macromolecular diffraction data. *Acta Crystallographica. Section D, Biological Crystallography*, 62(1):48–57, 1 2006.

[164] P. M. Lewis, M. P. Dunn, J. A. McMahon, M. Logan, J. F. Martin, B. St-Jacques, and A. P. McMahon. Cholesterol modification of Sonic hedgehog is required for long-range signaling activity and effective modulation of signaling by Ptc1. *Cell*, 105(5):599–612, 6 2001.

[165] X. Li, F. Lu, M. N. Trinh, P. Schmiege, J. Seemann, J. Wang, and G. Blobel. 3.3 Å structure of Niemann-Pick C1 protein reveals insights into the function of the C-terminal luminal domain in cholesterol transport. *Proceedings of the National Academy of Sciences*, 114(34):9116–9121, 8 2017.

[166] X. Li, P. Mooney, S. Zheng, C. R. Booth, M. B. Braunfeld, S. Gubbens, D. A. Agard, and Y. Cheng. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nature Methods*, 10(6):584–90, 6 2013.

[167] X. Li, J. Wang, E. Coutavas, H. Shi, Q. Hao, and G. Blobel. Structure of human Niemann-Pick C1 protein. *Proceedings of the National Academy of Sciences*, 113(29):8212–7, 7 2016.

[168] J. Liu and B. Rost. Comparing function and structure between entire proteomes. *Protein Science*, 10(10):1970–9, 10 2001.

[169] R. H. Liu, J. Jacob, and B. Tennant. Chemiluminescent detection of protein molecular weight markers in western blot techniques. *BioTechniques*, 22(4):594–5, 4 1997.

[170] S. K. Loftus, J. A. Morris, E. D. Carstea, J. Z. Gu, C. Cummings, A. Brown, J. Ellison, K. Ohno, M. A. Rosenfeld, D. A. Tagle, P. G. Pentchev, and W. J. Pavan. Murine model of Niemann-Pick C disease: mutation in a cholesterol homeostasis gene. *Science*, 277(5323):232–5, 7 1997.

[171] P. J. Loll. Membrane proteins, detergents and crystals: what is the state of the art? *Acta Crystallographica. Section F, Structural Biology Communications*, 70(Pt 12):1576–83, 12 2014.

[172] X. Lu, S. Liu, and T. B. Kornberg. The C-terminal tail of the Hedgehog receptor Patched regulates both localization and turnover. *Genes & Development*, 20(18):2539–51, 9 2006.

[173] G. Luchetti, R. Sircar, J. H. Kong, S. Nachtergaele, A. Sagner, E. F. Byrne, D. F. Covey, C. Siebold, and R. Rohatgi. Cholesterol activates the G-protein coupled receptor Smoothened to promote Hedgehog signaling. *eLife*, 5, 10 2016.

[174] S. J. Ludtke, I. I. Serysheva, S. L. Hamilton, and W. Chiu. The pore structure of the closed RyR1 channel. *Structure*, 13(8):1203–11, 8 2005.

[175] L. Lum and P. A. Beachy. The Hedgehog response network: sensors, switches, and routers. *Science*, 304(5678):1755–9, 6 2004.

[176] Y. Ma, A. Erkner, R. Gong, S. Yao, J. Taipale, K. Basler, and P. A. Beachy. Hedgehog-mediated patterning of the mammalian embryo requires transporter-like function of dispatched. *Cell*, 111(1):63–75, 10 2002.

[177] R. K. Maeda, F. Karch, C. Burks, and M. Levine. The ABC of the BX-C: the bithorax complex explained. *Development*, 133(8):1413–22, 4 2006.

[178] A. Manglik, B. K. Kobilka, and J. Steyaert. Nanobodies to study G protein-coupled receptor structure and function. *Annual Review of Pharmacology and Toxicology*, 57(1):19–37, 1 2017.

[179] V. Marigo, R. A. Davey, Y. Zuo, J. M. Cunningham, and C. J. Tabin. Biochemical evidence that Patched is the Hedgehog receptor. *Nature*, 384(6605):176–179, 11 1996.

[180] V. Marigo, R. L. Johnson, A. Vortkamp, and C. J. Tabin. Sonic hedgehog differentially regulates expression of GLI and GLI3 during limb development. *Developmental Biology*, 180(1):273–283, 11 1996.

[181] V. Martın, G. Carrillo, C. Torroja, and I. Guerrero. The sterol-sensing domain of Patched protein seems to control Smoothened activity through Patched vesicular trafficking. *Current Biology*, 11(8):601–607, 4 2001.

[182] H. R. Maun, X. Wen, A. Lingel, F. J. de Sauvage, R. A. Lazarus, S. J. Scales, and S. G. Hymowitz. Hedgehog pathway antagonist 5E1 binds Hedgehog at the pseudo-active site. *The Journal of Biological Chemistry*, 285(34):26570–80, 8 2010.

[183] F. R. Maxfield and I. Tabas. Role of cholesterol and lipid organization in disease. *Nature*, 438(7068):612–621, 12 2005.

[184] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read. Phaser crystallographic software. *Journal of Applied Crystallography*, 40(Pt 4):658–674, 2007.

[185] C. F. McDonagh, A. Huhalov, B. D. Harms, S. Adams, V. Paragas, S. Oyama, B. Zhang, L. Luus, R. Overland, S. Nguyen, J. Gu, N. Kohli, M. Wallace, M. J. Feldhaus, A. J. Kudla, B. Schoeberl, and U. B. Nielsen. Antitumor activity of a novel bispecific antibody that targets the ErbB2/ErbB3 oncogenic unit and inhibits Heregulin-induced activation of ErbB3. *Molecular Cancer Therapeutics*, 11(3):582–593, 1 2012.

[186] J. S. McLellan, X. Zheng, G. Hauk, R. Ghirlando, P. A. Beachy, and D. J. Leahy. The mode of Hedgehog binding to Ihog homologues is not conserved across different phyla. *Nature*, 455(7215):979–83, 10 2008.

[187] A. Merk, A. Bartesaghi, S. Banerjee, V. Falconieri, P. Rao, M. I. Davis, R. Pragani, M. B. Boxer, L. A. Earl, J. L. S. Milne, and S. Subramaniam. Breaking cryo-EM resolution barriers to facilitate drug discovery. *Cell*, 165(7):1698–1707, 6 2016.

[188] L. Milenkovic, M. P. Scott, and R. Rohatgi. Lateral transport of Smoothened from the plasma membrane to the membrane of the cilium. *The Journal of Cell Biology*, 187(3):365–74, 11 2009.

[189] J. A. Mindell and N. Grigorieff. Accurate determination of local defocus and specimen tilt in electron microscopy. *Journal of Structural Biology*, 142(3):334–347, 6 2003.

[190] S. A. Mireku, M. M. Sauer, R. Glockshuber, and K. P. Locher. Structural basis of nanobody-mediated blocking of BtuF, the cognate substrate-binding protein of the Escherichia coli vitamin B12 transporter BtuCD. *Scientific Reports*, 7(1):14296, 10 2017.

[191] I. Moraes, G. Evans, J. Sanchez-Weatherby, S. Newstead, and P. D. S. Stewart. Membrane protein structure determination - the next generation. *Biochimica et Biophysica Acta*, 1838(1 Pt A):78–87, 1 2014.

[192] S. Mulinari, U. Häcker, P. W. Ingham, X. J. Liu, and J. A. Lengyel. Hedgehog, but not Odd skipped, induces segmental grooves in the Drosophila epidermis. *Development*, 136(23):3875–80, 12 2009.

[193] J. L. Mullor and I. Guerrero. A gain-of-function mutant of Patched dissects different responses to the Hedgehog gradient. *Developmental Biology*, 228(2):211–224, 12 2000.

[194] S. Murakami, R. Nakashima, E. Yamashita, T. Matsumoto, and A. Yamaguchi. Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature*, 443(7108):173–179, 8 2006.

[195] S. Murakami, R. Nakashima, E. Yamashita, and A. Yamaguchi. Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature*, 419(6907):587–593, 10 2002.

[196] G. N. Murshudov, A. A. Vagin, and E. J. Dodson. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica. Section D, Biological Crystallography*, 53(3):240–255, 5 1997.

[197] S. Muyldermans, C. Cambillau, and L. Wyns. Recognition of antigens by single-domain antibody fragments: the superfluous luxury of paired domains. *Trends in Biochemical Sciences*, 26(4):230–235, 4 2001.

[198] S. Muyldermans and M. Lauwereys. Unique single-domain antigen binding fragments derived from naturally occurring camel heavy-chain antibodies. *Journal of Molecular Recognition*, 12(2):131–140, 3 1999.

[199] M. Nakamura, M. Kubo, K. Yanai, Y. Mikami, M. Ikebe, S. Nagai, K. Yamaguchi, M. Tanaka, and M. Katano. Anti-Patched-1 antibodies suppress Hedgehog signaling pathway and pancreatic cancer proliferation. *Anticancer Research*, 27(6A):3743–7, 11 2007.

[200] M. Nakamura, H. Tanaka, Y. Nagayoshi, H. Nakashima, K. Tsutsumi, T. Ohtsuka, S. Takahata, M. Tanaka, and H. Okada. Targeting the Hedgehog signaling pathway with interacting peptides to Patched-1. *Journal of Gastroenterology*, 47(4):452–460, 4 2012.

[201] Y. Nakano, I. Guerrero, A. Hidalgo, A. Taylor, J. R. S. Whittle, and P. W. Ingham. A protein with several possible membrane-spanning domains encoded by the Drosophila segment polarity gene Patched. *Nature*, 341(6242):508–513, 10 1989.

[202] J. A. Newman, P. Savitsky, C. K. Allerston, A. H. Bizard, Özer, K. Sarlós, Y. Liu, E. Pardon, J. Steyaert, I. D. Hickson, and O. Gileadi. Crystal structure of the Bloom's syndrome helicase indicates a role for the HRDC domain in conformational changes. *Nucleic Acids Research*, 43(10):5221–5235, 5 2015.

[203] H. Nikaido and Y. Takatsuka. Mechanisms of RND multidrug efflux pumps. *Biochimica et Biophysica Acta*, 1794(5):769–81, 5 2009.

[204] F. Noël, A. Malpertuy, and A. G. de Brevern. Global analysis of VHHs framework regions with a structural alphabet. *Biochimie*, 131:11–19, 12 2016.

[205] C. Nüsslein-Volhard and E. Wieschaus. Mutations affecting segment number and polarity in Drosophila. *Nature*, 287(5785):795–801, 10 1980.

[206] C. Ostermeier, S. Iwata, B. Ludwig, and H. Michel. Fv fragment-mediated crystallization of the membrane protein bacterial cytochrome c oxidase. *Nature Structural Biology*, 2(10):842–846, 10 1995.

[207] W. Palm, M. M. Swierczynska, V. Kumari, M. Ehrhart-Bornstein, S. R. Bornstein, and S. Eaton. Secretion and signaling activities of lipoprotein-associated Hedgehog and non-sterol-modified Hedgehog in flies and mammals. *PLoS Biology*, 11(3):e1001505, 2013.

[208] Y. Pan, C. Wang, and B. Wang. Phosphorylation of Gli2 by protein kinase A is required for Gli2 processing and degradation and the Sonic hedgehog-regulated mouse development. *Developmental Biology*, 326(1):177–89, 2 2009.

[209] D. Panáková, H. Sprong, E. Marois, C. Thiele, and S. Eaton. Lipoprotein particles are required for Hedgehog and Wingless signalling. *Nature*, 435(7038):58–65, 5 2005.

[210] T. Pape and T. R. Schneider. HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *Journal of Applied Crystallography*, 37(5):843–844, 10 2004.

[211] E. Pardon, T. Laeremans, S. Triest, S. G. F. Rasmussen, A. Wohlkönig, A. Ruf, S. Muyldermans, W. G. J. Hol, B. K. Kobilka, and J. Steyaert. A general protocol for the generation of nanobodies for structural biology. *Nature Protocols*, 9(3):674–93, 3 2014.

[212] J. L. Parker and S. Newstead. Current trends in α-helical membrane protein crystallization: an update. *Protein Science*, 21(9):1358–65, 9 2012.

[213] J. L. Parker and S. Newstead. Membrane protein crystallisation: current trends and future perspectives. *Advances in Experimental Medicine and Biology*, 922:61–72, 2016.

[214] M. Parmar, S. Rawson, C. A. Scarff, A. Goldman, T. R. Dafforn, S. P. Muench, and V. L. Postis. Using a SMALP platform to determine a sub-nm single particle cryo-EM membrane protein structure. *Biochimica et Biophysica Acta - Biomembranes*, 1860(2):378–383, 2 2018.

[215] S. Pathi, S. Pagan-Westphal, D. P. Baker, E. A. Garber, P. Rayhorn, D. Bumcrot, C. J. Tabin, R. Blake Pepinsky, and K. P. Williams. Comparative biological responses to human Sonic, Indian, and Desert hedgehog. *Mechanisms of Development*, 106(1-2):107–117, 8 2001.

[216] R. B. Pepinsky, P. Rayhorn, E. S. Day, A. Dergay, K. P. Williams, A. Galdes, F. R. Taylor, P. A. Boriack-Sjodin, and E. A. Garber. Mapping Sonic hedgehog-receptor interactions by steric interference. *The Journal of Biological Chemistry*, 275(15):10995–1001, 4 2000.

[217] R. B. Pepinsky, C. Zeng, D. Wen, P. Rayhorn, D. P. Baker, K. P. Williams, S. A. Bixler, C. M. Ambrose, E. A. Garber, K. Miatkowski, F. R. Taylor, E. A. Wang, and A. Galdes. Identification of a palmitic acid-modified form of human Sonic hedgehog. *The Journal of Biological Chemistry*, 273(22):14037–45, 5 1998.

[218] C. Perez, M. Köhler, D. Janser, E. Pardon, J. Steyaert, R. Zenobi, and K. P. Locher. Structural basis of inhibition of lipid-linked oligosaccharide flippase PglK by a conformational nanobody. *Scientific Reports*, 7(1):46641, 4 2017.

[219] J. D. Perlmutter, W. J. Drasler, W. Xie, J. Gao, J.-L. Popot, J. N. Sachs, and J. N. Sachs. All-atom and coarse-grained molecular dynamics simulations of a membrane protein stabilizing polymer. *Langmuir*, 27(17):10523–37, 9 2011.

[220] C. Peters, A. Wolf, M. Wagner, J. Kuhlmann, and H. Waldmann. The cholesterol membrane anchor of the Hedgehog protein confers stable membrane association to lipid-modified proteins. *Proceedings of the National Academy of Sciences*, 101(23):8531–6, 6 2004.

[221] E. Petrova, J. Rios-Esteves, O. Ouerfelli, J. F. Glickman, and M. D. Resh. Inhibitors of Hedgehog acyltransferase block Sonic hedgehog signaling. *Nature Chemical Biology*, 9(4):247–9, 4 2013.

[222] F. Peyvandi, M. Scully, J. A. Kremer Hovinga, S. Cataland, P. Knöbl, H. Wu, A. Artoni, J.-P. Westwood, M. Mansouri Taleghani, B. Jilma, F. Callewaert, H. Ulrichts, C. Duby, and D. Tersago. Caplacizumab for acquired thrombotic thrombocytopenic purpura. *The New England Journal of Medicine*, 374(6):511–522, 2 2016.

[223] A. Pollithy, T. Romer, C. Lang, F. D. Müller, J. Helma, H. Leonhardt, U. Rothbauer, and D. Schüler. Magnetosome expression of functional camelid antibody fragments (nanobodies) in Magnetospirillum gryphiswaldense. *Applied and Environmental Microbiology*, 77(17):6165–6171, 7 2011.

[224] F. D. Porter. Human malformation syndromes due to inborn errors of cholesterol synthesis. *Current Opinion in Pediatrics*, 15(6):607–13, 12 2003.

[225] J. A. Porter, K. E. Young, and P. A. Beachy. Cholesterol modification of Hedgehog signaling proteins in animal development. *Science*, 274(5285):255–9, 10 1996.

[226] A. Punjani, J. L. Rubinstein, D. J. Fleet, and M. A. Brubaker. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, 14(3):290–296, 3 2017.

[227] X. Qi, P. Schmiege, E. Coutavas, and X. Li. Two Patched molecules engage distinct sites on Hedgehog yielding a signaling-competent complex. *Science*, 362(6410):eaas8843, 8 2018.

[228] X. Qi, P. Schmiege, E. Coutavas, J. Wang, and X. Li. Structures of human Patched and its complex with native palmitoylated Sonic hedgehog. *Nature*, 560(7716):128–132, 8 2018.

[229] S. G. F. Rasmussen, H.-J. Choi, J. J. Fung, E. Pardon, P. Casarosa, P. S. Chae, B. T. Devree, D. M. Rosenbaum, F. S. Thian, T. S. Kobilka, A. Schnapp, I. Konetzki, R. K. Sunahara, S. H. Gellman, A. Pautsch, J. Steyaert, W. I. Weis, and B. K. Kobilka. Structure of a nanobody-stabilized active state of the $\beta 2$ adrenoceptor. *Nature*, 469(7329):175–180, 1 2011.

[230] S. Rawson, S. Davies, J. D. Lippiat, and S. P. Muench. The changing landscape of membrane protein structural biology through developments in electron microscopy. *Molecular Membrane Biology*, 33(1-2):12–22, 2016.

[231] R. D. Riddle, R. L. Johnson, E. Laufer, and C. Tabin. Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell*, 75(7):1401–1416, 12 1993.

[232] A. Rietveld, S. Neutz, K. Simons, and S. Eaton. Association of sterol- and glycosylphosphatidylinositol-linked proteins with Drosophila raft lipid microdomains. *The Journal of Biological Chemistry*, 274(17):12049–54, 4 1999.

[233] T. K. Rimkus, R. L. Carpenter, S. Qasem, M. Chan, and H.-W. Lo. Targeting the Sonic hedgehog signaling pathway: review of Smoothened and GLI inhibitors. *Cancers*, 8(2):22, 2 2016.

[234] X. Robert and P. Gouet. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, 42(W1):W320–W324, 4 2014.

[235] U. R. Rodgers, T. Lanyon-Hogg, N. Masumoto, M. Ritzefeld, R. Burke, J. Blagg, A. I. Magee, and E. W. Tate. Characterization of Hedgehog acyltransferase inhibitors identifies a small molecule probe for Hedgehog signaling by cancer cells. *ACS Chemical Biology*, 11(12):3256–3262, 10 2016.

[236] E. Roessler, E. Belloni, K. Gaudenz, P. Jay, P. Berta, S. W. Scherer, L.-C. Tsui, and M. Muenke. Mutations in the human Sonic hedgehog gene cause holoprosencephaly. *Nature Genetics*, 14(3):357–360, 11 1996.

[237] E. Roessler, E. Belloni, K. Gaudenz, F. Vargas, S. W. Scherer, L.-C. Tsui, and M. Muenke. Mutations in the C-terminal domain of Sonic hedgehog cause holoprosencephaly. *Human Molecular Genetics*, 6(11):1847–1853, 10 1997.

[238] R. Rohatgi, L. Milenkovic, M. P. Scott, K. Xiao, J. Kim, M. Chen, J. Wang, W. Chen, and R. J. Lefkowitz. Patched1 regulates Hedgehog signaling at the primary cilium. *Science*, 317(5836):372–376, 7 2007.

[239] A. Rohou and N. Grigorieff. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology*, 192(2):216–221, 11 2015.

[240] K. Rostislavleva, N. Soler, Y. Ohashi, L. Zhang, E. Pardon, J. E. Burke, G. R. Masson, C. Johnson, J. Steyaert, N. T. Ktistakis, and R. L. Williams. Structure and flexibility of the endosomal Vps34 complex reveals the basis of its function on membranes. *Science*, 350(6257):aac7365, 10 2015.

[241] U. Rothbauer, K. Zolghadr, S. Muyldermans, A. Schepers, M. C. Cardoso, and H. Leonhardt. A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Molecular & Cellular Proteomics*, 7(2):282–289, 10 2008.

[242] L. L. Rubin and F. J. de Sauvage. Targeting the Hedgehog pathway in cancer. *Nature Reviews Drug Discovery*, 5(12):1026–1033, 12 2006.

[243] C. M. Rudin, C. L. Hann, J. Laterra, R. L. Yauch, C. A. Callahan, L. Fu, T. Holcomb, J. Stinson, S. E. Gould, B. Coleman, P. M. LoRusso, D. D. Von Hoff, F. J. de Sauvage, and J. A. Low. Treatment of medulloblastoma with Hedgehog pathway inhibitor GDC-0449. *The New England Journal of Medicine*, 361(12):1173–8, 9 2009.

[244] G. S. Salama, M. A. Kaabneh, M. K. Al-Raqad, I. M. Al-Abdallah, A. G. Shakkoury, and R. A. Halaseh. Cyclopia: a rare condition with unusual presentation - a case report. *Clinical Medicine Insights. Pediatrics*, 9:19–23, 2015.

[245] H. Sasaki, Y. Nishizaki, C. Hui, M. Nakafuku, and H. Kondoh. Regulation of Gli2 and Gli3 activities by an amino-terminal repression domain: implication of Gli2 and Gli3 as primary mediators of Shh signaling. *Development*, 126(17):3915–24, 9 1999.

[246] S. H. Scheres. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, 180(3):519–530, 12 2012.

[247] M. D. Schroeder, M. Pearce, J. Fak, H. Fan, U. Unnerstall, E. Emberly, N. Rajewsky, E. D. Siggia, and U. Gaul. Transcriptional control in the segmentation gene network of Drosophila. *PLoS Biology*, 2(9):e271, 8 2004.

[248] A. F. Schubert, C. Gladkova, E. Pardon, J. L. Wagstaff, S. M. Freund, J. Steyaert, S. L. Maslen, and D. Komander. Structure of PINK1 in complex with its substrate ubiquitin. *Nature*, 552(7683):51–56, 12 2017.

[249] S. E. Sedykh, V. V. Prinz, V. N. Buneva, and G. A. Nevinsky. Bispecific antibodies: Design, therapy, perspectives. *Drug Design, Development and Therapy*, 12:195–208, 1 2018.

[250] G. M. Sheldrick. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallographica. Section D, Biological Crystallography*, 66(Pt 4):479–85, 4 2010.

[251] D. Shi, X. Lv, Z. Zhang, X. Yang, Z. Zhou, L. Zhang, and Y. Zhao. Smoothened oligomerization/higher order clustering in lipid rafts is essential for high Hedgehog activity transduction. *The Journal of Biological Chemistry*, 288(18):12605–14, 5 2013.

[252] F. Y. Siu, M. He, C. de Graaf, G. W. Han, D. Yang, Z. Zhang, C. Zhou, Q. Xu, D. Wacker, J. S. Joseph, W. Liu, J. Lau, V. Cherezov, V. Katritch, M.-W. Wang, and R. C. Stevens. Structure of the human glucagon class B G-protein-coupled receptor. *Nature*, 499(7459):444–449, 7 2013.

[253] S. M. Smith. Strategies for the purification of membrane proteins. In *Methods in Molecular Biology*, pages 485–496. Humana Press, 10 2011.

[254] H. Spemann and H. Mangold. Induction of embryonic primordia by implantation of organizers from a different species. 1923. *The International Journal of Developmental Biology*, 45(1):13–38, 2 2001.

[255] C. Spiess, Q. Zhai, and P. J. Carter. Alternative molecular formats and therapeutic applications for bispecific antibodies. *Molecular Immunology*, 67(2):95–106, 10 2015.

[256] A. Spirov, K. Fahmy, M. Schneider, E. Frei, M. Noll, and S. Baumgartner. Formation of the bicoid morphogen gradient: an mRNA gradient dictates the protein gradient. *Development*, 136(4):605–14, 2 2009.

[257] B. Z. Stanton, L. F. Peng, N. Maloof, K. Nakai, X. Wang, J. L. Duffner, K. M. Taveras, J. M. Hyman, S. W. Lee, A. N. Koehler, J. K. Chen, J. L. Fox, A. Mandinova, S. L. Schreiber, and S. L. Schreiber. A small molecule that binds Hedgehog and blocks its signaling in human cells. *Nature Chemical Biology*, 5(3):154–6, 3 2009.

[258] S. Steeland, R. E. Vandenbroucke, and C. Libert. Nanobodies as therapeutics: big opportunities for small antibodies. *Drug Discovery Today*, 21(7):1076–1113, 7 2016.

[259] A. Stetsenko and A. Guskov. An overview of the top ten detergents used for membrane protein crystallization. *Crystals*, 7(7):197, 7 2017.

[260] D. M. Stone, M. Hynes, M. Armanini, T. A. Swanson, Q. Gu, R. L. Johnson, M. P. Scott, D. Pennica, A. Goddard, H. Phillips, M. Noll, J. E. Hooper, F. de Sauvage, and A. Rosenthal. The tumour-suppressor gene Patched encodes a candidate receptor for Sonic hedgehog. *Nature*, 384(6605):129–134, 11 1996.

[261] H. Strutt, C. Thomas, Y. Nakano, D. Stark, B. Neave, A. Taylor, and P. Ingham. Mutations in the sterol-sensing domain of Patched suggest a role for vesicular trafficking in Smoothened regulation. *Current Biology*, 11(8):608–613, 4 2001.

[262] T. Tabata and T. B. Kornberg. Hedgehog is a signaling protein with a key role in patterning Drosophila imaginal discs. *Cell*, 76(1):89–102, 1 1994.

[263] O. Tacar, P. Sriamornsak, and C. R. Dass. Doxorubicin: an update on anticancer molecular action, toxicity and novel drug delivery systems. *Journal of Pharmacy and Pharmacology*, 65(2):157–170, 2 2013.

[264] J. Taipale, M. K. Cooper, T. Maiti, and P. A. Beachy. Erratum: Patched acts catalytically to suppress the activity of Smoothened. *Nature*, 418(6900):892–896, 8 2002.

[265] F. R. Taylor, D. Wen, E. A. Garber, A. N. Carmillo, D. P. Baker, R. M. Arduini, K. P. Williams, P. H. Weinreb, P. Rayhorn, X. Hronowski, A. Whitty, E. S. Day,

A. Boriack-Sjodin, R. I. Shapiro, A. Galdes, and R. B. Pepinsky. Enhanced potency of human Sonic hedgehog by hydrophobic modification. *Biochemistry*, 40(14):4359–4371, 4 2001.

[266] L. Thomas, J. Kahr, P. Schmidt, U. Krug, H. A. Scheidt, and D. Huster. The dynamics of the G protein-coupled neuropeptide Y2 receptor in monounsaturated membranes investigated by solid-state NMR spectroscopy. *Journal of Biomolecular NMR*, 61(3-4):347–359, 4 2015.

[267] N. Thonghin, V. Kargas, J. Clews, and R. C. Ford. Cryo-electron microscopy of membrane proteins. *Methods*, 147:176–186, 9 2018.

[268] C. Tribet, R. Audebert, and J. L. Popot. Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proceedings of the National Academy of Sciences*, 93(26):15047–50, 12 1996.

[269] T. N. Trinh, E. A. McLaughlin, C. P. Gordon, and A. McCluskey. Hedgehog signalling pathway inhibitors as cancer suppressing agents. *Medicinal Chemistry Communication*, 5(2):117–133, 1 2014.

[270] T. T. Tseng, K. S. Gratwick, J. Kollman, D. Park, D. H. Nies, A. Goffeau, and M. H. Saier. The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *Journal of Molecular Microbiology and Biotechnology*, 1(1):107–25, 8 1999.

[271] H. Tukachinsky, R. P. Kuzmickas, C. Y. Jao, J. Liu, and A. Salic. Dispatched and Scube mediate the efficient secretion of the cholesterol-modified Hedgehog ligand. *Cell Reports*, 2(2):308–20, 8 2012.

[272] H. Tukachinsky, K. Petrov, M. Watanabe, and A. Salic. Mechanism of inhibition of the tumor suppressor Patched by Sonic hedgehog. *Proceedings of the National Academy of Sciences*, 113(40):E5866–E5875, 10 2016.

[273] A. M. Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 237(641):37–72, 8 1952.

[274] M. W. Turner, R. Cruz, J. Mattos, N. Baughman, J. Elwell, J. Fothergill, A. Nielsen, J. Brookhouse, A. Bartlett, P. Malek, X. Pu, M. D. King, and O. M. McDougal. Cyclopamine bioactivity by extraction method from Veratrum californicum. *Bioorganic & Medicinal Chemistry*, 24(16):3752–7, 2016.

[275] I. Van Audenhove and J. Gettemans. Nanobodies as versatile tools to understand, diagnose, visualize and treat cancer. *EBioMedicine*, 8:40–48, 6 2016.

[276] G. R. van den Brink. Hedgehog signaling in development and homeostasis of the gastrointestinal tract. *Physiological Reviews*, 87(4):1343–1375, 10 2007.

[277] I. Vaneycken, N. Devoogdt, N. Van Gassen, C. Vincke, C. Xavier, U. Wernery, S. Muyldermans, T. Lahoutte, and V. Caveliers. Preclinical screening of anti-HER2 nanobodies for molecular imaging of breast cancer. *The FASEB Journal*, 25(7):2433–2446, 7 2011.

[278] M. T. Vanier. Complex lipid trafficking in Niemann-Pick disease type C. *Journal of Inherited Metabolic Disease*, 38(1):187–199, 1 2015.

[279] M. Varjosalo and J. Taipale. Hedgehog: functions and mechanisms. *Genes & Development*, 22(18):2454–72, 9 2008.

[280] M. Vosjan, J. Vercammen, J. A. Kolkman, M. Stigter-van Walsum, H. Revets, and G. van Dongen. Nanobodies targeting the hepatocyte growth factor: potential new drugs for molecular cancer therapy. *Molecular Cancer Therapeutics*, 11(4):1017–25, 4 2012.

[281] N. Vyas, A. Walvekar, D. Tate, V. Lakshmanan, D. Bansal, A. Lo Cicero, G. Raposo, D. Palakodeti, and J. Dhawan. Vertebrate Hedgehog is secreted on two types of extracellular vesicles with different signaling properties. *Scientific Reports*, 4:7357, 12 2014.

[282] P. Wadsten, A. B. Wöhri, A. Snijder, G. Katona, A. T. Gardiner, R. J. Cogdell, R. Neutze, and S. Engström. Lipidic sponge phase crystallization of membrane proteins. *Journal of Molecular Biology*, 364(1):44–53, 11 2006.

[283] A. Wagner, R. Duman, K. Henderson, and V. Mykhaylyk. In-vacuum long-wavelength macromolecular crystallography. *Acta Crystallographica. Section D, Structural Biology*, 72(3):430–439, 3 2016.

[284] A. C. Wallace, R. A. Laskowski, and J. M. Thornton. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering*, 8(2):127–34, 2 1995.

[285] T. S. Walter, J. M. Diprose, C. J. Mayo, C. Siebold, M. G. Pickford, L. Carter, G. C. Sutton, N. S. Berrow, J. Brown, I. M. Berry, G. B. E. Stewart-Jones, J. M. Grimes, D. K. Stammers, R. M. Esnouf, E. Y. Jones, R. J. Owens, D. I. Stuart, K. Harlos, and IUCr. A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization. *Acta Crystallographica. Section D, Biological Crystallography*, 61(6):651–657, 6 2005.

[286] M. S. Weiss. Global indicators of X-ray data quality. *Journal of Applied Crystallography*, 34(2):130–135, 4 2001.

[287] D. M. Whalen, T. Malinauskas, R. J. C. Gilbert, and C. Siebold. Structural insights into proteoglycan-shaped Hedgehog signaling. *Proceedings of the National Academy of Sciences*, 110(41):16420–5, 10 2013.

[288] L. Wolpert. Positional information and the spatial pattern of cellular differentiation. *Journal of Theoretical Biology*, 25(1):1–47, 10 1969.

[289] H. Wu, C. Wang, K. J. Gregory, G. W. Han, H. P. Cho, Y. Xia, C. M. Niswender, V. Katritch, J. Meiler, V. Cherezov, P. J. Conn, and R. C. Stevens. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science*, 344(6179):58–64, 4 2014.

[290] Z. Yan, X. Bai, C. Yan, J. Wu, Z. Li, T. Xie, W. Peng, C. Yin, X. Li, S. H. W. Scheres, Y. Shi, and N. Yan. Structure of the rabbit ryanodine receptor RyR1 at near-atomic resolution. *Nature*, 517(7532):50–55, 1 2015.

[291] P. G. Yancey, W. V. Rodrigueza, E. P. Kilsdonk, G. W. Stoudt, W. J. Johnson, M. C. Phillips, and G. H. Rothblat. Cellular cholesterol efflux mediated by cyclodextrins. Demonstration of kinetic pools and mechanism of efflux. *The Journal of Biological Chemistry*, 271(27):16026–34, 7 1996.

[292] Z.-J. Yang, T. Ellis, S. L. Markant, T.-A. Read, J. D. Kessler, M. Bourboulas, U. Schüller, R. Machold, G. Fishell, D. H. Rowitch, B. J. Wainwright, and R. J. Wechsler-Reya. Medulloblastoma can be initiated by deletion of Patched in lineage-restricted progenitors or stem cells. *Cancer Cell*, 14(2):135–45, 8 2008.

[293] R. L. Yauch, S. E. Gould, S. J. Scales, T. Tang, H. Tian, C. P. Ahn, D. Marshall, L. Fu, T. Januario, D. Kallop, M. Nannini-Pepe, K. Kotkow, J. C. Marsters, L. L. Rubin, and F. J. de Sauvage. A paracrine requirement for Hedgehog signalling in cancer. *Nature*, 455(7211):406–410, 9 2008.

[294] S. Yue, L.-Y. Tang, Y. Tang, Y. Tang, Q.-H. Shen, J. Ding, Y. Chen, Z. Zhang, T.-T. Yu, Y. E. Zhang, and S. Y. Cheng. Requirement of Smurf-mediated endocytosis of Patched1 in Sonic hedgehog signal reception. *eLife*, 3:e02555, 6 2014.

[295] R. Zalk, O. B. Clarke, A. des Georges, R. A. Grassucci, S. Reiken, F. Mancia, W. A. Hendrickson, J. Frank, and A. R. Marks. Structure of a mammalian ryanodine receptor. *Nature*, 517(7532):44–49, 1 2015.

[296] X. Zeng, J. A. Goetz, L. M. Suber, W. J. Scott, C. M. Schreiner, and D. J. Robbins. A freely diffusible form of Sonic hedgehog mediates long-range signalling. *Nature*, 411(6838):716–720, 6 2001.

[297] Q. Zhang, X. Ma, A. Ward, W.-X. Hong, V.-P. Jaakola, R. Stevens, M. Finn, and G. Chang. Designing facial amphiphiles for the stabilization of integral membrane proteins. *Angewandte Chemie International Edition*, 46(37):7023–7025, 9 2007.

[298] M. Zoonens and J.-L. Popot. Amphipols for each season. *The Journal of Membrane Biology*, 247(9-10):759–96, 10 2014.