

Towards a structure-based understanding of WNT receptor function

James Hillier

**A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of
Philosophy at the University of Oxford**

Linacre College

Nuffield Department of Clinical Medicine

Division of Structural Biology

University of Oxford

Trinity Term 2019



Abstract

The highly conserved WNT signalling pathways are vital during development and throughout the lives of animals, while aberrant WNT signalling is implicated in a range of degenerative diseases, as well as cancers. Cells detect WNT signals through members of the Frizzled (FZD₁₋₁₀) family of cell-surface receptor proteins, resulting in multiple possible outcomes including changes in gene expression and cell polarity. However, how WNT-FZD specificity is achieved, and how different WNT signalling modules are differentially activated by FZDs, remains unclear. It is also uncertain how the overall structure of FZD compares to other similar proteins and contributes to WNT signalling regulation.

In this thesis, work towards the determination of the first multidomain FZD structure is presented. The expression levels of a representative range of human FZDs are compared. The optimisation of expression and thermostability of candidate receptors for structural studies, through protein engineering, leading to poorly diffracting crystals, is detailed.

The crystal structure of the FZD₃ extracellular domain in complex with a highly specific nanobody (Nb) is described, suggesting a role for the latter in inhibiting signalling. Conversely, a synthetic WNT signalling agonist has been created by fusing this Nb to a WNT-inhibitory protein domain that binds to the WNT coreceptor, LRP6. Functional dissection of FZD₃ facilitated by this agonist confirms that the long C-terminal tail of FZD₃ is dispensable for ‘canonical’ WNT signalling, whereas disruption of the intracellular loops abolishes this activity. Moreover, the mutation of glycine^{5.64f}, perfectly conserved amongst FZD receptors and implicated in a developmental disorder, is shown to have no effect on ‘canonical’ signalling, and instead this residue may form part of a previously overlooked molecular ‘switch’ required for G protein-dependent signalling.

Finally, the crystal structures of the FZD₁₀ and FZD₃ extracellular domains contribute to a more complete understanding of the structural biology of these enigmatic receptors.

Acknowledgements

I would like to thank my supervisor Professor E. Yvonne Jones for the opportunity to undertake this project in the Division of Structural Biology. I also thank EYJ, my co-supervisor Dr Yuguang Zhao, Professors Christian Siebold and Simon Newstead for guidance and encouragement over the last four years.

I would also like to thank: Dr Tao-Hsin Chang for advice on protein construct design; Dr Tom Walter and YZ for advice on protein crystallisation; Dr Karl Harlos and YZ for advice on crystal handling; SN for access to lipidic cubic phase crystallisation facilities, as well as assistance with handling protein crystals; YZ for help with x-ray diffraction data collection; Dr Tomas Malinauskas and YZ for help with x-ray diffraction data processing; Dr Johnathan Elegheert and Weixian Lu for help with cell culture; Dr Christiane Kowatsch for assistance with bio-layer interferometry experiments and support; Margaret Jones and Oliver Hayward for help and advice in the laboratory.

I thank all members of the Division of Structural Biology for their support; Cancer Research UK for funding, and Diamond Light Source for the opportunity to collect x-ray diffraction data. Finally, I would like to thank my friends and family, particularly my wife, Gemma, for encouraging and supporting me throughout my studies.

Table of Contents

<i>Abstract</i>	2
<i>Acknowledgements</i>	3
<i>Table of Contents</i>	4
<i>Table of Figures</i>	9
<i>Chapter 1 Introduction</i>	11
1.1 Background	11
1.2 WNT-FZD signalling overview.....	11
1.3 β -catenin-dependent WNT-FZD signalling.....	11
1.4 β -catenin-independent WNT-FZD signalling.....	13
1.5 Integration of WNT-FZD signalling pathways	17
1.6 FZD structure and function.....	17
1.6.1 FZDs belong to the G protein-coupled receptor (GPCR) superfamily	17
1.6.2 FZDs are atypical GPCRs	18
1.6.3 FZD structural overview	20
1.6.4 FZD CRD structure and function.....	22
1.6.5 FZD LD structure and function.....	23
1.6.6 FZD TMD structure and function	26
1.6.7 FZD cytoplasmic tail structure and function.....	27
1.7 WNT pathway regulation at the cell surface.....	28
1.8 WNT-FZD pathway selectivity	32
1.9 The role of FZD in cancer	36
1.10 Therapeutic targeting of FZD receptors.....	37

1.11 Aims and outcomes of the project	42
<i>Chapter 2 Methods</i>	45
2.1 Molecular biology	45
2.1.1 Synthetic genes and expression vectors	45
2.1.2 Construct design and polymerase chain reaction	46
2.1.3 Restriction enzyme double digest, ligation and transformation.....	46
2.2 Nanobody production	47
2.3 Protein expression.....	48
2.3.1 Small-scale expression	48
2.3.2 Fluorescence size-exclusion chromatography analysis of multidomain FZD constructs	49
2.3.3 Western blot analysis of secreted constructs.....	50
2.3.4 Generation of lentiviral cell lines.....	51
2.3.5 Prokaryotic expression of Nbs	53
2.4 Purification protocols	53
2.4.1 Purification of Nbs from E. coli.....	53
2.4.2 Identification of FZD ₃ CRD-binding Nbs.....	54
2.4.3 Purification of FZD ₃ CRD:Nb ₁₄₄₇₈ complex and FZD ₁₀ CRD	54
2.4.4 Purification of transiently-expressed Rho-1D4 tagged multidomain FZD ₃	56
2.4.5 Purification of stably-expressed Twin-Strep-tagged multidomain FZD ₃	58
2.4.6 Purification of FZD CRD proteins for bio-layer interferometry experiments	59
2.5 Protein crystallisation.....	59
2.5.1 Crystallisation of FZD ₃ :Nb ₁₄₄₇₈ complex	59
2.5.2 Crystallisation of FZD ₁₀ CRD.....	60
2.5.3 Crystallisation of BRIL_FZD ₃	60
2.6 Crystal cryo-protection and x-ray diffraction data collection.....	61
2.6.1 X-ray diffraction data processing.....	61
2.6.2 Structure determination, model building, refinement and validation	61

2.7 Bio-layer interferometry	62
2.8 Signalling Assays.....	62
<i>Chapter 3 Structural studies of multidomain FZD proteins</i>	<i>64</i>
3.1 Introduction	64
3.2 Aims	64
3.3 Construct design and screening.....	66
3.3.1 FZD ₁ multidomain construct design and screening	66
3.3.2 FZD _{2/3/8/10} multidomain construct design and screening	68
3.4 Large-scale multidomain FZD production and optimisation	72
3.4.1 Initial large-scale FZD ₃ expression and crystallisation trials	72
3.5 Further optimisation of multidomain FZD production	74
3.5.1 Lentivirus-transduced cell lines for the expression of FZD ₃	74
3.5.2 Overcoming CRD cleavage and improving yield with use of N-terminal Twin-Strep tag.....	75
3.6 Thermostabilising point mutations	78
3.6.1 Design of point mutations	78
3.6.2 Point mutation expression screens	79
3.6.3 FZD ₃ TMD point mutations increase receptor thermostability.....	82
3.6.4 Theoretical basis for FZD ₃ thermostabilisation by TMD point mutations	84
3.7 Large scale purification and crystallisation trials of optimised multidomain FZD₃ construct	86
3.8 Conclusions and future perspectives.....	86
<i>Chapter 4 Nanobodies against FZD₃</i>	<i>87</i>
4.1 Introduction	87
4.2 Aims	87

4.3 Identification and characterisation of an anti-FZD3 CRD Nb.....	88
4.3.1 Nb14478 binds to the FZD ₃ CRD	88
4.3.2 Affinity of the FZD ₃ :Nb14478 interaction.....	89
4.3.3 Nb14478 binds with high specificity to FZD _{3/6}	90
4.4 Crystal structure of FZD₃ CRD:Nb₁₄₄₇₈ complex.....	90
4.4.1 Overview of the FZD ₃ CRD:Nb ₁₄₄₇₈ crystal structure.....	90
4.4.2 Potential implications of Nb ₁₄₄₇₈ binding on FZD ₃ signalling	98
4.4.3 Basis for the specificity of Nb ₁₄₄₇₈ binding to FZD _{3/6}	101
4.4.4 Nb ₁₄₄₇₈ binding in the context of full-length FZD ₃	108
4.5 Conclusions and future perspectives.....	111
<i>Chapter 5 Functional investigations of Class F GPCRs.....</i>	<i>112</i>
5.1 Introduction	112
5.2 Aims	112
5.3 Activity of FZD₃ in response to WNTs	113
5.3.1 WNT3A does not activate β -catenin-dependent signalling in HEK293T cells expressing FZD ₃	113
5.3.2 WNT5A activates a chimeric ROR2_LRP6 receptor independently of FZD ₃ in HEK293T cells ...	115
5.4 A synthetic agonist of β-catenin signalling permits functional dissection of FZD₃.....	116
5.4.1 A fusion protein comprising Nb ₁₄₄₇₈ fused to the DKK1 C-terminal domain functions as a FZD ₃ -dependent agonist of β -catenin-dependent signalling	116
5.4.2 The FZD ₃ C-terminal tail is dispensable for activation of the ‘canonical’ pathway.....	119
5.4.3 Replacement of the FZD ₃ ICL3 with a BRIL fusion abolishes signalling activity.....	119
5.4.4 The mutation of a conserved glycine does not affect FZD ₃ β -catenin-dependent signalling activity in response to Nb ₁₄₄₇₈ -DKK1C.....	121
5.4.5 OMOD2 may be caused by aberrations in a ‘non-canonical’ FZD G protein signalling activity conserved in SMO.....	121
5.4.6 FZD ₂ G434 ^{S.64FV} (and equivalent mutations) may stabilise a previously unidentified, Class F GPCR-specific ‘switch’ in an ‘off’ position, preventing interaction with G proteins	122

5.5 Further evidence that the FZD₂ G434^{5,64f}V mutation affects G protein signalling.....	125
5.6 Conclusions and future perspectives.....	126
<i>Chapter 6 Comparisons of FZD CRDs</i>	<i>128</i>
6.1 Introduction	128
6.2 WNTs exhibit high levels of conservation in amino acids that interact with FZD CRDs	129
6.3 Comparisons between the FZD₃ CRD and other FZD CRD structures.....	131
6.3.1 FZD ₃ possesses an unusual site 2 architecture	131
6.3.2 The FZD ₃ CRD crystal structure reveals an extended helix not seen in other FZDs.....	133
6.4 The crystal structure of the FZD₁₀ CRD compared to other FZD CRD structures.....	134
6.4.1 Crystal structure of the FZD ₁₀ CRD.....	134
6.4.2 The FZD ₁₀ dimerisation interface is more similar to FZD ₅ , 7, and 8 than FZD ₄	139
6.4.3 Why does Norrin bind to FZD ₄ but not FZD ₁₀ ?	140
6.5 Conclusions and future perspectives.....	142
<i>Conclusion.....</i>	<i>145</i>
<i>Appendices.....</i>	<i>147</i>
Appendix 1 Amino acid sequence alignment of human Class F GPCRs.....	147
Appendix 2 Disorder analysis of human FZD ₃ and FZD ₈ amino acid sequences	150
Appendix 3 Amino acid sequence alignment of anti-FZD ₃ Nbs	151
Appendix 4 Amino acid sequence alignment of human WNT proteins.....	152
Appendix 5 List of constructs in the pHLsec vector.....	154
Appendix 6 List of constructs in the pHR-CMV-TetO ₂ vector	158
Appendix 7 Useful abbreviations and symbols	160
<i>References</i>	<i>164</i>

Table of Figures

Figure 1.1 β -catenin-dependent WNT signalling overview.....	12
Figure 1.2 WNT-PCP signalling.....	15
Figure 1.3 WNT- Ca^{2+} signalling.....	16
Figure 1.4 Structural knowledge of FZD receptors	21
Figure 1.5 Comparison of the LD and ECLs of SMO and FZD.....	25
Figure 1.6 Natural regulators of WNT signalling.....	29
Figure 1.7 Auxiliary coreceptors ‘decode’ WNT-FZD complexes	34
Figure 1.8 Synthetic regulators of WNT signalling.....	38
Figure 2.1 Lentiviral transduction for the creation of stable cell lines	52
Figure 2.2 Purification Nbs from <i>E. coli</i>	55
Figure 2.3 Workflow for the purification of multidomain FZD protein.....	57
Figure 3.1 Multidomain FZD construct design.....	65
Figure 3.2 Multidomain Class F GPCRs and construct boundaries.	67
Figure 3.3 FSEC analysis of multidomain FZD ₁ constructs.....	69
Figure 3.4 FSEC analysis of multidomain FZD _{2/3/8/10} constructs.	71
Figure 3.5 Purification of multidomain FZD for crystallisation experiments.	73
Figure 3.6 N-terminal tag, in combination with lentiviral expression, facilitates purifications of large quantities of highly pure FZD ₃	76
Figure 3.7 FZD ₃ TMD point mutation expression screening	80
Figure 3.8 TMD point mutations increase FZD ₃ thermostability.	83
Figure 3.9 Potential basis for the stabilising effects of FZD ₃ TMD point mutations.	85
Figure 4.1 Nb ₁₄₄₇₈ binds to FZD ₃ CRD with high affinity.....	89
Figure 4.2 Nb ₁₄₄₇₈ specifically binds to the CRDs of the FZD _{3/6} subfamily	91

Figure 4.3 Crystal structure of Nb ₁₄₄₇₈ in complex with FZD ₃ CRD.....	92
Figure 4.4 Electron density of FZD ₃ CRD:Nb ₁₄₄₇₈ structure	94
Figure 4.5 Detail of the interactions between Nb ₁₄₄₇₈ and FZD ₃ CRD.....	96
Figure 4.6 Contributions of the Nb ₁₄₄₇₈ CDRs to the interaction with FZD ₃	97
Figure 4.7 Nb ₁₄₄₇₈ binding to FZD ₃	99
Figure 4.8 Nb ₁₄₄₇₈ appears to cause narrowing of hydrophobic groove on FZD ₃ CRD	100
Figure 4.9 Basis of the specificity of Nb ₁₄₄₇₈ for FZD _{3/6}	103
Figure 4.10 The CDRs of Nb ₁₄₄₇₈ do not interact with FZD ₃ LD.....	110
Figure 5.1 FZD ₃ is not a receptor for WNT3A or WNT5A in HEK293 cells.....	114
Figure 5.2 Nb ₁₄₄₇₈ _DKK1C is a synthetic WNT agonist.....	117
Figure 5.3 Functional dissection of FZD ₃	120
Figure 5.4 SMO G422 ^{5.64f} is located proximal to a conserved tryptophan, W339 ^{3.50f} , that appears to be involved in G protein signalling.	123
Figure 6.1 WNT-FZD interactions	130
Figure 6.2 Differences between the CRDs of FZD _{3/6} and other FZDs	132
Figure 6.3 The implications of the FZD ₃ :FZD ₃ interface observed in the FZD ₃ :Nb ₁₄₄₇₈ crystal packing on WNT binding.....	135
Figure 6.4 Different FZD CRD dimer configurations	136
Figure 6.5 Electron density map of FZD ₁₀ CRD structure	138
Figure 6.6 The FZD ₁₀ CRD structure strengthens understanding of the basis of Norrin:FZD ₄ specificity.....	141
Table 1 The 5 subfamilies of Class F GPCRs and their known natural and synthetic interaction partners.....	39
Table 2 Crystallographic data collection and refinement statistics for FZD ₃ CRD:Nb ₁₄₄₇₈	93
Table 3 Crystallographic data collection and refinement statistics for FZD ₁₀ CRD	137

Chapter 1 Introduction

1.1 Background

The first knowledge of the *WNT* family of genes emerged in 1982, when it was observed that mouse mammary tumour virus (MMTV)-driven oncogenesis frequently involved provirus integration into a specific region of the mouse genome, leading to the expression of a suspected proto-oncogene, which was accordingly named *int1* (Nusse and Varmus, 1982). This gene was later found to be related to a *Drosophila* gene, *wingless* (Rijsewijk *et al.*, 1987), mutations in which lead to abnormalities in wing development, indicating a potential role as a developmental growth factor (Sharma and Chopra, 1976). *int1* and related genes were eventually renamed ‘Wnt,’ a mnemonic based on ‘Wingless-related integration site’ (Nusse *et al.*, 1991). *WNT* genes are now known to encode a family of secreted glycoproteins that are lipid-modified on a conserved serine residue (Willert *et al.*, 2003; Janda *et al.*, 2012) and function by engaging the FZD seven-pass transmembrane receptor at the cell surface (Bhanot *et al.*, 1996) to elicit a range of downstream responses including modulation of gene expression and establishment of cell polarity. WNT signalling is commonly dysregulated in diseases such as cancer.

1.2 WNT-FZD signalling overview

1.3 β -catenin-dependent WNT-FZD signalling

The best-understood branch of the WNT signalling network is the β -catenin-dependent or ‘canonical’ pathway (**Figure 1.1**), activation of which leads to modulation of expression of target genes. In the absence of WNT, the transcriptional regulator β -catenin is sequentially phosphorylated, ubiquitinated and proteasomally degraded by a so-called ‘destruction complex’ (Li *et al.*, 2012). This complex comprises the scaffold proteins Dishevelled (DVL)

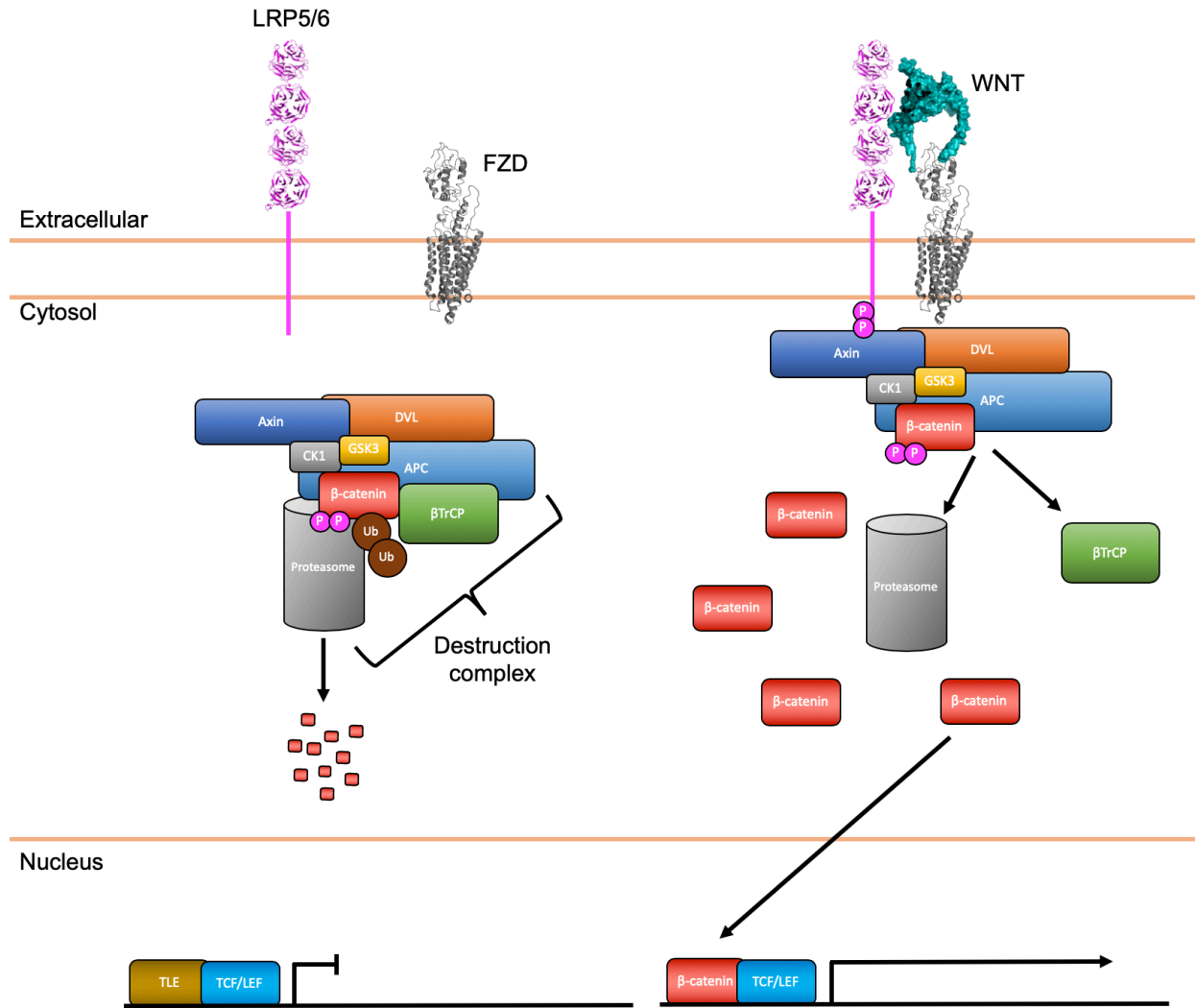


Figure 1.1 β -catenin-dependent WNT signalling overview

In the absence of a WNT ligand (left), β -catenin is phosphorylated and ubiquitinated by the destruction complex, which comprises Axin, DVL, CK1, GSK1, APC, β TrCP and proteasomal subunits, leading to its destruction. In the nucleus, TCF/LEF transcription factors repress transcription in conjunction with TLE. Upon WNT stimulation (right), the destruction complex is recruited to the plasma membrane, resulting in the loss of ubiquitination activity. Newly synthesised β -catenin subsequently accumulates and translocates to the nucleus to modulate expression of target genes. WNT (teal) from PDB ID 6AHY; FZD (grey) homology model using 5L7D as template; LRP5/6 (magenta), PDB 3S2K.

and Axin; the kinases glycogen synthase kinase 3 (GSK3) and casein kinase 1 (CK1); Adenomatous polyposis coli (APC); the β -transducin repeat-containing E3 ubiquitin protein ligase (β TrCP); and at least some components of proteasomes (Li *et al.*, 2012). The bringing into close proximity of a FZD receptor and a low-density lipoprotein receptor-related protein

5/6 (LRP5/6) co-receptor upon WNT binding (Cong, Schweizer and Varmus, 2004) leads to the sequential phosphorylation of the LRP5/6 intracellular domain by GSK3 and CK1, which facilitates axin recruitment to the cell membrane (Zeng *et al.*, 2005). Contrary to earlier models (reviewed by Clevers and Nusse, 2012), the destruction complex has been shown to remain mostly intact at this stage, with only β TrCP being absent (Li *et al.*, 2012). It therefore seems that the inactivation of the destruction complex is caused by the loss of β -catenin ubiquitination by β TrCP (Li *et al.*, 2012), rather than a more extensive disassembly of the complex. Consequently, the intact destruction complexes remain saturated with β -catenin, allowing newly-synthesised β -catenin to accumulate (Li *et al.*, 2012) and translocate to the nucleus, where it can modulate the expression of target genes in concert with members of the T-cell factor/lymphoid enhancer factor (TCF/LEF) family of transcription factors (Behrens *et al.*, 1996; Molenaar *et al.*, 1996). In the absence of signalling, TCF/LEF function as transcriptional repressors in complex with the Transducin-like enhancer of split 2 (TLE) protein (Daniels and Weis, 2005). Examples of target genes that are upregulated in response to WNT/FZD- β -catenin signalling include oncogenes such as the vascular endothelial growth factor (VEGF; Zhang, Gaspard and Chung, 2001) and the transcriptional regulator, c-MYC (He *et al.*, 2008).

1.4 β -catenin-independent WNT-FZD signalling

WNT-FZD binding can also activate pathways that exert their effects independently of β -catenin/TCF/LEF. Such pathways are traditionally referred to as ‘non-canonical,’ and are generally less well-characterised than the β -catenin-dependent pathway. Unlike the β -catenin-dependent pathway, non-canonical WNT pathways do not involve the LRP5/6 co-receptors, instead signalling through co-receptors including the retinoic acid-related orphan tyrosine kinase receptor (ROR1/2) and the related to tyrosine kinase receptor (RYK, Niehrs, 2012).

An example of such a pathway is the planar cell polarity (PCP) pathway, in which FZD activation leads to the establishment of cell polarity within a two-dimensional plane,

rearrangement of the cytoskeleton and/or transcription modulation (Niehrs, 2012). Another example of a non-canonical WNT pathway is the WNT/calcium (WNT/Ca²⁺) pathway, in which FZD activation triggers downstream effects via the release of calcium from intracellular compartments (Niehrs, 2012).

In the WNT/PCP pathway (**Figure 1.2**), WNT concentration gradients appear to direct the localisation of distinct protein complexes to opposite sides of the cell (reviewed by Yang and Mlodzik, 2015). More specifically, a FZD-containing complex that also comprises DVL, the cadherin EGF LAG seven-pass G-type receptor (CELSR) and Diversin, accumulates at the edge of the cell facing towards the source of the WNT ligand (Yang and Mlodzik, 2015). Conversely, a second complex that also contains CELSR; as well as Van Gogh-like (VANGL), a four-pass membrane protein; and Inversin, is directed to the opposite side of the cell, resulting in the establishment of an asymmetric axis (Yang and Mlodzik, 2015). Inversin and Diversin participate in positive and negative feedback mechanisms to enhance this arrangement of protein complexes. On the one hand, Inversin can bind to DVL, preventing it from being recruited to FZD and reducing its stability, whereas on the other hand Diversin competes with Inversin for binding to DVL and thereby stabilises the FZD-DVL interaction (Jenny *et al.*, 2005). While the FZD-DVL-CELSR-Diversin and CELSR-VANGL-Inversin complexes antagonise one another intracellularly, in contrast, apposing complexes on neighbouring cells stabilise one another through homophilic CELSR-CELSR and heterophilic VANGL-FZD interactions (Usui *et al.*, 1999; Wu and Mlodzik, 2008). Establishment of polarity of these core PCP complexes leads to changes in cytoskeletal arrangement and transcription via signalling proteins such as RhoA (Strutt, Weber and Mlodzik, 1997).

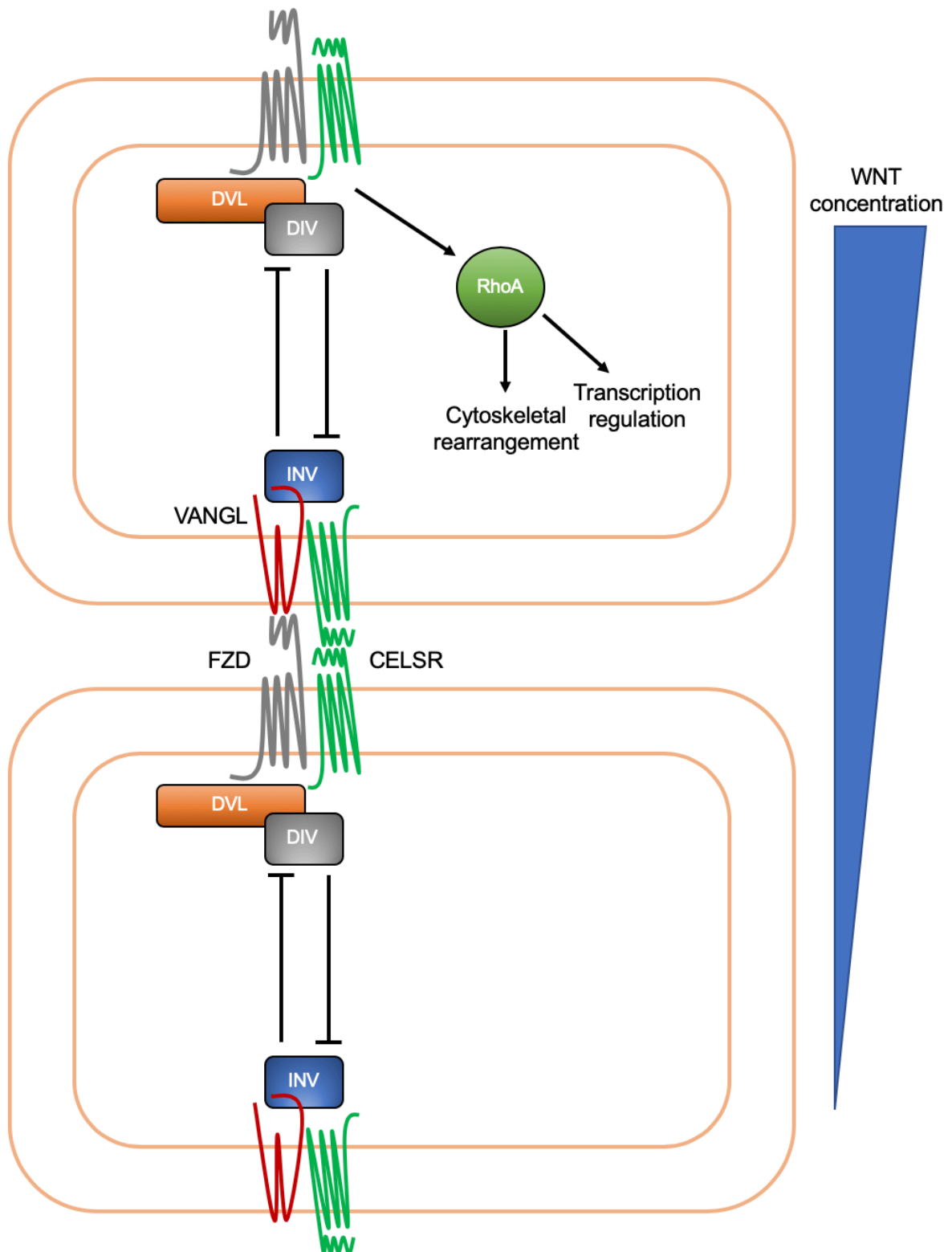


Figure 1.2 WNT-PCP signalling.

WNT gradients lead to the establishment of an asymmetrical distribution of distinct protein complexes (adapted from Yang and Mlodzik, 2015).

In the WNT/ Ca^{2+} pathway (**Figure 1.3**), WNT binding to a cognate FZD results in a release of Ca^{2+} from the endoplasmic reticulum (ER, Komiya and Habas, 2008). Intracellular Ca^{2+} -sensitive proteins, including protein kinase C (PKC) and calcium/calmodulin-dependent kinase II (CAMKII) are subsequently activated, resulting in a variety of downstream effects including inhibition of canonical signalling and tissue separation during development (Komiya and Habas, 2008). WNT/ Ca^{2+} signalling appears to rely on activation of G proteins by FZD (Slusarski, Corces and Moon, 1997).

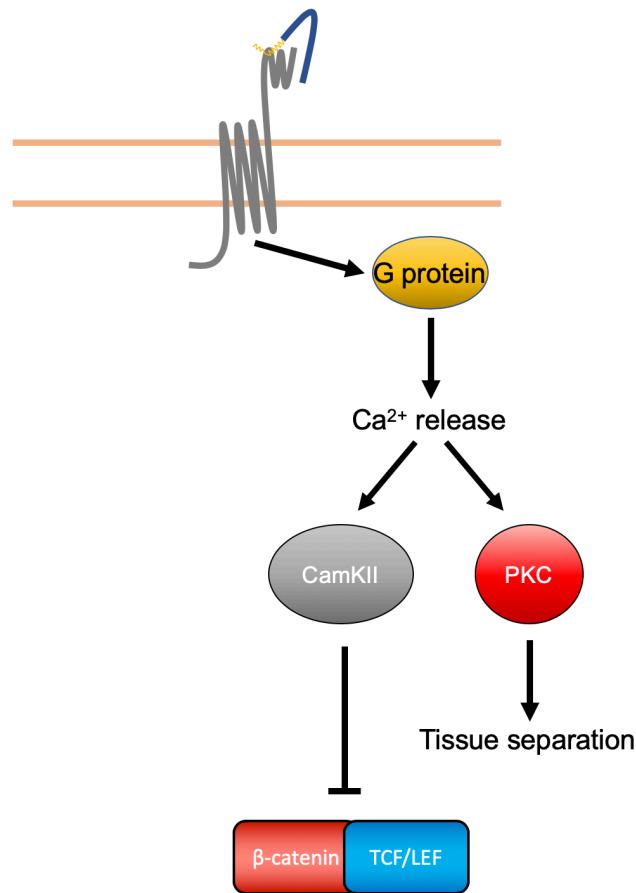


Figure 1.3 WNT- Ca^{2+} signalling

WNT binding to a FZD receptor leads to the release of Ca^{2+} from the ER, resulting in the activation of Ca^{2+} -dependent kinases (adapted from Komiya and Habas, 2008).

1.5 Integration of WNT-FZD signalling pathways

While classifying different WNT pathways, for example by their ability to stabilise β -catenin, helps to simplify understanding of this complicated signalling platform, it is becoming increasingly apparent that the different branches are not entirely separate. For example, non-canonical FZD signalling via Rho can antagonise canonical signalling by upregulating the expression of negative modulators of β -catenin-dependent signalling, such as Dickkopf (DKK Park *et al.*, 2015). Furthermore, FZD₆ has been shown to antagonise β -catenin stabilisation in mammalian cells by activating the mitogen-activated protein kinase kinase kinase 7 (MAP3K7)/Nemo-like kinase (NLK) pathway (Golan *et al.*, 2004), which leads to the phosphorylation of TCF/LEF transcription factors to prevent their β -catenin-dependent transcriptional effects (Ishitani *et al.*, 1999). ‘Canonical’ and ‘non-canonical’ WNTs have also been suggested to compete for FZD binding and thereby reciprocally inhibit one another (Grumolato *et al.*, 2010). Therefore, WNT signalling is perhaps better considered as a complex group of interconnected pathways with a broad spectrum of potential outcomes.

1.6 FZD structure and function

1.6.1 FZDs belong to the G protein-coupled receptor (GPCR) superfamily

GPCRs comprise the largest superfamily of membrane proteins and are responsible for interpreting the majority of cellular responses to hormones, neurotransmitters and environmental stimuli (Rosenbaum, Rasmussen and Kobilka, 2009). Vertebrate GPCRs are broadly divided into five classes based on sequence and structural similarities; Class A (rhodopsin family), Class B (secretin family), Class C (glutamate family), the adhesion family, and Class F (FZD family) (Alexander *et al.*, 2017). GPCRs are primarily identified by their seven-pass transmembrane domain (7TM), comprising seven membrane-spanning α -helices

(TM1-7) separated by alternating intracellular and extracellular loops (ICL/ECL; Rosenbaum, Rasmussen and Kobilka, 2009). Class F GPCRs, which include FZD and the closely-related Smoothed (SMO), which functions in the Hedgehog signalling pathway (Kong, Siebold and Rohatgi, 2019), contain an eighth helix (H8), also found in Class A GPCRs, which packs against the cell membrane intracellularly and mediates interactions with downstream signalling components (Wang *et al.*, 2013). GPCRs are of immense importance for drug design, and indeed the majority of drugs used clinically target this group of proteins (Alexander *et al.*, 2017).

A generic numbering system for the amino acids in GPCR TM helices, based on the position of highly conserved residues, simplifies comparisons between different receptors (Ballesteros and Weinstein, 1995). In brief, the most highly conserved residue in each helix is numbered X.50 (where X is the number of the helix, e.g. 1.50 in TM1) and other residues are numbered based on their position relative to that amino acid (Ballesteros and Weinstein, 1995). For example, an amino acid located five positions downstream of 1.50 will be denoted 1.55. For each GPCR family, the residues denoted as X.50 vary based on different patterns of conservation and are followed by a letter to indicate which family numbering system is being used (e.g. X.50f in Class F GPCRs, Isberg *et al.*, 2015). The positions of X.50 residues in Class F GPCRs (Wang *et al.*, 2014) are given in **Appendix 1**.

1.6.2 FZDs are atypical GPCRs

At first glance, the FZD 7TM domain is highly reminiscent of traditional GPCRs. However, closer analysis reveals that FZDs lack several features commonly found in other GPCRs, such as the Asp-Arg-Tyr motif or Asn-Pro-X-X-Tyr motif (where X can be any amino acid) that regulate receptor activation (Angers and Moon, 2009). Despite this, there is a wealth of evidence that FZDs indeed signal via G proteins (reviewed by Dijksterhuis, Petersen and Schulte, 2014).

For example, FZD₁-induced β -catenin stabilisation can be blocked by inactivating or depleting G proteins, indicating that canonical WNT signalling may involve G protein coupling (Liu *et al.*, 2001). Furthermore, overexpression of wild type G proteins, or expression of constitutively active G proteins (even in the absence of FZD), can activate both β -catenin-dependent and -independent signalling in *Drosophila* (Katanaev *et al.*, 2005). WNT3A has also been shown to stimulate guanine-nucleotide exchange on G proteins in a FZD-dependent manner (Koval and Katanaev, 2011). WNT-Ca²⁺ signalling appears to be dependent on activation of trimeric G proteins by FZD, as intracellular calcium release is abolished by inhibitors of G proteins such as pertussis toxin (Slusarski, Corces and Moon, 1997). Furthermore, WNT5A-induced phosphorylation of the protein phosphatase 1 regulatory subunit 1B is abolished upon knockdown of G α in a β -catenin-independent pathway (Hansen *et al.*, 2009). However, it was recently shown that WNT3A-induced β -catenin signalling can function even in cells lacking expression of any G proteins, which suggests that canonical WNT signalling is not strictly reliant on the GPCR activity of FZD, although G proteins may play a regulatory role in this pathway (Bowin, Inoue and Schulte, 2019). FZDs are therefore considered atypical GPCRs and are categorised in a separate class, Class F, alongside the closely-related SMO GPCR from the Hedgehog signalling pathway (Foord *et al.*, 2005).

While it remains unclear exactly how FZD activates G proteins, certain structural features that may be involved have been identified. For example, a molecular lock has been proposed to exist in class F GPCRs, involving a conserved interaction between arginine 6.32f (lysine in FZD_{4/9}) and tryptophan 7.55f, which may regulate G protein signalling (Wright *et al.*, 2019). This interaction can be observed in the FZD₄ TMD structure, as well as the inactive SMO structure, but appears to be broken in the putative active SMO structure, suggesting an inhibitory role (Kong, Siebold and Rohatgi, 2019). The lock is frequently abolished by mutations seen in cancer (for example R416Q in FZD₆), suggesting a model whereby the

arginine/tryptophan interaction could regulate downstream signalling by maintaining the receptor in a conformation that does not bind G proteins in the absence of a stimulus, whereas disruption of the lock may stabilise a conformation that constitutively activates G proteins (Wright *et al.*, 2019). Interestingly, mutation of the lock in FZD₅ was found to abolish DVL recruitment and β -catenin stabilisation, suggesting that the lock is required to be intact in order for FZD to activate DVL (Wright *et al.*, 2019).

1.6.3 FZD structural overview

The ten human FZD receptors comprise an N-terminal, extracellular cysteine-rich domain (CRD) containing approximately 120 amino acids and a GPCR 7TM or transmembrane domain (TMD; **Figure 1.4, a**; Bhanot *et al.*, 1996). The TMD is followed by a C-terminal cytoplasmic tail that varies greatly in length (MacDonald and He, 2012). The CRD and TMD are joined by a linker domain (LD), which is typically around 60 to 80 amino acids in length, however FZD₄ has a shorter linker of around 50 residues, while FZD₈ has a dramatically extended linker of around 120 residues. The FZD LD is generally much longer than the 40 amino acid linker found in SMO. Overall sequence identity between human FZDs ranges from around 34% (between FZD₆ and FZD₉) to 80% (between FZD₁ and FZD₂) but is typically around 45%. By contrast, the LD has much lower sequence conservation, ranging from around 10% (between FZD₃ and FZD₄)

to 57% (between FZD₁ and FZD₂). Overall sequence identity between FZDs and SMO is around 25%. The intracellular domain of FZD also varies greatly in length, between 25 amino acids in FZD₁ and 2, and 212 in FZD₆, although the SMO intracellular domain is longer still at 240 amino acids. Annotated sequence alignments of all human Class F GPCRs can be found in **Appendix 1**.

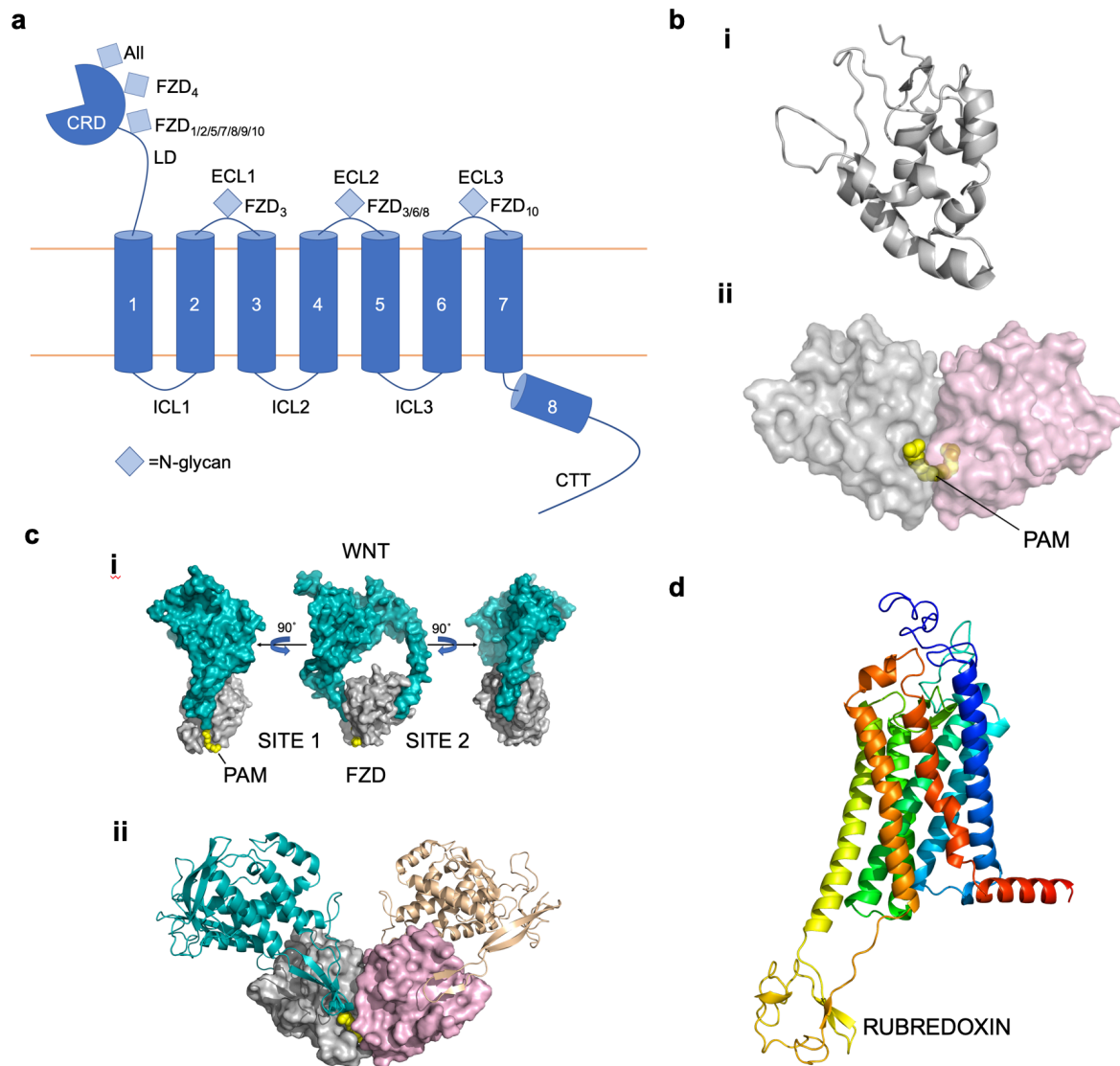


Figure 1.4 Structural knowledge of FZD receptors

a, Schematic diagram showing the major structural domains of FZD receptors. Known and putative N-linked glycosylation sites are indicated by light blue squares. The FZDs that possess each site are indicated. CTT=C-terminal tail. **b**, **i**, Cartoon of mouse FZD₈ CRD monomer (PDB 1IJY). **ii**, Surface diagram of dimeric FZD₅ CRD with palmitoleic acid bridging the dimerisation interface (PDB 5URY). PAM=palmitoleic acid. **c**, **i**, Crystal structure of human WNT3 in complex with mouse FZD₈ CRD with site 1 and site 2 indicated (PDB 6AHY). **ii**, A 2:2 WNT3:FZD₈ complex observed in the crystal lattice (PDB 6AHY). **d**, Crystal structure of FZD₄ TMD fused to rubredoxin (PDB 6BD4).

1.6.4 FZD CRD structure and function

The FZD CRD (**Figure 1.4, b, i**) is responsible for binding WNT ligands (Bhanot *et al.*, 1996) and, in the case of FZD₄, Norrin (Xu *et al.*, 2004). The CRD is the most structurally characterised domain of the FZD receptor, and the structures of five FZD CRDs (FZDs 2, 4, 5, 7 and 8) have been published to date (Dann *et al.*, 2001; Chang *et al.*, 2015; Nile *et al.*, 2017; Chen *et al.*, 2018), including two structures of mouse FZD₈ in complex with WNTs (Janda *et al.*, 2012; Hirai *et al.*, 2019), and the structure of FZD₄ in complex with Norrin, a FZD₄-specific ligand (Chang *et al.*, 2015; Shen *et al.*, 2015). A defining characteristic of the primarily α -helical CRD is its conserved arrangement of ten cysteines, which form five disulphide bonds (**Appendix 1**; Dann *et al.*, 2001). FZD CRDs also contain a conserved hydrophobic groove, which is able to accommodate the palmitoleic acid moiety appended to WNT ligands; this so-called ‘site 1’ interaction (**Figure 1.4, c, i**) forms a major part of the ligand-receptor interaction (Janda *et al.*, 2012). A second site, ‘site 2,’ on the opposite side of the CRD (**Figure 1.4, c, ii**), mediates protein-protein interactions with the WNT ‘finger,’ and, although this region is highly conserved among WNTs (see **Appendix 4**), it has been suggested that the substitution of certain residues may contribute to WNT-binding specificity of different FZDs (Janda *et al.*, 2012).

Crystal structures have also revealed a propensity for FZD CRDs to dimerise through an interface comprising mainly the first α -helix of the domain (Nile *et al.*, 2017). Interestingly, this dimerisation interface allows the grooves of two CRDs to form a single, extended U-shaped cavity, and an acyl chain can thereby bridge this interface (**Figure 1.4, b, ii** Nile *et al.*, 2017). The recent structure of FZD₈ CRD in complex with WNT3 confirmed the ability of WNTs to dimerise FZDs via this mechanism (Hirai *et al.*, 2019). Intriguingly, a 2:2 FZD-WNT complex was observed (**Figure 1.4, c, ii**), in which the palmitoleic acid moiety of both WNTs

could simultaneously occupy the U-shaped cavity (Hirai *et al.*, 2019), indicating flexibility in this region of the protein.

Interestingly, experiments in *Drosophila* showed that expression of FZD proteins lacking the CRD could rescue the phenotype of flies lacking FZDs, suggesting that the CRD may not be the only ligand binding site on the FZD receptor (Chen *et al.*, 2004), although it is required for efficient signalling in cell-based assays (Povelones and Nusse, 2005).

1.6.5 FZD LD structure and function

The structural and functional significance of the FZD LD remains unclear. However, clues can be gained from the structure and function of this domain in the closely related SMO receptor. In SMO, this region of the receptor is involved in the interaction with ligands such as cholesterol, and is needed to stabilise the inactive SMO conformation (Byrne *et al.*, 2016). SMO TMD structures have shown that the LD forms an extension of the TMD binding pocket and is involved in the binding of small molecule antagonists including SANT1 and SAG1.5 (Wang *et al.*, 2014). The structure of the multidomain SMO receptor in a cholesterol-bound, ‘resting’ state, revealed that the SMO CRD rests on top of the LD, which forms a ‘wedge’ between the CRD and TMD (Byrne *et al.*, 2016). It was also shown that residue V210 of the SMO LD is involved in orienting the tail of the cholesterol molecule (bound in the CRD hydrophobic groove) in an elongated conformation (Byrne *et al.*, 2016). The authors of this publication investigated the functional significance of the interactions between the CRD and LD by separately introducing N-linked glycosylation sites at two positions on the CRD that interact with the LD. The addition of bulky sugars at these positions would therefore disrupt these interactions. Both of these modifications conferred a level of constitutive activity on SMO, mimicking the effect of deleting the CRD entirely (Byrne *et al.*, 2016). A similar result was observed when two cysteines in the LD, which form a disulphide bond that stabilises the LD conformation, were mutated. This suggested that the LD was responsible for maintaining

unstimulated SMO in an inactive conformation. However, the same mutations also abolished SMO activation upon stimulation with the CRD-binding agonist, 20(S)-hydroxycholesterol, indicating that the LD is also required for full SMO activation, perhaps by communicating ligand binding to the TMD via conformational changes (Byrne *et al.*, 2016). These results demonstrate the functional significance for the SMO LD in regulating SMO activation, both in terms of preventing signalling in the absence of agonists and relaying conformational changes required for receptor activation.

The FZD LD is less well-understood. To date, only a single FZD TMD crystal structure has been published (Yang *et al.*, 2018). The construct used to produce this structure of the FZD₄ TMD contained an incomplete LD, which, although sharing similarities to SMO, possessed some subtle differences (Yang *et al.*, 2018; **Figure 1.5**). Specifically, while in SMO the LD is positioned against the third extracellular loop (ECL3) and stacks on top of the ECL1 (**Figure 1.5, b**), in FZD₄ the LD is sandwiched between ECL1 and ECL3 (Yang *et al.*, 2018; **Figure 1.5, c**). It has been demonstrated that while FZD₆ constructs lacking the CRD are well-expressed and localised to the cell surface, this is not the case when the LD is also deleted, arguing that the LD is essential for correct expression of the receptor (Valnohova *et al.*, 2018). Specifically, the residue C161, which is predicted to form a disulphide with another LD cysteine, based on the FZD TMD structure, is essential for correct expression at the plasma membrane (Valnohova *et al.*, 2018). Moreover, mutation of either of these cysteines, or a third in the LD that forms a disulphide with a cysteine in the ICL1 of SMO and FZD₄, also abolished correct localisation of FZD₆ to the cell surface (Valnohova *et al.*, 2018). These results highlight the importance of the LD secondary structure for FZD expression.

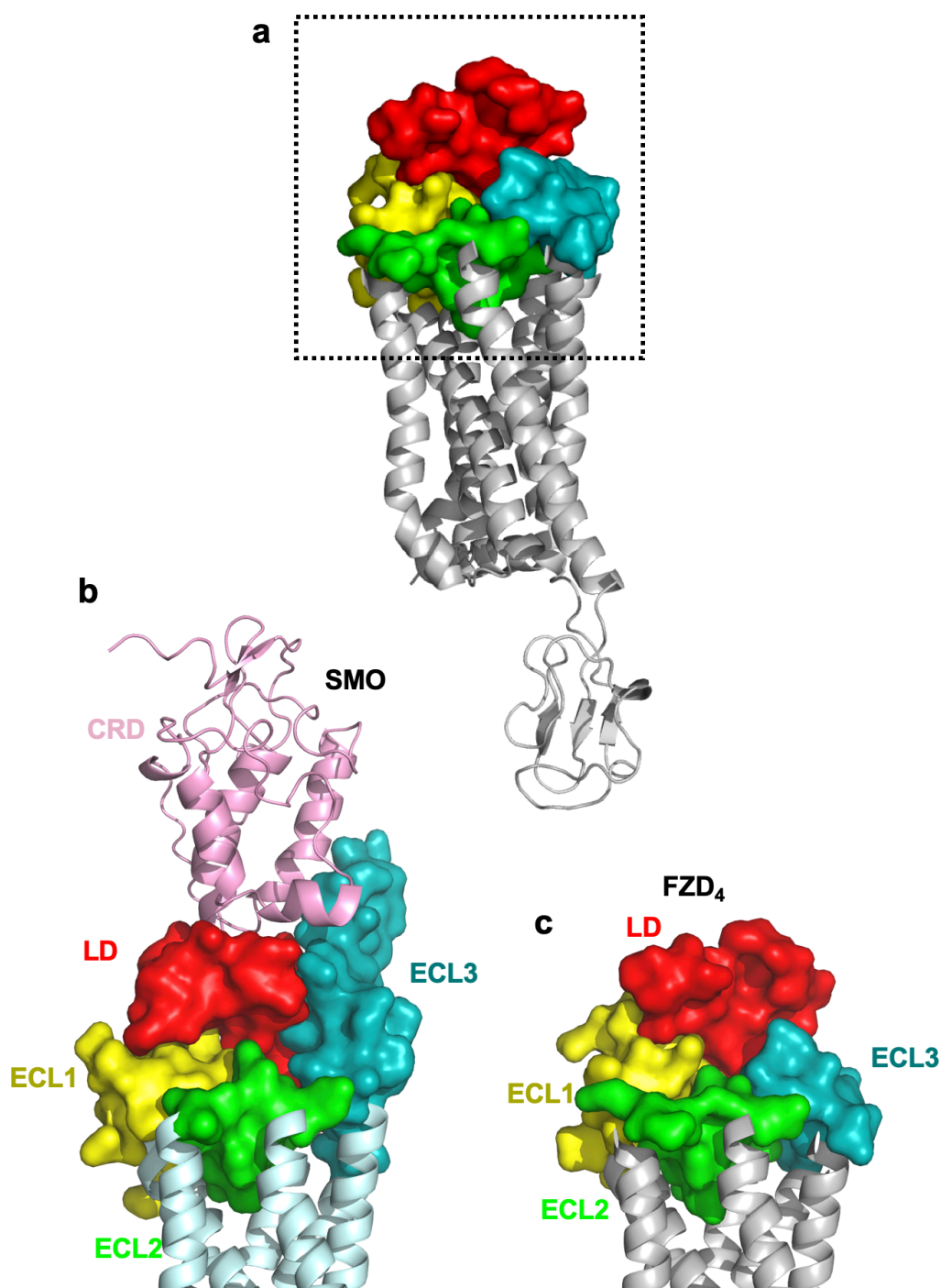


Figure 1.5 Comparison of the LD and ECLs of SMO and FZD

a, Crystal structure of the FZD₄ TMD (PDB 6BD4) with LD and ECLs indicated by dashed square. **b** and **c**, Crystal structure of SMO (PDB 5L7D, **a**) and FZD₄ (**b**) focusing on the extracellular domains. SMO has a substantially longer ECL3 than FZD₄, formed partially from a long extension of TM6, and this stabilises the conformation of the CRD. The SMO LD also rests against the larger ECL3 and stacks on top of ECL1, whereas in FZD₄, the LD is sandwiched between these two domains. Both receptors possess an ECL2 that forms a plug within the TM bundle (adapted from Yang *et al.*, 2018).

Evidence is growing for a role of the FZD LD in ligand binding, as observed for SMO. First of all, it has been shown that expression of FZD proteins lacking the CRD is able to rescue the phenotypes associated with a loss of FZD expression in *Drosophila* (Chen *et al.*, 2004). This result suggested that low-affinity ligand binding sites may exist in the FZD LD (or ECLs), which could be involved in ligand recognition and signal transduction. Furthermore, it was recently demonstrated that the affinity of the interaction between Norrin and the FZD₄ CRD was increased by almost an order of magnitude in the presence of the LD, however this effect was not observed when the LD was replaced with the FZD₅ LD, suggesting that the FZD₄ LD specifically is involved in Norrin recognition (Bang *et al.*, 2018). The deletion of the FZD₄ LD, or replacement with that of FZD₅, also led to a reduction in Norrin-induced β -catenin stabilisation (Bang *et al.*, 2018). It is therefore feasible that the poorly conserved FZD LD may be an important factor in WNT-FZD binding and specificity, which could explain how highly similar WNTs are capable of differentially stimulating specific FZDs.

1.6.6 FZD TMD structure and function

Structural knowledge of Class F GPCR TMDs was limited to SMO until 2018, when the FZD₄ TMD crystal structure was published (**Figure 1.4, d**). While the overall structure of FZD₄ is similar to SMO, several unique features of FZD TMDs were revealed. The most striking difference is that while FZD contains a conserved pocket in the TMD, a common feature of GPCRs (Wacker, Stevens and Roth, 2017), it is narrow and highly hydrophilic (Yang *et al.*, 2018), compared to the more hydrophobic cavity of SMO, which is more typical of GPCRs (Byrne *et al.*, 2016). This potentially explains why no endogenous or synthetic ligands targeting the FZD TMD pocket have been identified, and suggests that unlike typical GPCRs, FZD activation is not influenced by the binding of modulating agents in the TMD pocket. This is further supported by the finding that mutations of residues lining the pocket have little impact on FZD₄ activation (Yang *et al.*, 2018). Furthermore, the TM6 of FZD₄ does not appear to be

capable of the outward movement that is implicated in receptor activation in SMO and many Class A GPCRs (Huang *et al.*, 2018), and instead seems to remain closely packed inside the TM bundle (Yang *et al.*, 2018). In contrast, molecular dynamics simulations indicate that an outward swing of TM4 may contribute to FZD₄ activation by exposing a putative DVL-binding pocket (Strakova *et al.*, 2017; Yang *et al.*, 2018). The FZD₄ TMD also exhibits two unusual kinks in TM7, which can adopt bent and linear conformations, although the implications of this for receptor activation are unclear (Yang *et al.*, 2018).

The FZD₄ TMD structure showed that the extracellular regions of FZD are broadly similar to those of SMO, with the LD packing tightly with the extracellular loops and the ECL2 β -hairpin of FZD₄ forming a plug which partially occupies the TMD pocket (Yang *et al.*, 2018). However, the extracellular extension of TM6 of the FZD₄ TMD is markedly shorter than the same helix in SMO and the ECL3 in general is also shorter (Yang *et al.*, 2018; **Figure 1.5**). While FZD₄ has the shortest ECL of the FZDs, all FZDs are predicted to possess a shorter ECL3 than SMO, except for FZD₁, in which it is of equal length. This shorter ECL3 suggests that FZD₄ possesses a smaller interaction interface between the CRD and ECL3 compared to SMO, which, combined with the different LD arrangement described above, would lead to greater positional flexibility in the CRD in contrast to SMO (Yang *et al.*, 2018). The absence of small molecule ligands to stabilise the TMD and the predicted positional flexibility of the CRD represent two major hurdles in the path towards the determination of structures of FZD constructs encoding the CRD, LD and TMD.

1.6.7 FZD cytoplasmic tail structure and function

Knowledge of the FZD intracellular tail is relatively limited. Evidence currently available suggests that it functions mainly to mediate interactions with other proteins and may be subject to regulatory post-translational modifications. While most FZDs have a short intracellular domain between 25 and 82 amino acids in length, FZD₃ and FZD₆ have longer tails that are

168 and 212 amino acids long, respectively. It is currently unclear what function, if any, the extended intracellular domain of FZD_{3/6} may perform. A conserved KTXXXW motif in the FZD cytoplasmic tail has been shown to mediate an interaction with the PDZ domain of DVL (Wong *et al.*, 2003). All FZDs other than FZD_{3/6} harbour an S/TXV/L PDZ domain-binding site that has been proposed to interact with the Na⁺/H⁺ exchange regulatory factor 1 (NHERF1) to mediate tethering of FZDs to the cytoskeleton (Wheeler *et al.*, 2011). Phosphorylation of the FZD intracellular tail can lead to downregulation of signalling (Shafer *et al.*, 2011).

1.7 WNT pathway regulation at the cell surface

Given the roles of WNT signalling in cell proliferation and related processes, it is perhaps unsurprising that these pathways are subject to strict regulation. Regulation of the WNT pathway at the cell surface is largely achieved by controlling the availability and ability of signalosome components to form active complexes with FZD receptors at the cell surface (**Figure 1.6**).

For example, members of the secreted FZD-related protein (SFRP) family, which comprise an N-terminal FZD-like CRD and a C-terminal Netrin-like (NTR) domain (Bu *et al.*, 2017), are able to bind to and sequester WNT ligands (**Figure 1.6, b**), thereby preventing them from forming active complexes with FZD receptors (Cruciat and Niehrs, 2013). However, while the prevailing model is that the interaction with WNT is mediated by the SFRP CRD (Lin *et al.*, 1997), there is also evidence, surprisingly, that the NTR domain may be capable of binding WNTs and exerting WNT inhibitory effects in the absence of the CRD (Üren *et al.*, 2000). Furthermore, SFRPs have been shown to exert contradictory effects in cancers. For instance, SFRP1 expression is commonly lost through promoter methylation in colorectal cancer (Caldwell *et al.*, 2004) supporting the view of SFRPs as inhibitors of WNT signalling. Conversely, in metastatic renal cell carcinoma, SFRP1 is commonly overexpressed (Saini *et al.*, 2009), which challenges the classical model of SFRPs as simple WNT inhibitors. Possible

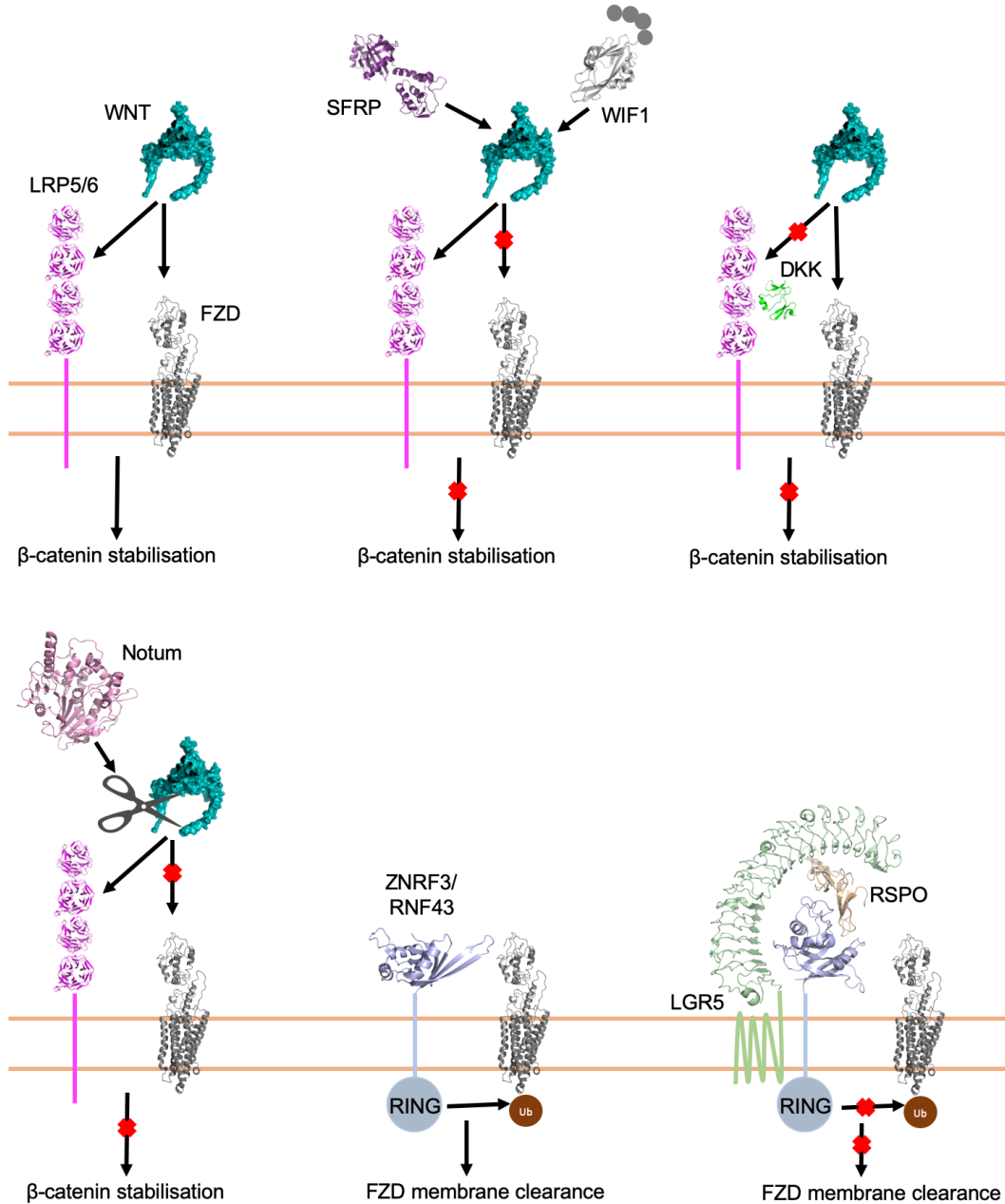


Figure 1.6 Natural regulators of WNT signalling

a, In the canonical pathway, WNT binds to FZD and LRP5/6, resulting in the downstream stabilisation of β -catenin. **b**, SFRP or WIF1 can bind to and sequester WNTs, so they cannot activate FZD, thereby inhibiting canonical signalling. **c**, DKK binds to the LRP5/6 extracellular domain preventing its interaction with WNT. **d**, Notum enzymatically removes the palmitoleic acid moiety from WNTs, thereby rendering them unable to bind to FZD. **e**, **i**, ZNRF3/RNF43 ubiquitinate FZD, leading to its clearance from the membrane. **ii**, RSPO forms a complex with LGR5 and ZNRF3/RNF43, leading to its clearance from membrane. FZD is stabilised. WNT (teal) from PDB ID 6AHY; FZD (grey) homology model using 5L7D as template; LRP5/6 (magenta) and DKK (green) based on 3S2K; SFRP (purple) from structure of *Xenopus* Sizzled, 5XGP; Notum (pink) from 4UZ1; ZNRF3 from 4C84; LGR5/RSPO/ZNRF3 from 4KNG.

explanations for these apparently tumorigenic actions of SFRPs include the facilitation of long-range diffusion of WNT ligands (Mii and Taira, 2009) or via the inhibition of non-canonical WNT signalling, which itself can antagonise WNT/ β -catenin signalling. Therefore, SFRPs are perhaps better understood as context-dependent modulators of WNT signalling, rather than strict WNT antagonists.

Another extracellular WNT-binding regulatory molecule is the WNT inhibitory factor 1 (WIF1), which comprises a WIF domain and five epidermal growth factor (EGF)-like domains (Malinauskas *et al.*, 2011). Similarly to SFRPs, the WIF1 WIF domain is capable of binding to and sequestering WNTs (**Figure 1.6, b**) to regulate FZD receptor activation (Hsieh *et al.*, 1999). Furthermore, a WIF domain has been identified in the ‘non-canonical’ WNT receptor, RYK, which also mediates binding to WNT ligands to transduce WNT/PCP signals (Macheda *et al.*, 2012). Thus, sequestration of WNT ligands by secreted ‘decoy’ receptors is an important layer of regulation in the WNT pathway.

A common feature of these two classes of extracellular regulators of WNT signalling is interaction with heparan sulphate proteoglycans (HSPGs), a class of extracellular proteins extensively modified with heparan sulphate glycosaminoglycan (GAG) chains (Lin, 2004). SFRPs have been shown to interact with HSPGs via the C-terminal NTR domain (Üren *et al.*, 2000), whereas WIF1 does so via its C-terminal EGF domains (Malinauskas *et al.*, 2011). Therefore, while these classes of proteins are structurally unrelated, they share a common modularity and mechanistic action, whereby the N-terminal portion of the protein binds WNTs while the C-terminal domains bind to HSPGs. The interaction with HSPGs may function to colocalise these modulators of WNT signalling with WNTs, which have also been shown to interact with glypicans (Reichsman, Smith and Cumberledge, 1996).

In contrast to the SFRPs and WIF1, which are capable of affecting multiple branches of WNT signalling due to their ability to bind to many different WNT ligands (Hsieh *et al.*, 1999;

Cruciat and Niehrs, 2013), DKK1, 2 and 4 bind to LRP5/6 (**Figure 1.6, c**) and therefore act specifically on β -catenin-dependent WNT signalling (Semenov *et al.*, 2001). DKKs are incapable of binding to WNT and cannot prevent its interaction with FZD receptors; rather, they exert their actions by preventing the formation of functional WNT/FZD/LRP5/6 complexes by binding to the LRP5/6 extracellular domain with low-nanomolar affinity to block its interaction with WNT (Semenov *et al.*, 2001). DKKs bind to LRP5/6 via a C-terminal CRD that is related to the colipase fold and therefore compete with WNT ligands for coreceptor binding (Brott and Sokol, 2002). However, while DKK1 functions exclusively as an inhibitor of WNT/ β -catenin signalling, DKK2 can also function as a weak activator of LRP5/6 (Brott and Sokol, 2002). Therefore, although it can antagonise WNT signalling through preventing the formation of functional WNT-FZD-LRP5/6 complexes, DKK2 may also activate WNT/ β -catenin signalling in certain cellular contexts (Brott and Sokol, 2002). As the C-terminal colipase folds of DKK1 and 2 in isolation are both able to exert this weak activating effect, the N-terminal DKK-specific CRD (DKK_N) appears to function as a regulatory domain that specifies the outcome of DKK-LRP5/6 interactions (Brott and Sokol, 2002). A further layer of complexity to the role of DKK is its ability to bind to and synergise with Kremen single-pass transmembrane receptors (KRM1/2, Mao *et al.*, 2002). Upon formation of the DKK/KRM/LRP5/6 ternary complex (Zebisch *et al.*, 2016), LRP5/6 rapidly undergoes endocytosis, thereby reducing the availability of these WNT co-receptors at the cell surface (Mao *et al.*, 2002). Therefore, the contradictory effects of DKK2 may be explained by a model in which Kremen causes endocytosis of DKK2/LRP5/6 complexes, which would otherwise activate β -catenin-dependent signalling, thereby switching DKK2 from an activator to an inhibitor (Mao and Niehrs, 2003).

In contrast to the regulatory molecules discussed above, which exert their inhibitory effects through protein-protein interactions with components of the signalling machinery to prevent

the formation of competent signalling complexes, the WNT deacylase enzyme, Notum, functions by covalently inactivating WNT ligands (**Figure 1.6, d**; Kakugawa *et al.*, 2015). Notum is an extracellular carboxyesterase enzyme, which cleaves the palmitoleic acid moiety on WNT ligands (Kakugawa *et al.*, 2015), thereby abolishing the interaction between this modification and the FZD CRD hydrophobic groove. The Notum crystal structure revealed a large, hydrophobic pocket that specifically recognises *cis*-unsaturated fatty acids, explaining why Notum is able to exclusively cleave the WNT palmitoleic acid appendage while not affecting the activity of other acylated morphogens, like Hedgehog (Kakugawa *et al.*, 2015). Notum has also been found to bind to HSPG GAGs, an interaction that is required for its function (Kakugawa *et al.*, 2015).

FZD itself is subject to regulation by a ubiquitin-dependent degradation pathway. The E3 ubiquitin ligases, zinc and ring finger 3 (ZNRF3) and ring finger protein 43 (RNF43), ubiquitinate FZD, leading to its clearance from the plasma membrane (**Figure 1.6, e, i**; Koo *et al.*, 2012), thereby preventing the activation of WNT signalling. In turn, ZNRF3 and RNF43 are regulated by a class of WNT agonists known as R-spondins (**Figure 1.6, e ii**). These proteins bind to cell surface receptors of the leucine-rich repeat-containing GPCR (LGR) class, such as a LGR5, which are present on proliferative stem cells (De Lau *et al.*, 2011), leading to the formation of a ternary complex with ZNRF3/RNF43 (Hao *et al.*, 2012). Consequently, these WNT antagonists are cleared from the membrane, allowing FZD to accumulate, leading to an increase in WNT signalling activity.

1.8 WNT-FZD pathway selectivity

Understanding WNT signalling is complicated by the fact that such varied outcomes can be triggered by overlapping sets of proteins. Furthermore, the apparent FZD-binding promiscuity that has been observed for certain WNTs raises the question of why the mammalian genome encodes such a large repertoire of WNT ligands and receptors. WNT-FZD specificity, and by

extension pathway selectivity, is generally believed to be dictated by varying affinities between different WNT-FZD CRD pairs (Janda *et al.*, 2012; Dijksterhuis *et al.*, 2015) and spatiotemporal availability of ligands and receptors, as certain WNTs have been demonstrated to bind multiple classes of FZD receptor and *vice versa* (Dijksterhuis *et al.*, 2015).

In particular, the ‘site 2’ interaction between the WNT ‘finger’ and the FZD CRD has been argued to be a key determinant in WNT-FZD binding specificity (Janda *et al.*, 2012). However, crystal structures of WNT3 or *Xenopus* WNT8 in complex with the mouse FZD₈ CRD reveal a high degree of conservation in the WNT amino acids involved in contacting the FZD site 2 (discussed in **section 6.1**). Therefore, while the variability of the FZD site 2 might contribute to different affinities between FZDs for WNTs generally, this site is unlikely to facilitate discrimination between different WNTs, since the interacting residues on the WNT finger are practically invariant.

So how is WNT-FZD specificity actually achieved? One possibility is that the FZD LD may be involved in the WNT-FZD interaction and could allow discrimination between different WNTs (see **section 1.6.5**). Another possible explanation is that more WNT coreceptors may exist. It has recently been shown that WNT9A can elicit β -catenin stabilisation through FZD₉ but not FZD₈ (Grainger *et al.*, 2019). Intriguingly, swapping the CRD and LD of these two receptors did not affect the ability of FZD₉ (or the inability of FZD₈) to activate the canonical pathway when stimulated with WNT9A, indicating that WNT9A can likely bind to the CRDs of both receptors but only activate one, suggesting that pathway activation is specified by another part of the FZD (Grainger *et al.*, 2019). On the other hand, swapping the ICL3 and C-terminal tail of the two receptors was sufficient to reverse their ability to respond to WNT9A, indicating that these regions may somehow be responsible for pathway selectivity (Grainger *et al.*, 2019). Surprisingly, the epidermal growth factor receptor (EGFR) was found to form a

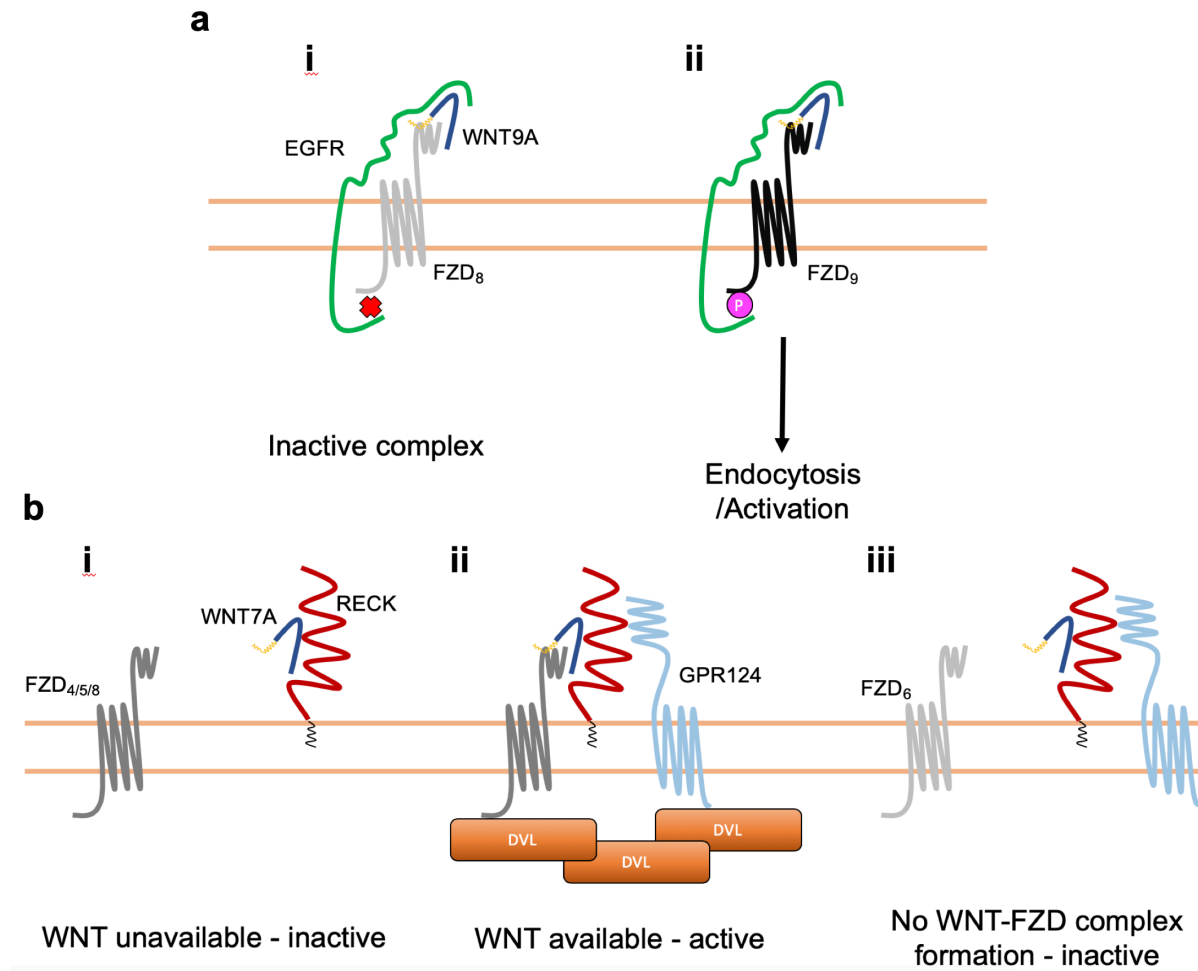


Figure 1.7 Auxiliary coreceptors ‘decode’ WNT-FZD complexes

a, WNT9A can bind to multiple FZDs, however only FZD₉ is phosphorylated in its C-terminal tail by EGFR upon WNT9A binding. This leads to internalisation of the complex and activation of β -catenin-dependent signalling. **b**, RECK binds to WNT7A and, in the absence of GPR124, sequesters the ligand away from FZD receptors. In the presence of both RECK and GPR124, active signalling complexes can be formed with FZDs 4, 5 and 8 but not with FZD₆, despite FZD₆ being able to bind WNT7A.

complex with WNT9A and FZD₉, and to phosphorylate the receptor on a conserved tyrosine residue in the C-terminal tail (**Figure 1.7, a**; Grainger *et al.*, 2019). Furthermore, this complex was observed to undergo clathrin-mediated endocytosis, which was shown to be necessary for β -catenin stabilisation (Grainger *et al.*, 2019). In this context, the prerequisite for EGFR as an additional coreceptor essentially uncouples WNT binding from FZD activation.

Another supplementary coreceptor has been identified that regulates the interaction between WNT7A and FZD (**Figure 1.7**; Eubelen *et al.*, 2018). In this instance, the Reversion-inducing cysteine-rich protein with Kazal motifs (RECK) was demonstrated to function as a receptor for WNT7A, by specifically recognising its N-terminal linker domain, a region that exhibits strong conservation between orthologues in different species, while diverging substantially between WNTs within a species (Eubelen *et al.*, 2018). RECK in turn was found to interact with the orphan GPCR, GPR124, which connects, via DVL, to FZD (Eubelen *et al.*, 2018). Interestingly, in the absence of GPR124, RECK was observed to act as a potent inhibitor of WNT7A/FZD₅ signalling (Eubelen *et al.*, 2018), indicating that GPR124 and RECK together may function as a switch to either permit or prevent a cell from responding to WNT7A. While WNT7A can bind to the CRDs of FZD_{4,5,6} and ₈, only FZD_{4,5} and ₈ can form active complexes with RECK and GPR124, suggesting that either RECK or GPR124 are able to discriminate between different FZDs (Cho *et al.*, 2019). Replacing the FZD₆ CRD with the CRD from FZD₄ or ₈ substantially increased the ability of FZD₆ to activate WNT7A signalling in the presence of GRP124 and RECK, suggesting that discrimination between FZDs is achieved via recognition of the CRD (Cho *et al.*, 2019).

Unlike EGFR, which enzymatically activates WNT9A/FZD₉ complexes, RECK/GPR124 control signalling activation by regulating the availability of a specific WNT and seemingly guiding it to specific FZDs. In both cases, the coreceptor functions to ‘decode’ WNT-FZD combinations.

These examples raise the possibility that other as-yet unidentified WNT coreceptors may exist that could facilitate the activation of certain specific WNT-FZD complexes or permit signalling only in specific contexts. While this is unlikely to be a universal feature of WNT-FZD signalling, as WNT-independent co-localisation of FZD and LRP5/6 is sufficient to trigger β -

catenin-dependent signalling through some FZDs (Janda *et al.*, 2017), such mechanisms would certainly help to explain how seemingly promiscuous WNTs can activate specific receptors.

1.9 The role of FZD in cancer

Given the role of WNT signalling in processes such as cell proliferation, migration and angiogenesis, it is hardly surprising that dysregulation of this signalling network is involved in diseases such as cancer, (reviewed by Zeng, Chen and Fu, 2018). Indeed, the first insights into the function of the WNT gene came from the observation that viral induction of its expression was a common feature of tumours in mice (Nusse and Varmus, 1982). While oncogenic WNT signalling commonly results from the loss of function of downstream regulatory components, such as APC (Cross *et al.*, 1997), or activating mutations of β -catenin that allow it evade its normal fate of proteasomal degradation (Morin *et al.*, 1997), overexpression of FZD receptors at the cell surface is another common feature of cancers (Zeng, Chen and Fu, 2018).

For example, FZD₇ is commonly overexpressed in triple-negative breast cancer (TNBC, King *et al.*, 2012). Inhibition of FZD₇-dependent β -catenin stabilisation was observed to reduce the rate of growth and transformation in TNBC cell lines (King *et al.*, 2012). Furthermore, FZD₂ is overexpressed in some late-stage cases of hepatocellular carcinoma (HCC) and lung cancer and its expression is correlated with poor patient survival (Gujral *et al.*, 2014). In this case, FZD₂ appears to activate a ‘non-canonical’ signalling pathway and knockdown or inhibition of FZD₂ was found to reduce growth and metastasis of HCC cells in a xenograft model (Gujral *et al.*, 2014). FZD₃ has also recently been shown to be overexpressed in some cases of melanoma, and appears to drive tumour progression in a β -catenin-independent manner (Li *et al.*, 2019). Proliferation and colony-forming potential of melanoma cells overexpressing FZD₃ are reduced when FZD₃ is knocked down (Li *et al.*, 2019). Results such as these highlight the role of FZD in tumorigenesis and demonstrate the antitumor potential of therapies that target FZD receptors.

1.10 Therapeutic targeting of FZD receptors

Given the role of FZD in tumorigenesis, combined with its localisation at the cell surface, FZD represents an attractive candidate for therapies of diseases caused by aberrant WNT signalling (**Figure 1.8** and **Table 1**). However, while many GPCRs, including SMO, have been successfully targeted by TMD-binding small molecule anticancer inhibitors (reviewed by Rimkus *et al.*, 2016; Pietrobono and Stecca, 2018), no such inhibitors of FZD have been identified to date. As described in **1.6.6**, the FZD₄ TMD structure revealed a TMD pocket that was narrow and hydrophilic (Yang *et al.*, 2018), which may explain why FZD has so far evaded the kind of therapeutic targeting that has been seen with SMO.

Inhibiting FZD is further complicated by the fact that the 10 human FZDs exhibit a high level of sequence conservation; have numerous functions, which are often overlapping; and are expressed in virtually all cells (Wang *et al.*, 2016). Thus, potential anti-FZD therapies must be tailored to be specific to a single FZD (or small group thereof) so as not to cause widespread off-target effects.

This requirement for specificity, combined with the lack of known small molecule binding sites, has made antibodies a weapon of choice against FZD receptors (**Figure 1.8, b, i**). The first inhibitory antibody against FZD was identified from a phage display library and found to bind to a conserved epitope now known to be proximal to the palmitoleic acid-binding groove on the CRD (Gurney *et al.*, 2012). This monoclonal antibody, known as ‘vantictumab,’ was shown to bind to FZD_{1, 2, 5, 7} and ₈ with low-nanomolar affinity, and was observed to potently inhibit β -catenin-dependent signalling in cell-based assays (Gurney *et al.*, 2012). Furthermore, the antibody was demonstrated to be effective in inhibiting growth of human tumour xenografts, but was not effective against those harbouring mutations in APC or β -catenin, which is unsurprising as these proteins function downstream of FZD (Gurney *et al.*, 2012).

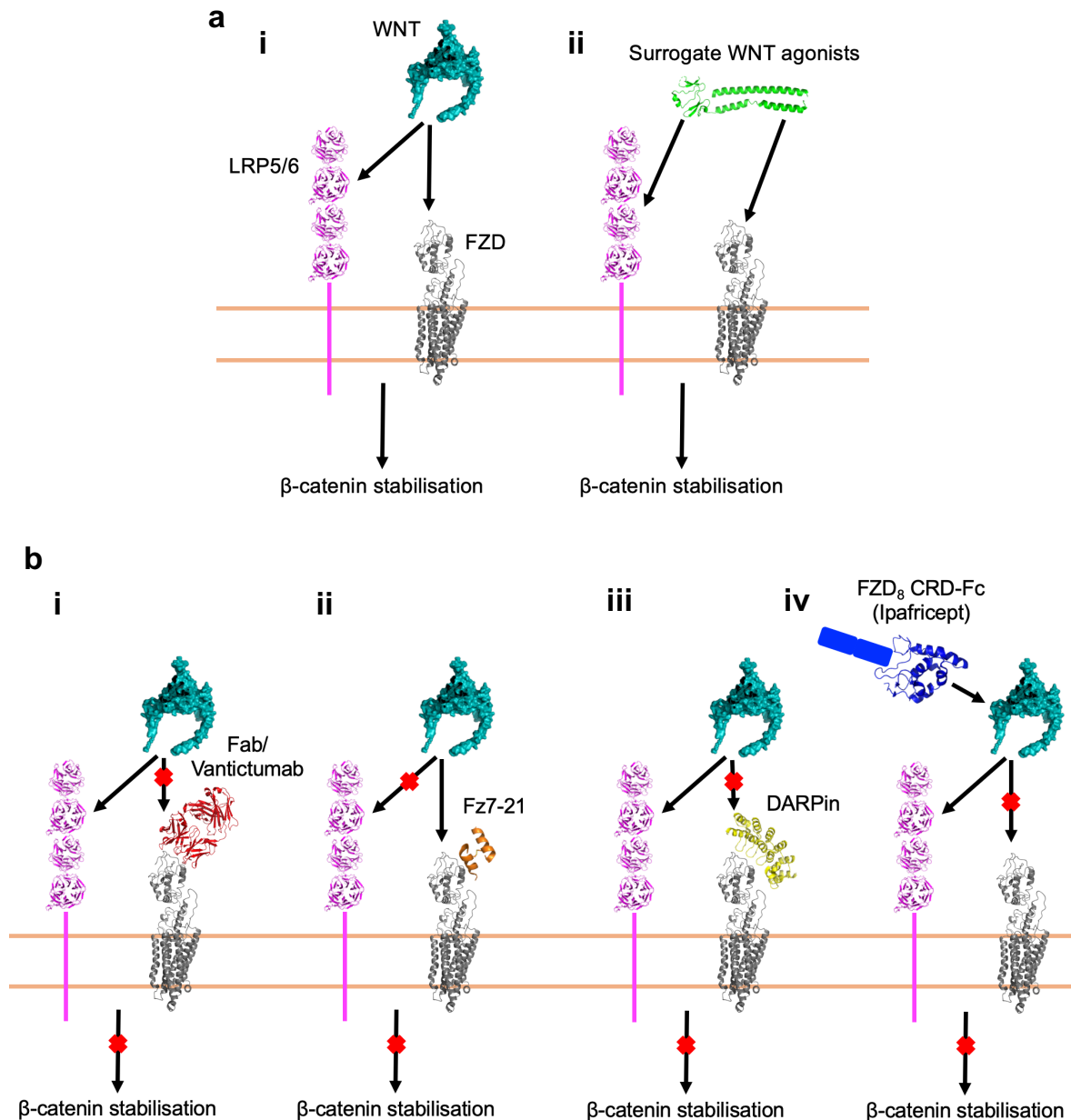


Figure 1.8 Synthetic regulators of WNT signalling

a, activators of WNT signalling, **i**, In the canonical WNT pathway, WNT binds to FZD and LRP5/6, leading to stabilisation of β -catenin. **ii**, Surrogate WNT agonists mimic WNT by combining FZD and LRP5/6-binding proteins. **b**, antagonists of WNT signalling, **i**, Antibodies, such as vantictumab that bind to FZD CRDs prevent the WNT-FZD interaction, thereby inhibiting WNT signalling. **ii**, Similarly, antibodies against FZD CRD also inhibit this interaction. **iii**, The FZD₇-inhibitory peptide, Fz7-21 binds to FZD CRD and appears to prevent the interaction of the WNT-FZD complex with LRP5/6, through an as-yet unclear mechanism. **iv**, Soluble FZD₈ CRD fused to the human Fc domain sequesters WNTs, thereby preventing the formation of active WNT-FZD-LRP5/6 complexes. WNT (teal) from PDB ID 6AHY; FZD (grey) homology model using 5L7D as template; LRP5/6 (magenta) based on 3S2K; Surrogate WNT agonist (green) based on PDB 5UN6 and 3S2K; DARPin from 6NDZ; F2.1 Fab (red) from 6O39; Fz7-21 from 5WBS; FZD₈ CRD (blue) from 1IJY.

Table 1 The 5 subfamilies of Class F GPCRs and their known natural and synthetic interaction partners

Subfamily	Protein	Receptor for WNT?	Receptor for Norrin?	Known small molecule and synthetic protein binders
FZD _{1/2/7}	FZD ₁	✓	✗	DRPB_Fz7/8, DRPB_Fz7, Vantictumab, F ^P , F6, F2.I, F7.B
	FZD ₂	✓	✗	DRPB_Fz7/8, DRPB_Fz7, Vantictumab, F ^P , F6, F2.I, F7.B
	FZD ₇	✓	✗	Fz7-21 peptide, DrPb_Fz7/8, DRPB_Fz7, Vantictumab, F ^P , F6, F2.I, F7.B
FZD _{3/6}	FZD ₃	✓	✗	None
	FZD ₆	✓	✗	None
FZD _{4/9/10}	FZD ₄	✓	✓	FzM1, FzM1.8 DRPB_Fz4, F2.I, F7.B
	FZD ₉	✓	✗	None
	FZD ₁₀	✓	✗	None
FZD _{5/8}	FZD ₅	✓	✗	DRPB_Fz8, Vantictumab, F ^P , F6, F2.I, F7.B
	FZD ₈	✓	✗	DRPB_Fz8, DRPB_Fz7/8, Vantictumab, B12, F ^P , F6, F2.I, F7.B
SMO	SMO	✗	✗	Cholesterol, Taladegib, Vismodegib*, Sonidegib*, Cyclopamine, ANTA XV, SANT-1, SAG1.5

*approved for clinical use by the United States Food and Drug Administration as of February

2020

While vantictumab has shown promise in early stage clinical trials, its adverse effects on bone metabolism highlight the perils of targeting such a ubiquitous signalling network (Tam *et al.*, 2017). Crystal structures of three more antibody inhibitors of FZD_{1,2,4,5,7} and/or 8 (F6, F2.I and F7.B; **Table 1**), all of which bind at the hydrophobic groove of the FZD CRD and exhibit strong antitumor potential in xenograft studies (Raman *et al.*, 2019), have underscored the potential of targeting the WNT- FZD interaction with antibodies. Antibodies against FZD₅ are

also able to inhibit growth of tumour cells with inactivating mutations in the natural FZD antagonist, RNF43 (Steinhart *et al.*, 2017).

Direct FZD inhibition has also been achieved with a peptide that specifically targets the FZD_{1/2/7} subfamily (**Figure 1.8, b, ii**; Nile *et al.*, 2018). Although this peptide, known as Fz7-21, was not demonstrated to exhibit anticancer potential, it was shown to impair the function of intestinal stem cells, which is reliant on FZD₇ (Nile *et al.*, 2018). The crystal structure of this peptide, in complex with the FZD₇ CRD, shows that it forms a disulphide-linked dimer that binds at the FZD CRD dimerisation interface, causing a substantial displacement in the relative orientation of the CRD monomers (Nile *et al.*, 2018). Intriguingly, Fz7-21 does not prevent binding of WNT3A to the FZD₇ CRD but rather appears to inhibit canonical WNT signalling by preventing the formation of a ternary complex with LRP5/6 (Nile *et al.*, 2018). It is possible that, as the peptide functions only in its dimeric form (Nile *et al.*, 2018), it stabilises a FZD₇ CRD dimer configuration that precludes LRP5/6 recruitment to the WNT:FZD complex.

Another class of synthetic inhibitors, known as designed ankyrin repeat proteins (DARPin)s, have also shown potential for targeting FZD (**Figure 1.8, b, iii**; Dang *et al.*, 2019). Three FZD subtype-specific DARPins were generated that bound FZD_{1/2/7} (DRPB_Fz7), FZD₄ (DRPB_Fz4), and FZD_{5/8} (DRPB_Fz8), respectively, while a fourth bound FZD_{1/2/7} and FZD₈ (DRPB_Fz7/8; Dang *et al.*, 2019). These proteins were shown to be potent inhibitors of WNT- β -catenin signalling, and those that bound FZD_{5/8} or FZD_{1/2/7} and FZD_{5/8} (but not FZD_{1/2/7} only) were observed to severely impair intestinal homeostasis (Dang *et al.*, 2019). Crystal structures of these DARPins in complex with their cognate FZDs confirm that these inhibitors insert loops into the WNT-binding groove on the FZD CRDs, likely preventing them from binding WNTs (Dang *et al.*, 2019).

Interestingly, FZD inhibition can also be achieved by soluble FZD CRD, which acts as a decoy receptor to sequester WNTs, preventing them from activating signalling (**Figure 1.8, b, iv**; DeAlmeida *et al.*, 2007). The FZD₈ CRD, fused to the human Fc region, was demonstrated to be effective against tumour xenografts (DeAlmeida *et al.*, 2007). Clinical trials are underway for this reagent, known as ipafricept, however as with vantiactumab, bone metabolism side effects have been reported (Tam *et al.*, 2017).

A small molecule chaperone specific to FZD₄, known as FzM1, has also been shown to inhibit WNT signalling in cell-based assays (Generoso *et al.*, 2015). This inhibitor, which has been suggested to bind to a site involving the intracellular loops (Generoso *et al.*, 2015), has also been modified to yield a ‘non-canonical’ agonist of FZD₄, called FzM1.8 (Riccio *et al.*, 2018). While the inhibition of WNT signalling is of potential therapeutic value against cancers, insufficient pathway activation is implicated in certain degenerative disorders, such as Alzheimer’s disease (Tapia-Rojas and Inestrosa, 2018). Consequently, WNT agonists are also of potential value in the treatment of disease. Synthetic WNT agonists have been created by fusing FZD CRD-binding proteins with proteins that bind to LRP5/6 (**Figure 1.8, a, ii**; Janda *et al.*, 2017). In this case, two WNT inhibitors, the anti-FZD antibody, vantiactumab; and the C-terminal LRP5/6-binding domain of DKK1, were fused together, generating a potent agonist of β -catenin-dependent WNT signalling (Janda *et al.*, 2017). This raises the possibility that more WNT agonists that exhibit FZD subtype specificity can be created by combining other FZD-binding modules with the DKK1 LRP5/6-binding domain. A second WNT agonist was created by fusing the DKK1 C-terminal domain with ‘B12,’ a designed FZD CRD-binding protein comprising a peptide that bound the FZD₈ lipid-binding groove integrated into a scaffold that exhibited shape complementarity for FZD₈ (Janda *et al.*, 2017). Remarkably, the potency of the vantiactumab-based agonist was increased by approximately 200-fold in the presence of the WNT potentiator RSPO2 (see **section 1.7**) at a concentration of 75 nM, an

effect that was mimicked by fusion of RSPO2 to the C-terminus of the agonist itself (Janda *et al.*, 2017). FZD stimulation has also been achieved by fusing the DKK C-terminal LRP5/6-binding domain to ‘non-canonical’ WNTs, providing further evidence that FZD-LRP5/6 heterodimerisation is sufficient for β -catenin-dependent signalling, at least in certain contexts (Grumolato *et al.*, 2010).

More recently, synthetic WNT agonists were described that comprise two FZD-binding antibodies and two LRP5/6-binding antibodies (Tao *et al.*, 2019). These tetravalent antibodies function similarly to the agonists involving DKK fusions (Janda *et al.*, 2017), bringing together FZD and the LRP5/6 co-receptor to form an active complex. Interestingly, the most potent agonist activity was achieved by a tetravalent antibody that simultaneously binds to two different LRP5/6 epitopes, one that is required for WNT3A binding and one that is required for WNT1 binding (Tao *et al.*, 2019). Mutating one FZD-binding domain resulted in moderate agonist activity, hinting that while dimerisation of FZD by WNT may not be a prerequisite for β -catenin-dependent WNT signalling, it may be required for maximal signalling activity (Tao *et al.*, 2019). Unlike DKK-based agonists, however, bivalent antibodies were incapable of activating β -catenin-dependent WNT signalling and instead functioned as potent antagonists (Tao *et al.*, 2019).

1.11 Aims and outcomes of the project

The primary aim of this project was to determine for the first time a crystal structure of a FZD protein containing the CRD, LD and TMD. Such a structure is needed to improve understanding of how FZD receptors function. For example, it is unclear whether there are stabilising interactions between the CRD and the LD or ICLs, as seen in SMO, or whether the FZD CRD has more flexibility, as has been suggested (Yang *et al.*, 2018). A multidomain FZD structure may also provide insights into WNT signalosome assembly and the role of dimerisation in FZD activation (Zhang, Dong and Xu, 2018). Efforts towards this goal included

screening a representative panel of FZD receptors for expression and monodispersity and improving the expression and stability of these proteins through the use of fusion proteins and thermostabilising point mutations. Large-scale expression and purification strategies were also fine-tuned to further optimise FZD production and to overcome an issue of sample heterogeneity. While a multidomain FZD structure remains to be determined, this work, discussed in **Chapter 3** substantially improved FZD production and the suitability of the sample for crystallisation, resulting in initial crystal hits.

A closely related aim of the project was to generate anti-FZD single-domain antibodies or ‘nanobodies’ (Nbs) to both stabilise FZD receptors for structural studies and to investigate as potential modulators of WNT signalling. **Chapter 4** describes the determination of a high resolution FZD₃ CRD crystal structure, facilitated by a Nb, the first structure of a FZD from the FZD_{3/6} subfamily and the first FZD:Nb structure. This structure revealed a potential inhibitory role for the antibody. This work also involved biophysical characterisation of the affinity and specificity of this Nb for FZD₃.

Following biophysical and structural characterisation of this anti-FZD₃ Nb, efforts were undertaken to investigate whether it could inhibit FZD₃. However, FZD₃ was found to not activate β -catenin-dependent signalling in response to the WNTs available for testing. Consequently, work was performed to investigate whether the anti-FZD₃ CRD Nb could be engineered to activate FZD₃- β -catenin signalling via fusion to an LRP-binding protein. The success of this fusion protein in activating signalling permitted experiments to determine the ability of different FZD₃ variant proteins to transduce signals. Interestingly, a disease-causing mutation in a perfectly conserved Class F GPCR residue was shown to not affect ‘canonical’ signalling. Instead, by examining this residue in FZD₄ and SMO TMD structures, a potential molecular ‘switch’ was identified, suggesting that this mutation may prevent FZD-G protein signalling. These results are described in **Chapter 5**.

The final aim of the project was to improve the current structural understanding of FZD CRDs. To this end, the crystal structure of the FZD₁₀ CRD was determined, the first structure of a FZD from the FZD_{4/9/10} subfamily other than FZD₄, which is unusual due to its ability to be activated by Norrin (Smallwood *et al.*, 2007). This revealed a more typical FZD CRD dimerisation interface in FZD₁₀, unlike that seen in FZD₄. This structure also provides a clearer understanding of the determinants of Norrin-FZD₄ specificity. Furthermore, the FZD₃ CRD structure allows structural comparisons between a FZD_{3/6} member and other FZDs for the first time, exposing a substantial rearrangement of the WNT binding site 2 in this subfamily. Details of this work can be found in **Chapter 6**.

Chapter 2 Methods

2.1 Molecular biology

2.1.1 Synthetic genes and expression vectors

Synthetic genes encoding human FZD₁, ₂, ₃, ₈ Δ 173-203, Δ 342-376, Δ 620-694, C617S and ₁₀ (Uniprot ID Q9UP38, Q14332, Q9NPG1, Q9H461 and Q9ULW2, respectively) were obtained from Invitrogen.

The pHLsec vector (Aricescu, Lu and Jones, 2006) was used for transient expression of recombinant multidomain FZD proteins in mammalian cells. DNA encoding FZD constructs was cloned into the pHLsec vector. In the case of constructs containing a native signal peptide, the EcoRI/KpnI restriction sites were used. The forwards primer for each construct encoded a Kozak sequence, such that the final construct would encode a Kozak sequence, followed by the target protein (including native signal peptide) and C-terminal tags. In the case of constructs lacking a native signal peptide, the AgeI/KpnI restriction sites were used. These constructs therefore encoded a Kozak sequence, followed by the pHLsec signal peptide, gene of interest, and C-terminal tags. In all cases, the signal peptide is cleaved during protein secretion, resulting in a mature protein containing an N-terminal ETG sequence and C-terminal GT sequence, followed by the tags. For large-scale expression, the constructs were cloned into a pHLsec vector with C-terminal 1D4 tag only.

The pHR-CMV-TetO₂ vector (Elegheert *et al.*, 2018) was used to produce stable cell lines via lentiviral transduction. DNA sequences encoding multidomain FZD constructs, FZD CRD constructs, the Nb NB₁₄₄₇₈ and biotin ligase were cloned into this vector using the EcoRI/AgeI and KpnI restriction sites, as for pHLsec.

2.1.2 Construct design and polymerase chain reaction

For multidomain FZD construct design see **section 3.3**. DNA encoding the gene (or region thereof) of interest was amplified using the polymerase chain reaction (PCR, Mullis et al., 1986). PCR reaction volumes of 50 μL were used, which typically contained 0.5 μL of template DNA (100-1000 ng μL^{-1}), 0.5 μL forward primer (100 μM), 0.5 μL reverse primer (100 μM), 5 μL 10x Pyrobest buffer (Takara), 4 μL dNTP mixture (2.5 mM of each) 0.5 μL PyrobestTM DNA polymerase (Takara, 5 units μL^{-1}), 29 μL H₂O and 10 μL 5x PCR enhancer mix (2.7 M betaine, 6.7% DMSO, 6.7 mM DTT, 55 mg mL^{-1} , Ralser et al., 2006).

A T100 thermal cycler (BioRad) was used for PCR thermocycling using the following protocol:

1. 94°C, 180 seconds to denature template DNA duplex
2. 94°C, 30 seconds to denature template DNA duplex
3. 55°C, 30 seconds to allow primer annealing to template DNA
4. 72°C, 1 minute per kb of amplified sequence for primer extension/DNA synthesis
5. Repeat from step 2. 20 times
6. 72°C, 300 seconds to ensure complete DNA synthesis

Overlap extension PCR (Ho et al., 1989) was used for site-directed mutagenesis, insertion of the thermostabilised apocytochrome b₅₆₂ containing mutations M7W, H102I and R106L (BRIL) into constructs and generation of the chimeric ROR2/LRP6 construct.

Agarose gel electrophoresis was performed to confirm the presence of the expected DNA product. Bands were excised from the gel and purified using the QIAquick Gel Extraction Kit (QIAGEN).

2.1.3 Restriction enzyme double digest, ligation and transformation

Purified DNA products and desired vectors were digested for 1-16 hours using the necessary restriction enzymes (New England Biolabs) according to the manufacturer's instructions.

Digested PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN), while digested vector DNA was purified by agarose gel electrophoresis and gel extraction as described in section 2.1.2. Ligation reactions typically involved 100 ng of vector DNA and a ten times molar excess of insert DNA. These were mixed with 1 μ L of T4 DNA Ligase Buffer (10X, NEB) and 0.3 μ L of T4 DNA Ligase (NEB) and the reaction volume was adjusted to 10 μ L with H₂O. Ligation reactions were incubated at room temperature for 10 minutes. 20 μ L of Subcloning EfficiencyTM DH5 α competent cells (Invitrogen) were then added to the ligation reaction and incubated on ice for five minutes. Cells were then heat shocked at 42°C for 30 seconds and spread on lysogeny broth (LB) plates containing carbenicillin (MELFORD) at a concentration of 100 μ g mL⁻¹. Plates were incubated at 37°C for 18 hours. Individual colonies were picked and inoculated into 10 mL of LB containing 100 μ g μ L⁻¹ carbenicillin. Cultures were incubated at 37°C in a shaking incubator set to 180 rpm for 18 hours. Plasmid DNA was then purified using the QIAprep Spin Miniprep Kit (QIAGEN). Purified plasmid DNA was validated by Sanger sequencing (Eurofins). Lists of constructs in the pHLsec and pHR-CMV-TetO₂ vectors can be found in **Appendix 5** and **Appendix 6**, respectively.

2.2 Nanobody production

Camelid single chain antibodies (Nanobodies/Nbs) against FZD₃ were produced by the Steyaert laboratory in the Vlaams Instituut voor Biotechnologie at Vrije Universiteit Brussel according to established protocols (Pardon *et al.*, 2014). Approximately 1 mg of FZD₃ 26-517 S215A S334A G393V was produced for immunisation and another 1 mg of each protein for selection. A single animal was repeatedly immunised with FZD₃ 26-517 S215A S334A G393V. Following library generation, Nbs were selected by phage display by trapping biotinylated FZD₃ 26-517 S215A S334A G393V or by the direct adsorption of FZD₃ protein with the CRD replaced with BRIL.

2.3 Protein expression

2.3.1 Small-scale expression

FZD, NB₁₄₄₇₈, ROR2/LRP6 chimera and biotin ligase constructs were expressed in human embryonic kidney (HEK) cells. For structural studies, the α -1,3-mannosyl-glycoprotein 2- β -N-acetylglucosaminyltransferase-deficient variant (GnTI), HEK293S GnTI- (Reeves *et al.*, 2002), was used. This cell line lacks the GnTI activity necessary to produce complex N-linked glycans, limiting N-linked glycosylation to a simple, homogeneous Man₅-GlcNac₂ structure, which can be enzymatically removed by the endo- β -N-acetylglucosaminidase F1 (Endo F1) enzyme, thus avoiding the presence of complex glycans, which can interfere with protein crystallisation. HEK293S GnTI- cells are also amenable to suspension culture, which can support higher densities compared to adherent cell culture. This has the advantage of requiring less incubator space and growing suspension cultures in reusable glass flasks mitigates some of the environmental impact of using numerous disposable plastic roller bottles. Cells in suspension are also more easily harvested than adherent cells and can be adapted to serum-free media. FZD proteins for interaction studies were produced in HEK293T cells so that they would contain complex glycans.

Adherent cells were grown and maintained in a humidified incubator at 37°C with 5% CO₂ in complete medium, comprising Dulbecco's Modified Eagle's Medium – high glucose (DMEM, Sigma), supplemented with L-glutamine and non-essential amino acids (GIBCO), and 10% foetal bovine serum (FBS, Sigma).

Initial small-scale expression tests were performed using HEK293T cells in Costar® tissue culture-treated 6-well plates (Corning). $\sim 0.5 \times 10^6$ cells were seeded in 2 mL of medium per well. After 24 hours, media was replaced with DMEM containing 2% FBS. 4 μ g of plasmid DNA was mixed with 8 μ L Lipofectamine 2000 transfection reagent (Invitrogen) in 100 μ L of serum-free DMEM and incubated at room temperature for 20 minutes. Each transfection

mixture was then added to one well. Proteins were expressed for 48 hours incubation at 37°C in a humidified incubator containing 5% CO₂. For membrane protein constructs, cells were gently suspended in the culture medium by pipetting, pelleted by centrifugation at 800 x g for 4 minutes, washed with 1 mL phosphate-buffered saline (PBS) and pelleted as before. If cells were not to be processed immediately, they were flash-frozen using liquid nitrogen and stored at -80°C until needed. For secreted constructs, media was harvested and centrifuged at 15,000 x g for 15 minutes to remove debris. Membrane protein constructs were analysed by fluorescence size-exclusion chromatography (SEC), whereas secreted proteins were analysed by western blot.

2.3.2 Fluorescence size-exclusion chromatography analysis of multidomain FZD constructs

Each cell pellet was resuspended in 400 µL suspension buffer (25 mM HEPES, pH 7.5, 300 mM NaCl, protease inhibitor cocktail (Sigma)). 60 µL of 10% n-Dodecyl β-D-maltoside (DDM, low alpha, Anatrace) in a 5:1 ratio (w/w) with cholesteryl hemisuccinate (CHS, DDM 5:1 CHS) was then added and samples were incubated at 60 rpm, 4°C for 90 minutes. Samples were then centrifuged at 15,000 x g for 10 minutes to remove insoluble material. 200 µL of suspension buffer was then added to dilute the detergent. 75 µL of slurry (~30 µL beads) containing CNBr-activated sepharose 4B beads (GE Healthcare) coupled to anti-Rho-1D4 antibody (University of British Columbia) was then added and samples were rotated at 60 rpm, 4°C for 3 hours. Beads were then pelleted by centrifugation at 600 x g for 30 seconds and washed twice with 1 mL of suspension buffer containing 0.03 % DDM 5:1 CHS (w/v). Proteins were eluted by overnight incubation with 100 µL of suspension buffer containing 0.03% DDM 5:1 CHS and 500 µM TETSQVAPA peptide. 10 µL of supernatant from each sample was then loaded on a Superose 6 Increase 3.2/300 column (GE Healthcare) equilibrated with 10 mM

HEPES, pH 7.5, 150 mM NaCl and 0.03% DDM 5:1 CHS using an HPLC system with automatic micro-volume sample loader and fluorescence detection (Shimadzu).

2.3.3 Western blot analysis of secreted constructs

15 µL of conditioned media containing each construct was mixed with 5 µL of reducing loading buffer and proteins were resolved by size using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) using established protocols. The polyhistidine-tagged Benchmark™ Protein Ladder and Benchmark™ Pre-Stained Protein Ladder (Invitrogen) molecular weight standards were used to confirm the molecular weight of detected proteins and to monitor transfer of samples to the membrane. Proteins were transferred from the polyacrylamide gel to a polyvinylidene difluoride (PVDF) membrane (Hybond P, 0.45 µm pore size, Amersham) using a Mini Blot Module (Invitrogen) in NuPAGE Transfer Buffer (Invitrogen), 20V for 75 minutes. The membrane was then blocked with 5% non-fat milk powder in PBS containing 0.05% Tween 20 (PBS-T) for 2 hours at room temperature or at 4 °C for 16 hours. For C-terminally polyhistidine-tagged proteins, the membrane was then probed with a monoclonal mouse anti-PentaHistidine primary antibody (QIAGEN, 1 in 1000 dilution in PBS-T) for 1 hour at room temperature. The membrane was then washed four times for five minutes with PBS-T at room temperature and probed with a goat anti-mouse IgG antibody conjugated to horseradish peroxidase (HRP) for one hour at room temperature. The membrane was then washed four times with PBS-T for five minutes at room temperature and HRP was detected using the Clarity™ Western Enhanced Chemiluminescence (ECL) Blotting Substrate kit (Bio-Rad) and visualised with a ChemiDoc imager (Bio-Rad). For Twin-Strep-tagged and biotinylated proteins, the membrane was probed with streptavidin conjugated to HRP and detected as above.

2.3.4 Generation of lentiviral cell lines

Lentiviral cell lines were produced as described previously (Elegheert *et al.*, 2018) and the workflow is summarised in **Figure 2.1**. Cells were maintained in DMEM/Nutrient Mixture F-12 (DMEM/F12, Gibco) supplemented with 10% FBS, L-glutamine and non-essential amino acids (Gibco). A lentiviral producer cell line (HEK 293 Lenti-X) was transfected in 6-well, T75 or T175 format at ~80% confluency. 1.3 μg (6-well), 10 μg (T75) or 25 μg (T175) each of the second-generation lentiviral envelope plasmid (pMDG2.G, Addgene plasmid no. 12259), the second-generation lentiviral packaging plasmid (psPAX2, Addgene plasmid no. 12260) and the pHR-CMV-TetO₂ transfer plasmid encoding the gene of interest were mixed with 10 μL , 75 μL or 190 μL , respectively, of 25 kDa branched polyethyleneimine (PEI, 1 mg mL⁻¹, Sigma) and serum-free DMEM/F12 to a volume of 50 μL , 500 μL or 1 mL, respectively, vortexed for 10 seconds and incubated at room temperature for 15 minutes. Media on the cells was then replaced with 2 mL (6-well), 12 mL (T75) or 30 mL (T175) 2% FBS DMEM/F12 and transfection mixtures were added to cells. After 72 hours, virus-containing media was harvested from the cells, filtered using a 0.45 μm polyethersulphone syringe filter unit (Merck), mixed with 0.5 vols 2% FBS DMEM/F12 and polybrene infection reagent (hexadimethrine bromide, Merck) to a final concentration of 10 $\mu\text{g mL}^{-1}$. Media was removed from the host cell line and replaced with virus-containing medium. After three days, cells were split as required. For multidomain FZD₃ constructs, the host cell line was a HEK293S GnTI- expressing the Tet repressor protein (HEK293S GnTI- TetR), which facilitated inducible target gene expression. Virus from a single well of a 6-well plate, T75 flask or T175 flask was used to infect a single well of a 6-well plate, T75 or T175 flask, respectively, of host cells. For secreted constructs for structural studies the target cell line was HEK293S GnTI- to facilitate constitutive protein expression. Typically, transfections were performed in T75 flasks and virus-containing medium was diluted to a total volume of 120 mL in 5% FBS DMEM/F12 and used to infect

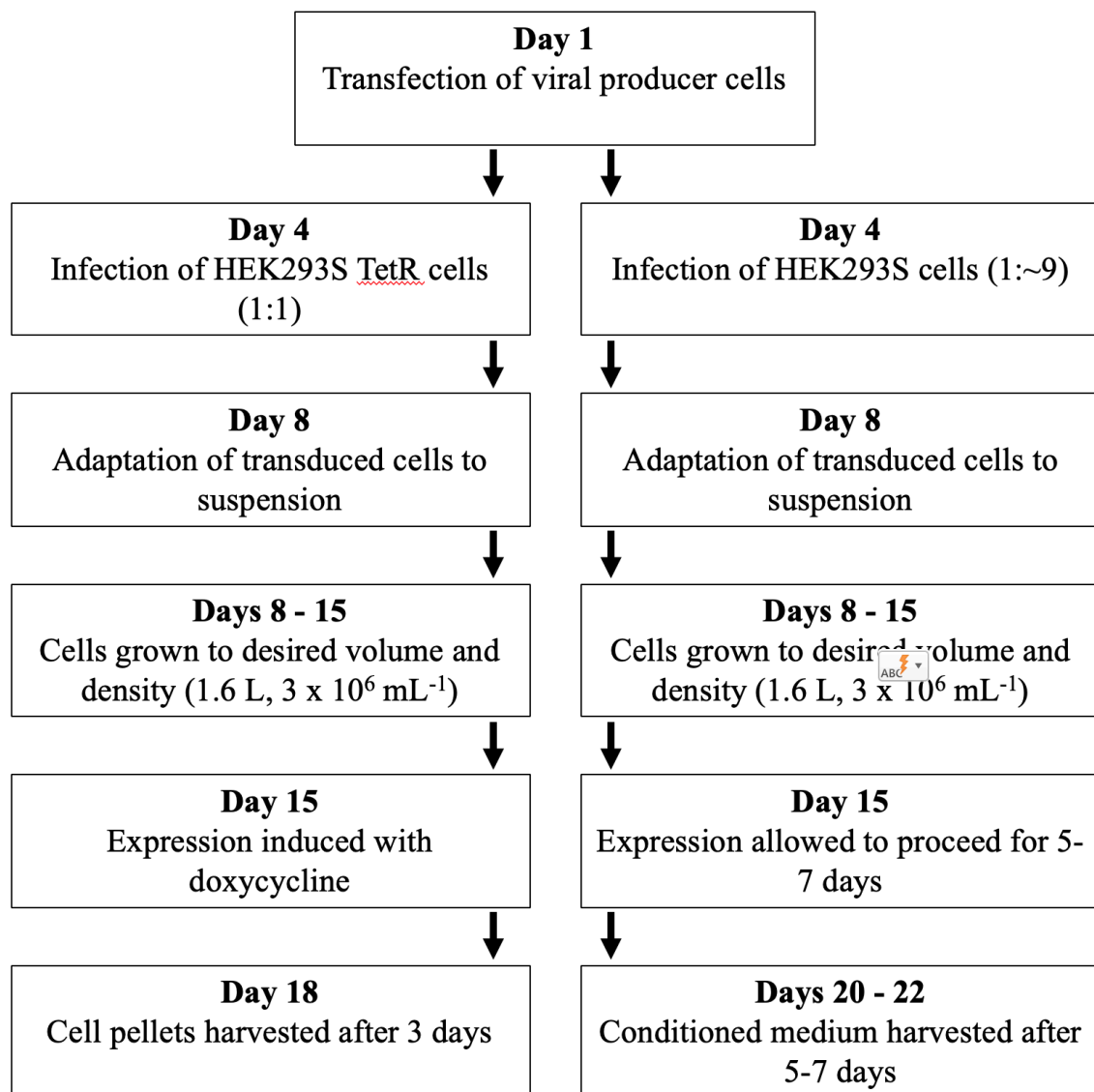


Figure 2.1 Lentiviral transduction for the creation of stable cell lines

Summary of the workflow for the creation of stable cell lines using lentiviral transduction for the inducible (left) and constitutive (right) expression of proteins.

four T175 flasks. This is because secreted constructs were generally smaller and therefore incorporated into viral particles with high efficiency, leading to a high multiplicity of infection. This, combined with constitutive expression, appeared to lead to the death of the majority of the host cells when higher concentrations of virus were used. While the surviving cells were able to express large quantities of protein, this increased the time it took to reach target culture volumes. When the virus was diluted, cells received fewer viral particles each on average and

increased cell death was not observed, meaning that transduced cells could be rapidly expanded to the desired volume, while maintaining high expression levels. For biotinylated constructs, 293T cells were first transduced with ER-localised biotin ligase (Predonzani *et al.*, 2008). The resulting cell line was then transduced with AVI-tagged FZD CRD constructs for *in vivo* biotinylation.

2.3.5 Prokaryotic expression of Nbs

Constructs encoding Nbs (Nbs) were transformed into WK6 *E. coli* cells, which were then plated on carbenicillin agar plates containing 2% glucose (w/v). Individual colonies were then picked and grown overnight in 10 mL of lysogeny broth (Sigma-Aldrich) supplemented with 2% glucose (w/v) and 1 mM MgCl₂ and 100 µg mL⁻¹ carbenicillin. This overnight culture was then added to 500 mL of terrific broth medium (Thermo Fisher Scientific) and shaken at 180 rpm in an incubator at 37°C until OD₆₀₀ = 0.6 was reached. Expression was induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and proceeded at 22°C overnight.

2.4 Purification protocols

2.4.1 Purification of Nbs from *E. coli*

Cell pellets were harvested by centrifugation at 4000 x g for 10 minutes at 4 °C and resuspended in 12 mL STE buffer (0.5 M sucrose, 0.2 M Tris, pH 8.0, 0.5 mM EDTA). The resuspended pellets were rotated at 4 °C for 4 hours at 120 rpm. 20 mL of ice-cold deionised water was then added to each sample and samples were rotated at 4 °C at 120 rpm overnight. Insoluble material was removed by centrifugation at 8000 x g for 30 minutes at 4 °C. Each sample was then incubated with 800 µL of Nickel-NTA Agarose slurry (Qiagen) and rotated at 4°C, 120 rpm for one hour. Protein-bound resin was then collected and washed with 20 column volumes of wash buffer 1 (50 mM HEPES, pH 7.5, 0.5 M NaCl), followed by 20 column volumes of wash

buffer 2 (50 mM HEPES, pH 7.5, 0.2 M NaCl, 10 mM imidazole). Nbs were eluted with 10 column volumes of elution buffer containing 20 mM HEPES, pH 7.5, 150 mM NaCl, 250 mM imidazole. Samples were concentrated to a volume of 1 mL and further purified by size-exclusion chromatography (SEC) using a HiLoad Superdex 75 prepgrade 16/600 column (GE Healthcare) equilibrated with 10 mM HEPES, pH 7.5, 150 mM NaCl (**Figure 2.2**). Typical yields were 2-4 mg per litre of culture.

2.4.2 Identification of FZD₃ CRD-binding Nbs

Nb samples were mixed in approximately equimolar ratios with FZD₃ CRD samples and incubated on ice for one hour. Size exclusion chromatography (SEC) was then performed using a Superdex 75 10/300 GL column (GE healthcare), equilibrated with 10 mM HEPES, pH 7.5, 150 mM NaCl. Elution profiles were then compared to those of FZD₃ CRD and a Nb when run separately. Complex formation was indicated by the appearance of a protein peak at a lower elution volume.

2.4.3 Purification of FZD₃ CRD:Nb₁₄₄₇₈ complex and FZD₁₀ CRD

For the FZD₃ CRD:Nb₁₄₄₇₈ complex, a HEK293S GnTI- cell line was generated by lentiviral cotransduction, which constitutively co-expressed FZD₃ CRD₂₆₋₁₃₈ in frame with a C-terminal Rhinovirus 3C protease cleavage site followed by monomeric Venus fluorescent protein and a Twin-Strep tag; and untagged NB₁₄₄₇₈. For the FZD₁₀ CRD, HEK293S GnTI- cells were transduced with a constitutively expressed construct encoding FZD₁₀ CRD₃₁₋₁₅₇ with the same tags as FZD₃ CRD. Cells were subsequently adapted to suspension culture in serum-free Freestyle 293 medium (Gibco) supplemented with 1% MEM non-essential amino acids (NEAA, Gibco) and grown to a density of approximately $3 \times 10^6 \text{ mL}^{-1}$ in a culture volume of 800 mL. Conditioned medium was harvested after a further 5-7 days and loaded onto a Strep-Tactin XT Superflow high capacity cartridge (IBA) at a rate of 3 mL min^{-1} . The resin was then

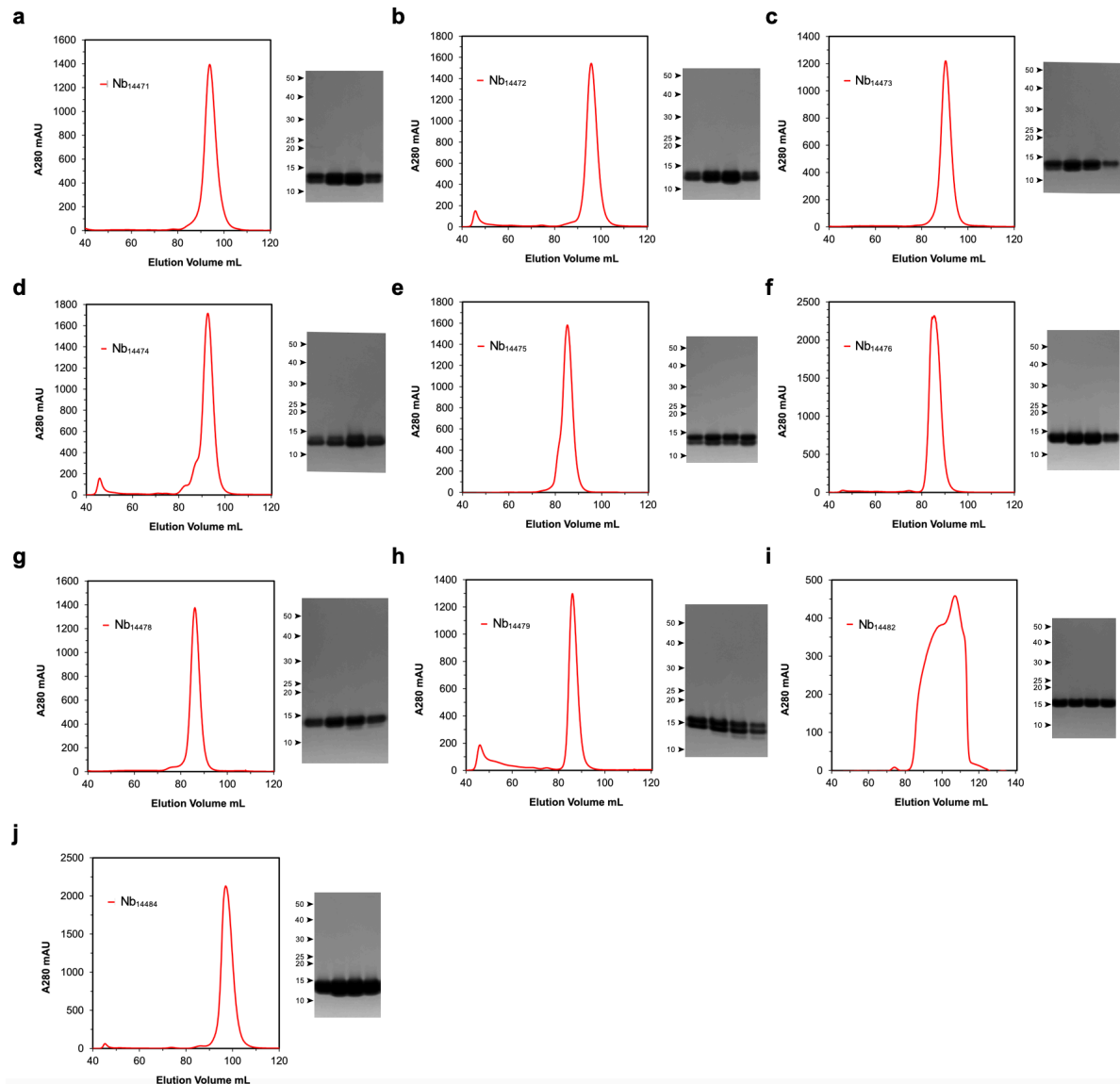


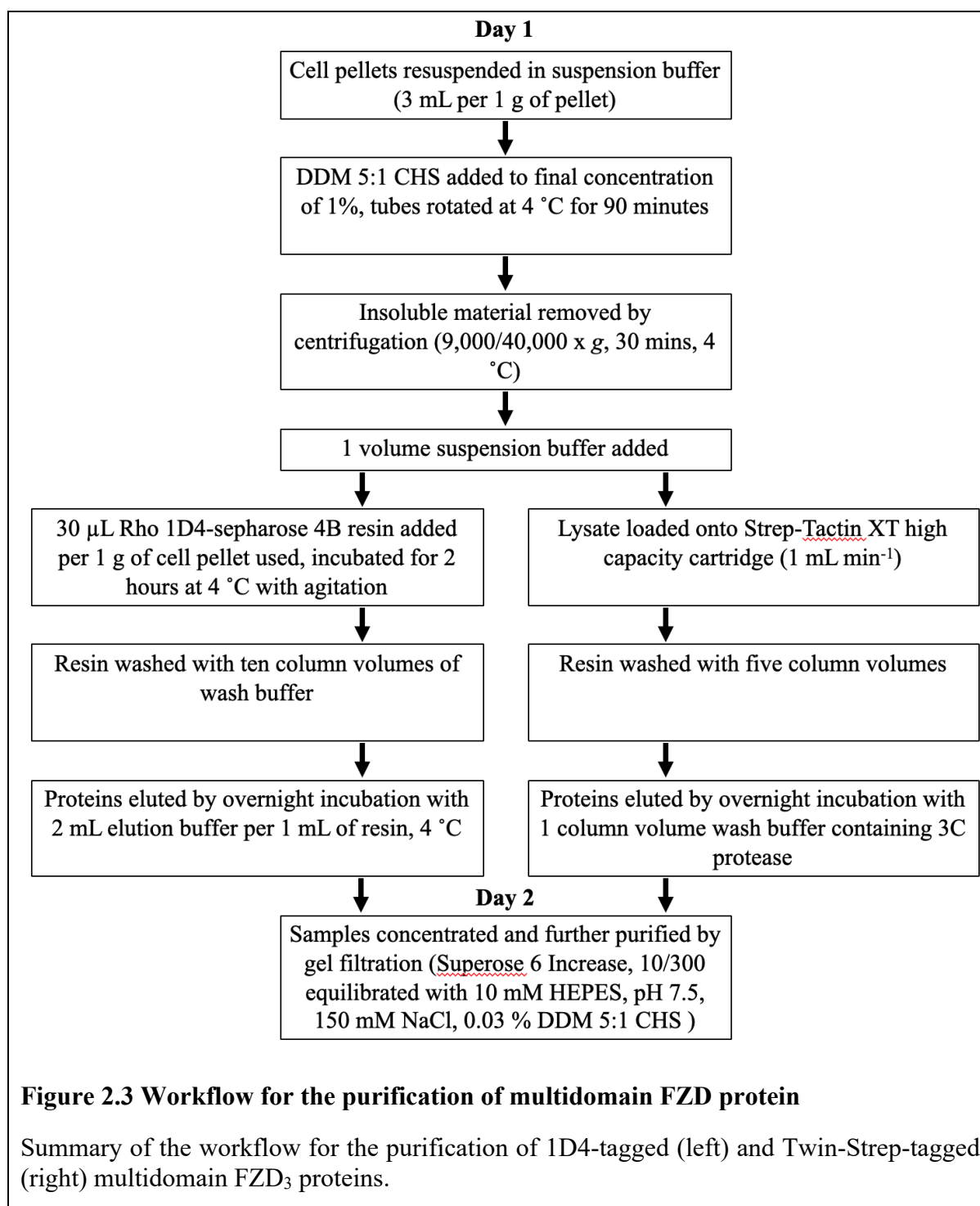
Figure 2.2 Purification Nbs from *E. coli*

Gel filtration profiles (left) and SDS-PAGE analysis (right) of Nb₁₄₄₇₁ (a) – Nb₁₄₄₈₄ (j) purified from *E. coli*. Nb₁₄₄₈₂ (i) produced an unusually broad gel filtration elution peak, indicating a large range of molecular weights, however SDS-PAGE analysis showed that the only species present were of the expected molecular weight. As the elution started at the expected volume, it seems likely that Nb₁₄₄₈₂ was interacting non-specifically with the column, resulting in a protracted elution. Nb₁₄₄₈₂ was also observed to precipitate upon concentration, which may suggest that it was unable to fold properly in *E. coli*.

washed with 5 column volumes of wash buffer (25 mM TRIS pH 8.0, 150 mM NaCl, 1 mM EDTA). The resin was then incubated overnight at 4 °C with 1 column volume of wash buffer containing Rhinovirus 3C protease (~1 mg mL⁻¹) at a dilution of 1:2500 (v/v) and endoglycosidase F1 (~1 mg mL⁻¹) at a dilution of 1:1000 (v/v) to facilitate elution and deglycosylation (when desired) of the protein(s) of interest. Size-exclusion chromatography (SEC) was then performed using a Superdex 75 10/300 GL column (GE Healthcare) equilibrated with 10 mM HEPES pH 7.5, 50 mM NaCl.

2.4.4 Purification of transiently-expressed Rho-1D4 tagged multidomain FZD₃

The work flow for the purification of Rho-1D4-tagged multidomain FZD₃ is outlined in **Figure 2.3** (left). Cell pellets were harvested 72 hours post-transfection by centrifugation at 2000 x g for 5 minutes and snap frozen using liquid nitrogen. Pellets were thawed in tubes in water at 12 °C. 3 mL of suspension buffer containing 10 mM HEPES, pH 7.5, 150 mM NaCl was added per 1 g of cell pellet and cells were resuspended with vigorous pipetting. Membrane proteins were then solubilised by the addition of 10% DDM 5:1 CHS to a final concentration of 1%, before rotation at 15 rpm at 4 °C for 90 minutes. Insoluble material was then cleared by centrifugation at 9,000 x g for 30 minutes at 4 °C. 1 volume of suspension buffer was then added to dilute the detergent in the supernatant. Anti-Rho-1D4 antibody (University of British Columbia) coupled to CNBr-activated sepharose 4B beads (GE Healthcare) was then added at an amount of 30 µL of beads per 1 g of cell pellet and incubated with the supernatant with gentle agitation at 4 °C for 2 hours. Protein-bound beads were then washed with 10 column volumes of wash buffer (suspension buffer containing 10 % glycerol and 0.03 % DDM 5:1 CHS). FZD₃ was eluted by overnight incubation at 4 °C with 2 mL of elution buffer (wash buffer supplemented with 500 µM TETSQVAPA peptide (Genscript) per mL of beads. The sample was then concentrated using an Ultra centrifugal filter with a molecular weight cut-off of 100 kDa (Amicon) and further purified by SEC using a Superose 6 Increase 10/300 column



(GE Healthcare) equilibrated with wash buffer.

2.4.5 Purification of stably-expressed Twin-Strep-tagged multidomain FZD₃

The work flow for the purification of Twin-Strep-tagged multidomain FZD₃ is outlined in **Figure 2.3** (right). A HEK293S GnTI- cell line was generated that inducibly expressed a multidomain FZD₃ construct comprising residues 26-511 with ICL3 replaced with a BRIL fusion (BRIL_FZD₃) in-frame with an N-terminal Twin-Strep tag, followed by a monomeric Venus fluorescent protein and Rhinovirus 3C protease cleavage site. Cells were subsequently adapted to suspension culture in Freestyle 293 (Gibco) medium supplemented with 1% MEM NEAA and grown to a density of approximately $3 \times 10^6 \text{ mL}^{-1}$ in a culture volume of 1.6 L. Expression was then induced by the addition of doxycycline at a final concentration of $1 \mu\text{g mL}^{-1}$.

Cell pellets were harvested 72 hours post-induction by centrifugation for 5 minutes at $2000 \times g$ and snap-frozen using liquid nitrogen. Cell pellets were thawed in tubes in water at 12°C . 3 mL of suspension buffer containing 25 mM TRIS pH 8.0, 150 mM NaCl, 1 mM EDTA was added for each 1 g of cell pellet and cells were resuspended with vigorous pipetting. Membrane proteins were then solubilised by the addition of 10% n-Dodecyl β -D-maltoside (DDM, low alpha, Anatrace) in a 5:1 ratio (w/w) with cholesteryl hemisuccinate (CHS, DDM 5:1 CHS) to a final concentration of 1%, before rotation at 15 rpm at 4°C for 90 minutes. Insoluble material was then cleared by centrifugation at $40,000 \times g$ for 30 minutes at 4°C . 1 volume of suspension buffer was then added to dilute the detergent in the sample, before loading onto a Strep-Tactin XT Superflow high capacity cartridge (IBA) at a rate of 1 mL min^{-1} . The resin was then washed with 5 column volumes of wash buffer (suspension buffer supplemented with 0.03 % DDM 5:1 CHS). Elution was performed by incubating the resin with 1 column volume of wash buffer containing Rhinovirus 3C protease (1 mg mL^{-1}) at a dilution of 1:1000 (v/v) overnight at 4°C .

SEC was then performed using a Superose 6 Increase 10/300 column (GE Healthcare) equilibrated with 10 mM HEPES, pH 7.5, 150 mM NaCl, 0.03 % DDM 5:1 CHS.

2.4.6 Purification of FZD CRD proteins for bio-layer interferometry experiments

HEK 293T cells stably expressing biotin ligase (see **section 2.3.4**) were separately transduced with viruses encoding the FZD₁₋₁₀ CRDs (for construct boundaries see **Appendix 6**). Each resulting cell line was scaled up to a T175 flask and protein expression was allowed to proceed for 7 days in the presence of 100 μ M biotin. Media was then harvested by centrifugation at 9,000 x g for 30 minutes in a 50 mL tube (Greiner). Imidazole was added to each sample to a final concentration of 10 mM. 1 mL of Nickel-NTA Agarose slurry (Qiagen) was then added to each sample, and tubes were rotated at 4 °C for one hour. The resin was then washed with 20 mM imidazole and proteins were eluted with 500 mM imidazole. As these proteins were specifically biotinylated, this level of purity was sufficient for bio-layer interferometry (BLI) experiments, and gel filtration was not necessary.

2.5 Protein crystallisation

2.5.1 Crystallisation of FZD₃:Nb₁₄₄₇₈ complex

Peak fractions containing purified FZD₃ CRD-Nb₁₄₄₇₈ complexes were pooled and concentrated to 6-8 mg mL⁻¹ for crystallisation using an Ultra centrifugal filter with a molecular weight cut-off of 10 kDa (Amicon). Sitting drop vapour diffusion crystallisation trials were performed whereby 100 nL of protein solution was mixed with 100 nL of reservoir solution by robots (Cartesian Technologies) in 96-well crystallisation plates (Greiner). Trials were set up in duplicate, with one plate for each screen stored at 6.5 °C and one at 20 °C in TAP Homebase storage vaults.

Initial crystals of deglycosylated FZD₃ CRD-Nb₁₄₄₇₈ complex grew at a protein concentration of 6.7 mg mL⁻¹ at 20 °C after 7 days in a condition containing 0.8 M ammonium sulphate and

0.1 M citric acid, pH 4.0. These were then optimised using a custom screen in which the ammonium sulphate concentration was varied from 0.4 M to 1.2 M and the citric acid pH was varied from 3.6 to 4.4. Crystals grew across the entire pH range tested and at ammonium sulphate concentrations ranging from 690 mM and 1.2 M at 20°C. Crystals appeared at times ranging from <1 day to 6 days post-setup. The crystal used for data collection was grown in a condition containing 690 mM ammonium sulphate and 0.1 M citric acid, pH 3.6.

2.5.2 Crystallisation of FZD₁₀ CRD

Crystallisation trials were performed as for the FZD₃:Nb₁₄₄₇₈ complex. Initial crystals appeared after two weeks in a condition containing 25 % (w/v) PEG3350, 0.2 M MgCl₂ and 0.1 M Tris pH 8.5. Following optimisation, crystals used for data collection grew in a condition containing 0.1 M Tris pH 9.0, 150 mM MgCl₂ and 34 % (w/v) PEG3350.

2.5.3 Crystallisation of BRIL_FZD₃

Peak fractions containing purified multidomain FZD₃ were pooled and concentrated to 20-50 mg mL⁻¹ using an Ultra centrifugal filter with molecular weight cut-off of 100 kDa (Amicon). BRIL_FZD₃ was then reconstituted into lipidic cubic phase (LCP) by mixing with molten lipid containing a 9:1 ratio (w/w) of monoacylglycerol (MAG, Sigma-Aldrich) and cholesterol (Sigma-Aldrich) at a protein to lipid ratio (v/v) of 2:3, using a mechanical syringe. LCP crystallisation trials were performed using a Gryphon robot (Art Robins instruments), which dispensed a 50 nL bolus of LCP mixture, then 800 nL of precipitant solution, at each of 96 positions on a hydrophobically-coated glass plate, which was then covered with a glass coverslip to create a “sandwich” and stored at 20°C. Small (~5-10 µm in the longest dimension) birefringent crystals grew after approximately 4 days in conditions containing 0.1 M TRIS pH 9.0, 30% PEG500 dimethyl ether (DME) and 0.1 M of either sodium acetate, sodium malonate, sodium nitrate or potassium formate. Larger (~15 µm in the longest dimension) birefringent

crystals grew in conditions containing 0.1 M HEPES, pH 7.0, 30% PEG300 and 0.1 M of either diammonium phosphate or ammonium sulphate.

2.6 Crystal cryo-protection and x-ray diffraction data collection

To reduce radiation damage, crystals were cryo-protected, and x-ray diffraction was performed in cryogenic conditions. Crystals grown by vapour diffusion were transferred from the crystallisation drops to drops containing reservoir solution containing 30% glycerol, to reduce the formation of ice, and then flash frozen using liquid nitrogen. LCP crystals were grown in conditions containing high PEG concentrations and therefore did not require additional cryo-protectants. X-ray diffraction data were collected at the Diamond Light Source (Oxfordshire, UK). The I03 beamline was used for the FZD₃ CRD:Nb₁₄₄₇₈ and the I04 beamline was used for the FZD₁₀ CRD. Crystals were maintained in nitrogen gas cryostream at 100 K during data collection. Diffraction data were collected through a 360° rotation of the crystal (oscillation angle 0.1°), using an EIGER2 X 16M single photon-counting detector (DECTRIS).

2.6.1 X-ray diffraction data processing

For the FZD₃ CRD:Nb₁₄₄₇₈ complex, x-ray diffraction data were processed using the automated autoPROC pipeline (Vonrhein *et al.*, 2011), including the programs XDS (Kabsch, 2010) for data integration and AIMLESS (Evans and Murshudov, 2013) for scaling. For the FZD₁₀ CRD, the XIA2 pipeline (Winter, 2010) was used, including DIALS (Winter *et al.*, 2018) for data integration and AIMLESS (Evans and Murshudov, 2013) for scaling.

2.6.2 Structure determination, model building, refinement and validation

The program Phaser (McCoy *et al.*, 2007) was used to determine structure factor phases using the molecular replacement method (Rossman and Blow, 1962). The crystal structure of the FZD₃ CRD in complex with Nb₁₄₄₇₈ was solved by molecular replacement, first by using the mouse FZD₈ CRD (PDB 1IJY) and a llama-derived Nb against the J-base binding protein 1

(PDB 6FPV) as homology search models. The FZD₁₀ CRD crystal structure was solved by molecular replacement using the mouse FZD₈ CRD (PDB 1IJY) as the homology search model. The PHENIX AutoBuild Wizard (Adams *et al.*, 2010) was then used to rebuild and refine the models. The structures were then further refined using iterative cycles using Coot (Emsley and Cowtan, 2004) and PHENIX (Adams *et al.*, 2010). Data collection and refinement statistics are reported in the relevant results chapters.

2.7 Bio-layer interferometry

BLI qualification of protein interactions was performed using a BLItz System (ForteBio). Streptavidin-coated biosensors (ForteBio) were dipped into samples of biotinylated FZD CRD proteins, followed by samples of Nb₁₄₄₇₈ at a concentration of 5 μ M.

Quantitative measurement of protein interactions was performed using an Octet Red 96e system (ForteBio). Proteins were diluted in a buffer containing 10 mM HEPES, pH 7.5, 150 mM NaCl, 0.02 % Tween-20). Streptavidin-coated biosensors were dipped into samples of biotinylated FZD CRD proteins ($\sim 10 \mu\text{g mL}^{-1}$). After a 30s baseline step, biosensors were dipped into wells containing a serial dilution of Nb₁₄₄₇₈ with concentrations ranging from 200 nM to 3.125 nM. Curves were fitted using a 1:1 binding model and K_ds calculated using ForteBio Data Analysis software.

2.8 Signalling Assays

Luciferase assays were performed using HEK293T cells and HEK293T FZD_{1/2/7}⁻ cells (Voloshanenko *et al.*, 2017) in 96-well plates (Thermo Fisher Scientific). 4×10^4 cells were seeded per well in 100 μ L of complete medium and kept in a humidified incubator at 37°C, with 5 % CO₂. After 24 hours, cells were transfected with various combinations of plasmids (30 ng per plasmid per well), with the total amount of DNA being adjusted to 120 ng with empty vector DNA. Transfections were performed using PEI (0.33 μ L per well).

For assays using WNTs, medium was replaced on all cells after 24 hours. Complete medium was added to control cells, while to stimulated cells, sterile filtered conditioned medium from mouse L fibroblasts expressing WNT3A/WNT5A was added. Nb₁₄₄₇₈ was diluted to medium as required, followed by sterile filtration.

48 hours post-transfection, media on cells was replaced with PBS. Cells were lysed and luciferase activity measured using the Steady-Glo Luciferase Assay System (Promega) according to the manufacturer's instructions.

Chapter 3 Structural studies of multidomain FZD proteins

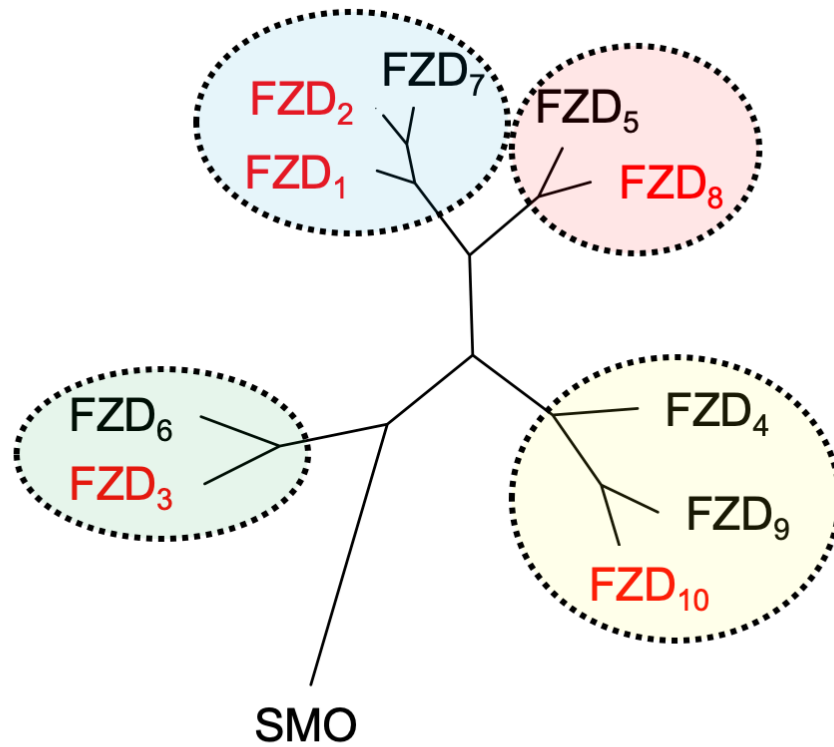
3.1 Introduction

While structural knowledge of FZD receptors has been boosted in recent years by a flurry of CRD structures (Dann *et al.*, 2001; Chang *et al.*, 2015; Shen *et al.*, 2015; DeBruine *et al.*, 2017; Nile *et al.*, 2017; Chen *et al.*, 2018), including two WNT:FZD CRD complexes (Janda *et al.*, 2012; Hirai *et al.*, 2019); as well as a single TMD structure (Yang *et al.*, 2018), understanding of the FZD LD and the interactions between the various FZD domains remains limited. Although SMO structures (Wang *et al.*, 2013; Byrne *et al.*, 2016; Huang *et al.*, 2018; Qi *et al.*, 2019) are a useful supplement to structural understanding of FZD, the functional and structural differences between SMO and FZD, such as the longer LD found in FZD, highlight the need for a multidomain FZD structure to improve knowledge of these receptors, and of WNT-FZD signal transduction more broadly.

At the start of the project, the ten human FZDs were divided between a postdoctoral researcher in the same laboratory, Dr Tao-Hsin Chang, and myself, based on their evolutionary relatedness. As such, this thesis focuses on FZD₁ and FZD₂ from the FZD_{1/2/7} subfamily; FZD₃ from the FZD_{3/6} subfamily; FZD₈ from the FZD_{5/8} subfamily and FZD₁₀ from the FZD_{4/9/10} subfamily (**Figure 3.1, a**).

3.2 Aims

The first aim of the work described in this chapter was to identify a suitable FZD receptor candidate for structural studies. FZD receptors were assessed based on their expression level and monodispersity using gel filtration analysis. The second aim was to optimise FZD expression constructs for structural studies, for example by improving expression levels by

a**b**

	400	410	420	430	440	450
SMO	RAG	FVLAP	IGLVLVIV	GYFLIR	GVM	LFSIKSNHPGLLSEKAASKINETML
FZD3	LR	FVLAP	LCLYVVV	GVSLLLA	GIIS	LNVRVRIEIP..LEKENQDKLVKFM
FZD1	LR	FVLAP	LFVYLF	GTSFLLA	GFVS	LFRIIRTIMKH..DGTKTEKLEKLMV
FZD2	LR	FVLAP	LFVYLF	GTSFLLA	GFVS	LFRIIRTIMKH..DGTKTEKLEKLMV
FZD8	LR	FVLAP	LFVYLF	GTMFLLA	GFVS	LFRIIRSVIKQQDGPTKTHKLEKLM
FZD10	LT	FVLIP	LACYLVI	GTSFILS	GFVA	LFHIRRVMKT..GGENTDKLEKLM

c

	530	540	550	560	570	580
SMO	I	AMSTW	WTKA	TLLI	WRR	TWCRLTGQSDDEPKRIKSKMIAKAFSKRHELL
FZD3	I	PSVFW	WGS	KKTC	FEW	ASFFHGRRKKEIVNESRQVLQEPDFAQSLLRDPNT
FZD1	I	TSGFW	WIS	GKTL	NSW	RKFYTRLTNSKQGETTV.....
FZD2	I	TSGFW	WIS	GKTL	LHS	WRFYTRLTNSRHGETTV.....
FZD8	I	TSGVW	WIS	GKTL	ESW	RSRLCTRCCWASKGAAVGGGAGATAAGGGGGPGGGG
FZD10	I	TSGM	WIS	TSK	TLS	QVCSRRLKKKSRRKPASVITSGGIYKKAQHPQKT

C-terminal Boundary

Figure 3.1 Multidomain FZD construct design.

a, Unrooted phylogenetic tree showing the evolutionary relationships between and subfamilies of human Class F GPCRs. FZDs for which multidomain constructs were made during this project are coloured red. **b**, Sequence alignment of FZD_{1,2,3,8,10} and SMO with ICL3 indicated by the green rectangle and BRIL insertion site indicated by arrows. **c**, Sequence alignment of FZD_{1,2,3,8,10} and SMO with C-terminal boundary of multidomain FZD constructs indicated.

screening different expression boundaries and boosting expression and stability through the use of fusion proteins (Chun *et al.*, 2012). Following identification of a suitable construct, large-scale expression was performed to produce suitable quantities of FZD for crystallisation trials.

As diffraction-quality crystals were not obtained following crystallisation trials, work was then undertaken to identify thermostabilising point mutations in the FZD TMD, overcome issues with sample heterogeneity and improve the efficiency of FZD production through the use of lentiviral cell lines, in an attempt to permit the determination of a multidomain FZD structure.

3.3 Construct design and screening

3.3.1 FZD₁ multidomain construct design and screening

Initially, a strategy was selected whereby several N- and C-terminal boundaries would be screened for a single FZD, in an attempt to identify an optimal expression construct. Therefore, several FZD₁ constructs were designed (**Figure 3.2, a**) and cloned to screen for boundaries that yielded well-expressed proteins. One construct included the full N-terminus of FZD₁ including the long native signal peptide, whereas others started around the predicted start of the CRD. The FZD₁ sequence was aligned with SMO and the C-terminal boundary was selected based on that used for SMO structure determination, BRIL_SMO₃₂₋₅₅₅ (SMO_{XTAL}, **Figure 3.1, c**; **Figure 3.2, f**; Byrne *et al.*, 2016).

For each construct, another was cloned with the thermostabilised apocytochrome b₅₆₂ containing mutations M7W, H102I and R106L (BRIL) replacing ICL3, a common strategy for stabilisation of GPCRs for crystallisation studies (Chun *et al.*, 2012), which has been used successfully for SMO (Byrne *et al.*, 2016). Again, the FZD₁ sequence was aligned with SMO

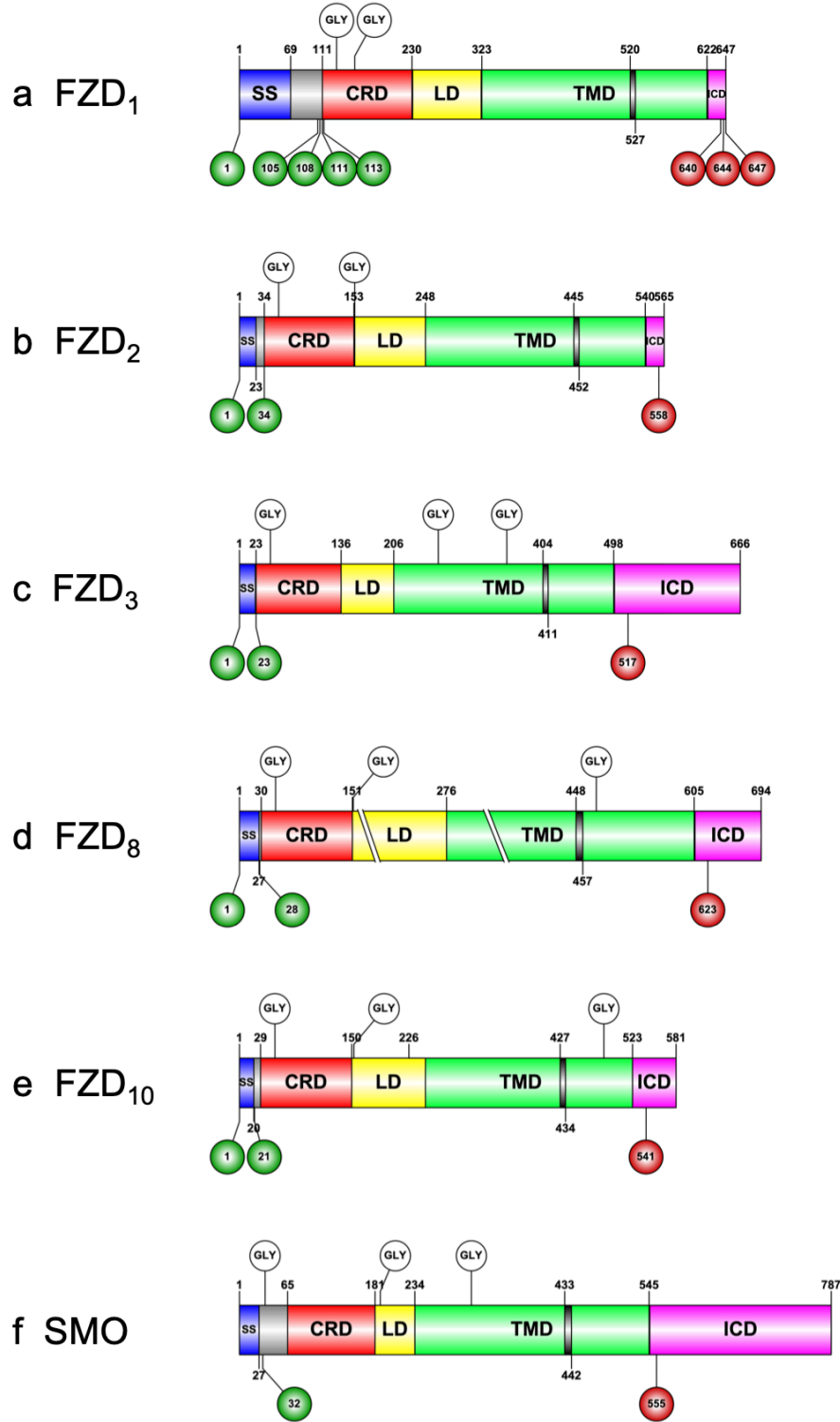


Figure 3.2 Multidomain Class F GPCRs and construct boundaries.

Schematic diagrams of FZD₁ (a), FZD₂ (b), FZD₃ (c), FZD₈ (d), FZD₁₀ (e) and SMO (f). Known and predicted N-linked glycosylation sites are indicated by white circles. N- and C-terminal construct boundaries are depicted by green and red circles, respectively. Internal truncations are indicated by diagonal slashes (FZD₈ only). Sites of BRIL insertions are shown as black bars. SS, signal sequence.

and boundaries for BRIL insertion were selected based on those used for SMO_{XTAL} (**Figure 3.1, b**).

Fluorescence size-exclusion chromatography (FSEC) was used to screen the expression and monodispersity of multidomain FZD₁ constructs (**Figure 3.3**), which were cloned into the pHLsec vector in frame with a C-terminal monomeric Venus fluorescent protein (mVenus) and Rho 1D4 tag. SMO_{XTAL} was used as a positive control so that FZDs could be compared to a closely-related protein that has previously been shown to produce diffraction-quality crystals (Byrne *et al.*, 2016). FZD₁ constructs lacking a BRIL fusion were substantially less well-expressed than SMO_{XTAL} with relative expression levels around one third (**Figure 3.3, a**). BRIL fusion had a positive impact on expression, with these constructs giving expression levels around 80% that of SMO_{XTAL} (**Figure 3.3, b**). However, in all cases FZD₁ constructs produced a peak that appeared to be split, indicating the presence of two distinct species, indicating a possible degradation product. Given that both of these species are detected by FSEC and therefore maintain the C-terminal fluorescent tag, it is very likely that the secondary peak is produced by FZD₁ protein that had undergone CRD cleavage, a phenomenon later observed for FZD₃ (see **section 3.5.2**). Overall, expression levels and elution profiles of FZD₁ constructs did not compare favourably to SMO_{XTAL}. Furthermore, little difference was observed between constructs with different boundaries. For these reasons, the strategy was changed from a thorough screen of an individual FZD to a broader screen for FZD_{2/3/8/10}, with a smaller number of constructs created for each, in order to find a FZD that was better expressed.

3.3.2 FZD_{2/3/8/10} multidomain construct design and screening

For each of FZDs _{2, 3, 8} and ₁₀, a single construct was cloned that maintained the entire N-terminus of the FZD, including the native signal peptide, while another was cloned that utilised the pHLsec signal peptide, with the FZD sequence starting around the predicted start of the

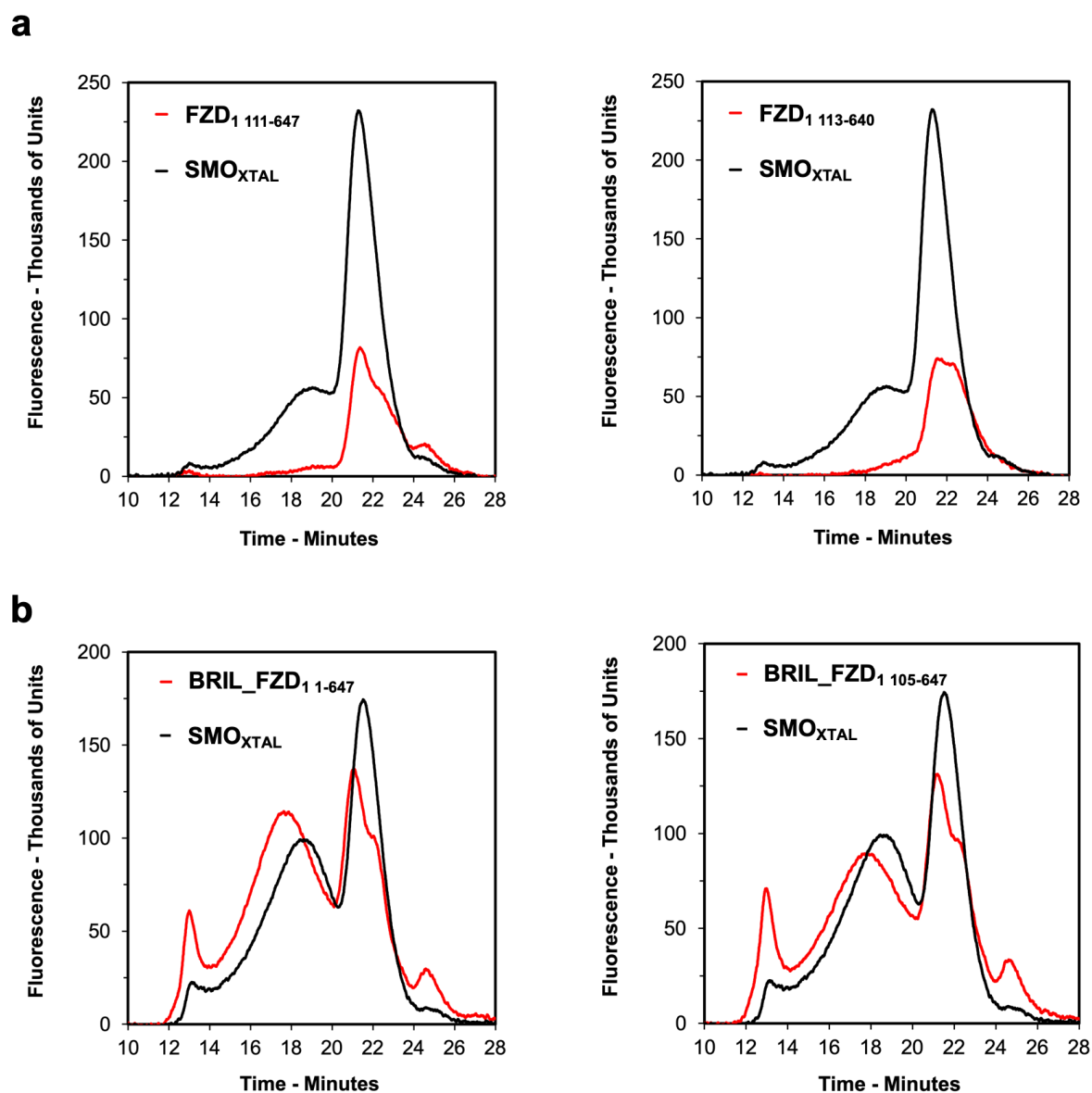


Figure 3.3 FSEC analysis of multidomain FZD₁ constructs.

Representative FSEC profiles from multidomain FZD₁ constructs both without **(a)** and with **(b)** BRIL fusions. SMO_{XTAL} was expressed and purified in parallel to allow comparisons to be made.

CRD (**Figure 3.2 b-e**). This allowed comparisons to be made between the two signal peptides used for each FZD to determine whether higher expression levels could be achieved by using one rather than the other. The C-terminus of each FZD was truncated after the predicted end of H8, based on alignments to SMO_{XTAL} (**Figure 3.1, c**), as the region after this is predicted to be disordered and therefore likely to hinder attempts to crystallise FZD (**Appendix 2**).

For FZDs _{3, 8} and ₁₀, where the sequence of the C-terminal residues of the construct diverge significantly from SMO, mutations were introduced to make the C-terminus more similar to SMO, to remove free cysteines in this region that could have led to intermolecular disulphide bonds that may have increased sample heterogeneity, and to potentially aid protein expression and crystallisation. Residues 512 to 517 of FZD₃ were mutated to SRLTGQ, residues 620 to 623 of FZD₈ were mutated to LTGQ and residues 538 to 541 of FZD₁₀ were mutated to LTGQ. The free cysteine residues C504, in FZD₃; C617, in FZD₈; and C535, in FZD₁₀, were mutated to serine. FZD₈ constructs were further modified by the truncation of two unique, long, unstructured loops (**Appendix 2**) in the LD (Δ 173-203) and ECL1 (Δ 342-376), as these regions would likely impede any attempts to crystallise FZD₈. Except where otherwise stated, all multidomain FZD constructs contain these respective modifications.

For each of these two constructs for each FZD, a second was cloned wherein ICL3 was replaced with BRIL, with boundaries selected based on sequence alignments to SMO_{XTAL}, as for FZD₁ (**Figure 3.1, b**).

FZD₂ constructs were expressed at very low levels, approximately 15% of that observed for SMO (**Figure 3.4, a**). In this case, BRIL fusion did not improve expression noticeably.

FZD₃ constructs without BRIL were expressed at a level of around 25% that of SMO_{XTAL}, however BRIL insertion conferred dramatically increased expression on FZD₃, producing a peak with greater intensity than SMO_{XTAL} (**Figure 3.4, b**)

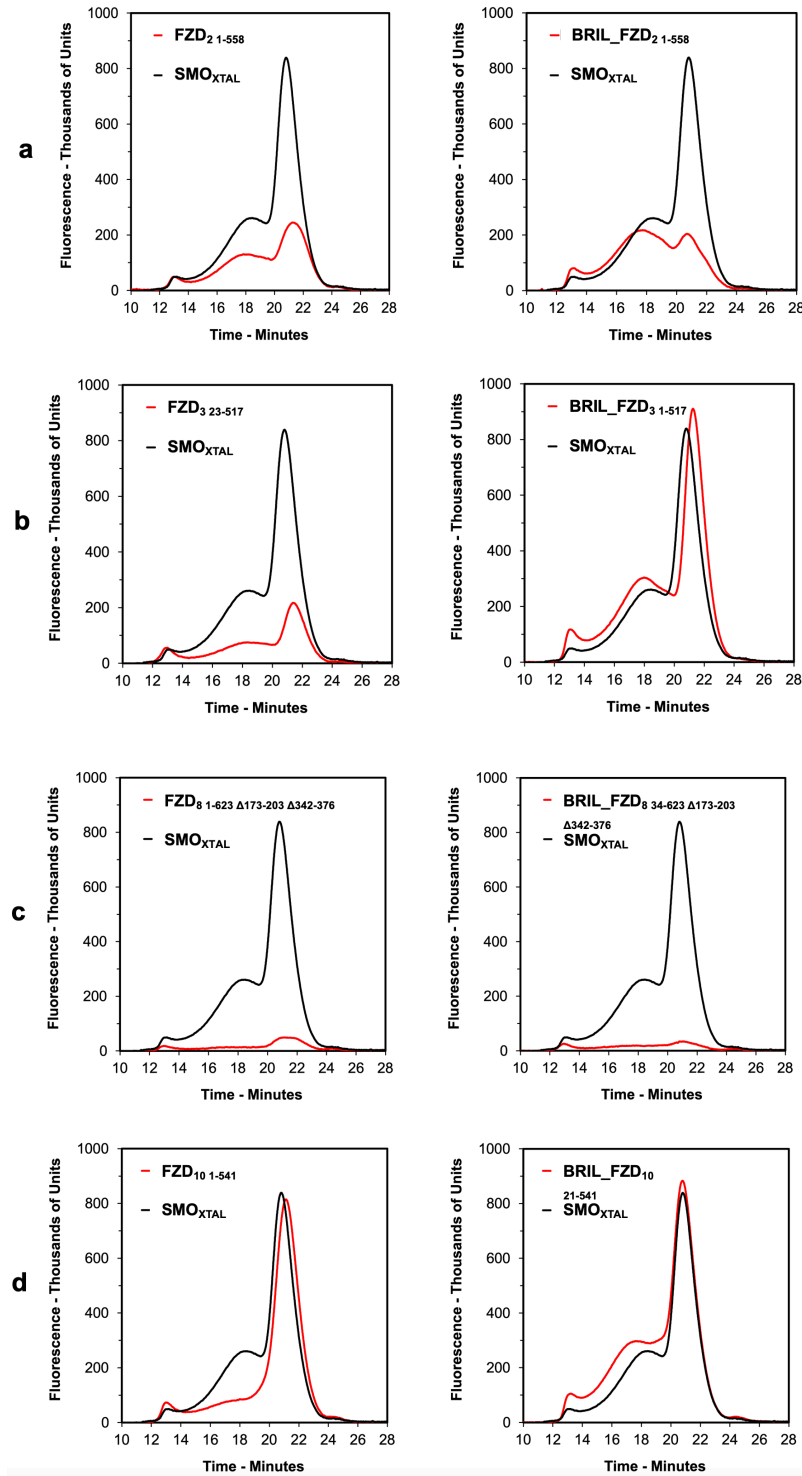


Figure 3.4 FSEC analysis of multidomain FZD_{2/3/8/10} constructs.

Representative FSEC profiles for multidomain FZD₂ (a), FZD₃ (b), FZD₈ (c) and FZD₁₀ (d) constructs without (left) and with (right) BRIL insertions compared to SMO_{XTAL}, which was expressed and purified in parallel.

FZD₈ constructs exhibited almost no expression, even with BRIL (**Figure 3.4, c**). It is possible that the extensive truncation of the long extracellular loops of FZD₈ interfered with its stability or ability to fold correctly.

FZD₁₀ constructs were expressed to a similar level to SMO_{XTAL}, both with and without BRIL fusions (**Figure 3.4, d**).

Overall, FZD₁ and ₃ constructs containing a BRIL fusion were generally better expressed than those without. The shape of the elution profiles observed for FZDs were similar to that observed for SMO_{XTAL}. FZD₂ and ₈ constructs were poorly-expressed regardless of BRIL fusion, whereas FZD₁₀ constructs with and without BRIL were well-expressed. Replacing the native secretion signal with the pHLsec secretion signal had little impact on expression levels. Therefore, constructs using the pHLsec secretion signal were favoured, since in these constructs residues between the end of the native signal peptide and the actual start of the CRD, which could be a source of flexibility that would interfere with crystallisation, are absent.

Constructs that compared favourably with SMO_{XTAL} in terms of expression level and elution profile were observed for FZDs ₃ and ₁₀, whereas FZDs ₁, ₂ and ₈ did not compare favourably. FZD₃ and ₁₀ were therefore identified as candidates for large-scale expression based on their favourable comparison to SMO in terms of expression level and elution profile.

3.4 Large-scale multidomain FZD production and optimisation

3.4.1 Initial large-scale FZD3 expression and crystallisation trials

BRIL_FZD₃ 23-517 (herein referred to as BRIL_FZD₃) was cloned into the pHLsec vector in-frame with a C-terminal Rho 1D4 tag and expressed transiently in HEK293S GnTI- cells grown in suspension. Culture volumes were typically around 9 litres, which would yield a cell pellet of around 80 g. From this amount of starting material, typical yields were approximately 1 mg of BRIL_FZD₃ (**Figure 3.5, a**). While this was a suitable yield for LCP crystallisation trials,

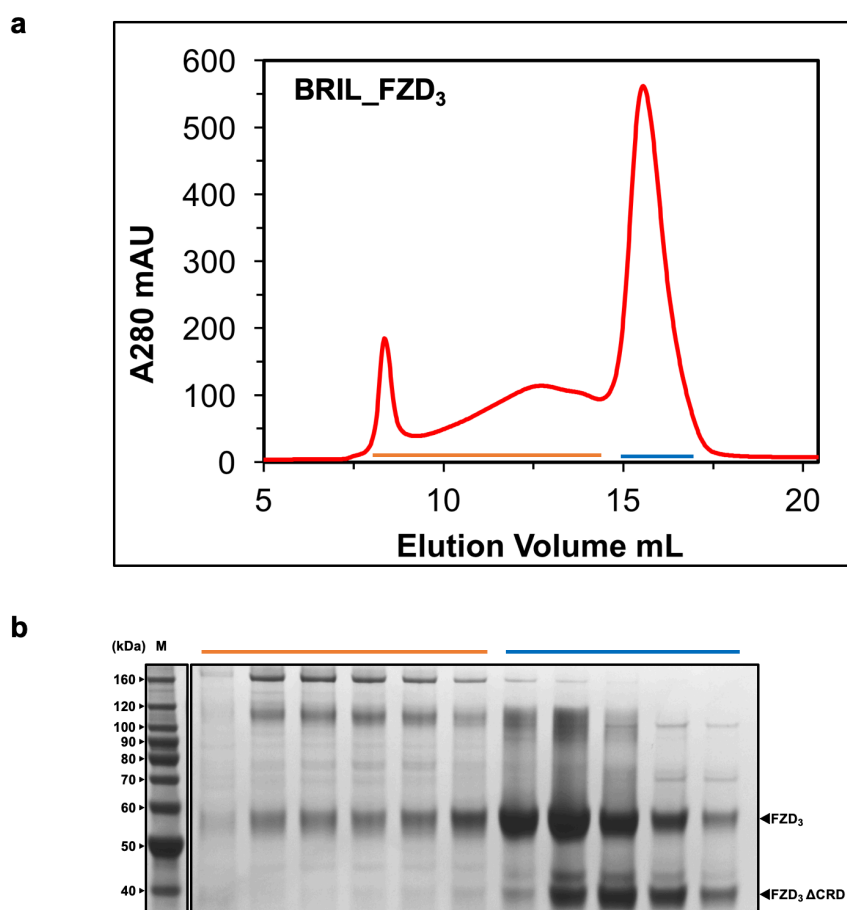


Figure 3.5 Purification of multidomain FZD for crystallisation experiments.

Typical SEC profile (a) and SDS-PAGE (b) of BRIL_FZD₃ 26-517 expressed transiently and purified via a C-terminal 1D4 tag. Cleavage of FZD₃ CRD leads to a band at apparent molecular weight of ~40 kDa. Blue bars represent fractions pooled for crystallisation.

the large culture volumes meant that transfection was a time- and labour-intensive process and large cell pellets were cumbersome to process. A more efficient expression strategy was therefore sought. Furthermore, SDS-PAGE analysis of purified samples showed the presence of a lower molecular weight species (**Figure 3.5, b**). This was also undesirable as a highly homogeneous sample was sought for crystallisation.

Despite these issues, LCP crystallisation was attempted with this construct, at concentrations between 20 and 40 mg mL⁻¹, however while small crystals (~10 μm in length) were observed,

these did not diffract x-rays. Optimisation of the crystallisation conditions did not yield an improvement in crystal size or diffraction quality.

3.5 Further optimisation of multidomain FZD production

3.5.1 Lentivirus-transduced cell lines for the expression of FZD₃

Due to the expensive and cumbersome nature of large-scale transient transfections, a more efficient expression strategy was sought for multidomain FZD constructs. Lentivirus transduction, a technique recently adopted in this laboratory, was therefore explored as a more efficient and lower-cost strategy for the generation of stable cell lines expressing FZD constructs.

In contrast to transient transfection, whereby cultures are grown to a desired volume prior to introduction of the gene of interest, stable cell lines can be produced at a smaller scale and then grown to a desired culture volume. Lentiviral transduction is one such method for generating stable cell lines (Elegheert *et al.*, 2018). This technique involves an initial transfection step, in which plasmids encoding the gene of interest, as well as essential viral proteins, are introduced to virus producer cells. The resulting virus particles are then capable of infecting target cells with extremely high efficiency.

In contrast to transient transfection whereby the transfected plasmid remains episomal (i.e. Extrachromosomal) and is rapidly degraded and diluted by cell division (Subramanian and Srienc, 1996), lentiviral transduction leads to incorporation of the target gene into the host cell genome. This has multiple advantages. Firstly, once the gene is stably incorporated into the host genome, it is maintained after cell division. This means that a relatively small number of transduced cells can be grown to a desired culture volume and cell density. As such, the need for cumbersome large-scale transfections is eliminated. Secondly, lentiviral transduction leads to a greater proportion of cells receiving the target gene compared to transient transfection

(Elegheert *et al.*, 2018). In this way, culture volumes can be reduced as there are fewer cells lacking the target gene, which are essentially useless.

For these reasons, expression of multidomain FZD constructs was switched to lentivirus-transduced cell lines. Indeed, combined with other strategies discussed below, lentivirus-transduced cell lines allowed the purification of milligram quantities of FZD₃ from as little as 3 g of cell pellet, or approximately 400 mL of culture (**Figure 3.6, a and b**). This substantial improvement in efficiency increased the overall amount of FZD₃ that could be purified and eliminated many of the difficulties associated with processing large cell pellet masses.

3.5.2 Overcoming CRD cleavage and improving yield with use of N-terminal Twin-Strep tag

SDS-PAGE analysis of purified multidomain FZD₃ indicated the presence of a major contaminating lower molecular weight species, which migrated to an indicated molecular weight of approximately 39 kDa (**Figure 3.5, b**). The presence of this band indicated that the FZD₃ sample was not very homogeneous, which may have hindered its propensity to crystallise. The lower molecular weight species was predicted to correspond to FZD₃ that had undergone cleavage of the CRD, based on the observation that it migrated to an indicated molecular weight approximately 18 kDa smaller than the indicated ~58 kDa of intact BRIL_FZD₃. The fact that the degradation product was still present after 1D4 affinity chromatography indicated that its C-terminus remained intact and N-terminal degradation was likely to be the cause. The highly-specific anti-1D4 antibody purification method also meant there was a low risk of contaminating proteins.

Cleavage of the CRD was confirmed as the source of the lower molecular weight species by N-terminal sequencing (NTS, performed by Cambridge Peptides). The cleavage site of the major degradation product observed was determined to be immediately after A156 within the

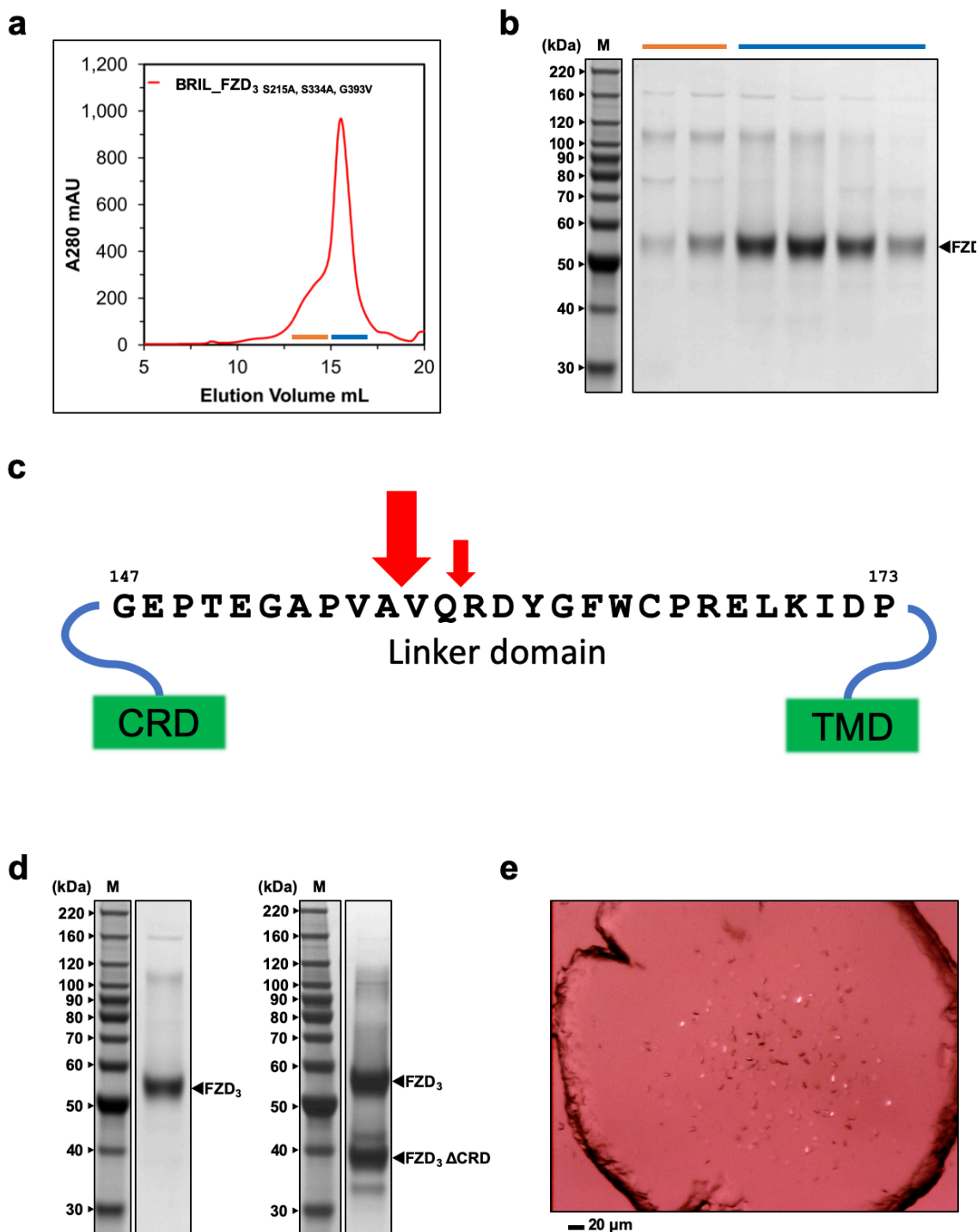


Figure 3.6 N-terminal tag, in combination with lentiviral expression, facilitates purifications of large quantities of highly pure FZD₃.

a, b, SEC (**a**) and SDS-PAGE (**b**) analysis of BRIL_FZD₃ S215A, S334A, G393V. **c**, Cleavage sites in the FZD₃ LD identified by NTS. **d**, Comparison of N-terminally tagged FZD₃ (left) and C-terminally tagged FZD₃ (right). Purification via an N-terminal tag overcomes the issue of CRD cleavage that increases the heterogeneity when a C-terminal tag is used. **e**, Crystals of BRIL_FZD₃ S215A, S334A, G393V grown in LCP.

linker domain between the CRD and TMD (**Figure 3.5, c**). This region is predicted to be unstructured (**Appendix 2**) and is therefore likely to be flexible and accessible to extracellular proteases. As the same phenomenon of CRD cleavage was also observed for FZD₄ and FZD₆ by colleagues (data not shown), and appeared to be affecting FZD₁ based on FSEC elution profiles (see **3.3.1**), it is predicted that it will affect most or all FZDs. This is despite the fact that the linker sequence is very poorly conserved between FZDs. These observations suggest that mutation of this sequence would be unlikely to overcome the sample heterogeneity introduced by CRD cleavage. A second, minor degradation product was also detected by NTS, immediately after Q158 of FZD₃. The presence of multiple cleavage sites and the fact that CRD cleavage was observed for other FZDs, despite the lack of conservation in the linker domain, suggests that this region is highly prone to proteolytic cleavage.

A strategy for addressing this problem was the use of an N-terminal affinity tag. In this approach, only FZDs maintaining an intact N-terminus, including the CRD, would be purified, while proteins that had undergone CRD cleavage would not be, as this would result in the loss of the affinity tag as well. However, the Rho 1D4 epitope tag is only recognised with high affinity by the anti-Rho 1D4 antibody when present at the C-terminus of proteins, due to the requirement of a free carboxylate group (Hodgess *et al.*, 1988). As such, an alternative tag was required that remained functional when present at the N-terminus.

The Twin-Strep tag was chosen based on its high affinity, high specificity, high binding capacity and its location-independent function (Schmidt *et al.*, 2013). When purified with an N-terminal Twin-Strep tag, SDS-PAGE analysis showed the sample to be highly homogeneous with no lower molecular weight band (**Figure 3.6, d**), indicating that this strategy had successfully overcome the issue of CRD cleavage.

3.6 Thermostabilising point mutations

3.6.1 Design of point mutations

Another common technique employed to aid structural studies of GPCRs is to introduce point mutations into the TMD, which increase the expression and thermostability of the protein (Serrano-vega *et al.*, 2008). An example of this is the V329^{3.40f}F (superscript numbers in X.YYf format denote Ballesteros-Weinstein numbering of Class F GPCRs; Ballesteros and Weinstein, 1995; Wang *et al.*, 2014; see section 1.6.1 and **Appendix 1**) mutation used for the first published multidomain crystal structure of the SMO receptor (Byrne *et al.*, 2016). In this case, the introduction of a bulkier residue into the bottom of the ligand-binding pocket in the TMD locks the receptor in an inactive conformation, thereby reducing the conformational heterogeneity of the sample (Byrne *et al.*, 2016).

Interestingly, sequence alignments show that the equivalent residue is serine in all FZDs other than FZD₆, in which it is a threonine (**Appendix 1**). Nevertheless, mutation of this residue to phenylalanine (S301F) was investigated as a strategy to improve the stability of FZD₃.

A commonly used strategy to identify thermostabilising point mutations in GPCRs is alanine scanning, whereby each residue in turn is mutated to alanine (and alanine is mutated to leucine) because alanine and leucine occur frequently in alpha helices (Pace and Scholtz, 1998). However, this intensive approach is very time-consuming. As such, a targeted approach was sought to try and identify stabilising point mutations with greater efficiency.

Mutation of non-conserved polar residues in the TMD of GPCRs has previously been shown to be a viable strategy for discovery of thermostabilising point mutations in GPCRs (Chen *et al.*, 2012). A multiple sequence alignment of the 11 Class F GPCR sequences was performed to assess the conservation of polar TMD residues (**Appendix 1**). Three polar amino acids were identified in the FZD₃ TMD that were not conserved in any other Class F GPCR: S215^{1.45f},

S334^{4.48f} and T341^{4.55f}. These amino acids were separately mutated to alanine (S215^{1.45f}A, S334^{4.48f}A and T341^{4.55f}A) in BRIL_FZD₃.

A further mutation candidate was selected based on its implication in disease. A *de novo* mutation in FZD₂, G434^{5.64}V, was previously identified in a patient with autosomal dominant omodysplasia (OMOD2), a rare disease characterised by abnormal skeletal development (Türkmen *et al.*, 2017). Interestingly, this mutation appears to phenocopy a nonsense mutation in FZD₂, W548* (Türkmen *et al.*, 2017), which leads to truncation of the KTXXXW motif required for DVL binding and activation of ‘canonical’ WNT signalling (Umbhauer *et al.*, 2000), suggesting that the mechanism in both cases is loss of FZD₂ function. Interestingly, this glycine residue is perfectly conserved in all Class F GPCRs, including SMO (**Appendix 1**), suggesting that it is indispensable for the structure or function of these proteins.

The authors that identified the FZD₂ G434^{5.64}V disease-causing mutation concluded that introduction of a bulky amino acid at this position likely abrogates the ability of FZD₂ to fold (Türkmen *et al.*, 2017). However, it is also plausible that this mutation inactivates FZD by locking it in an inactive conformation in a similar manner to the V329^{3.40f}F mutation in SMO (Byrne *et al.*, 2016), a theory that warranted further investigation to determine if this mutation would be useful for stabilising FZD. The equivalent mutation, G393^{5.64f}V, was therefore introduced into FZD₃.

3.6.2 Point mutation expression screens

Interestingly, FSEC analysis showed that the equivalent mutation to SMO V329^{3.40f}F, FZD₃ S301^{3.40f}F, severely impacted the expression level of FZD₃ (**Figure 3.7, a**), with almost no monomer peak detected, suggesting it had a severe destabilising effect on the receptor. The expression level of each non-conserved polar residue mutant was analysed by FSEC (**Figure 3.7, b-d**). While this method does not directly measure thermostability, the V329^{3.40f}F mutation

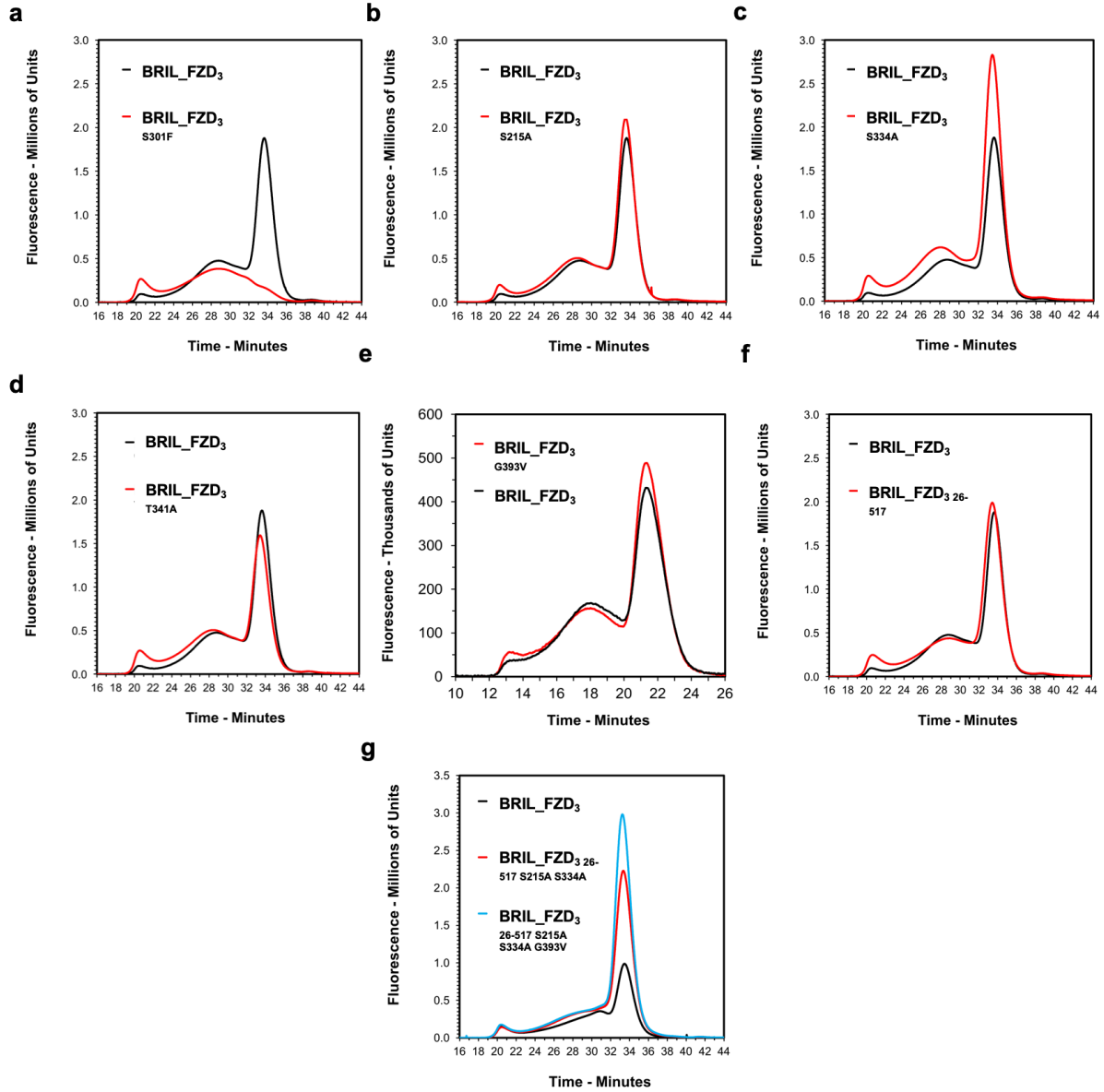


Figure 3.7 FZD₃ TMD point mutation expression screening

a-d, FSEC analysis of the BRIL_FZD₃ TMD point mutations **(a)** S301^{3.40f}F; **(b)** S215^{1.45f}A; **(c)** S334^{4.48f}A; and **(d)** G393^{5.64f}V. **e**, FSEC analysis of BRIL_FZD₃ 26-517. **f**, FSEC analysis of the mutation combination constructs BRIL_FZD₃ 26-517 S215A, S334A and BRIL_FZD₃ 26-517 S215A, S334A, G393V. In all cases, BRIL_FZD₃ (with no point mutations) was expressed and purified in parallel to facilitate comparison.

that confers heightened thermostability on SMO was also observed to boost its expression (Byrne *et al.*, 2016).

The S215^{1.45f}A and S334^{4.48}A mutations were both found to boost expression of FZD₃. S215^{1.54f}A yielded a modest increase in expression, with a peak fluorescence intensity approximately 10% greater than FZD₃ with wild type TMD helices. On the other hand, S334^{4.48f}A lead to an increase in peak intensity of approximately 50%. T341^{5.55}A caused a small decrease in yield, indicating decreased stability.

A construct combining S215^{1.45f}A and S334^{4.48}A was then generated (BRIL_FZD₃ S215A, S334A) to see if combination of these two mutations could yield a further-stabilised protein. When combined, the expression level as measured by FSEC was more than double that of FZD₃ with wild type TMD helices (**Figure 3.7, g**), indicating a synergistic effect.

The equivalent mutation to the disease-causing FZD₂ G434^{5.64f}V mutation, G393^{5.64f}V also produced a very modest increase in expression (**Figure 3.7, e**). However, when introduced into (Byrne *et al.*, 2016), meaning it can cautiously be considered to be an indirect measure of thermostability.

The S215^{1.45f}A and S334^{4.48f}A mutations were both found to boost expression of FZD₃. S215^{1.54f}A yielded a modest increase in expression, with a peak fluorescence intensity approximately 10% greater than FZD₃ with wild type TMD helices. On the other hand, S334^{4.48f}A lead to an increase in peak intensity of approximately 50%. T341^{5.55}A caused a small decrease in yield, indicating decreased stability.

A construct combining S215^{1.45f}A and S334^{4.48f}A was then generated (BRIL_FZD₃ S215A, S334A) to see if combination of these two mutations could yield a further-stabilised protein. When BRIL_FZD₃ S215A, S334A (BRIL_FZD₃ S215A, S334A, G393V) the triple mutant was expressed at a level approximately 35% higher than the double mutant (**Figure 3.7, g**), again suggesting a synergistic effect of these mutations on FZD₃ expression levels. This supported the hypothesis

that the mutation FZD₂ G434^{5.64f}V may cause disease by stabilising the receptor in an inactive conformation.

Truncation of the first 2 residues of the FZD₃ CRD (BRIL_FZD₃ 26-517) was also found to modestly increase the expression of the receptor (**Figure 3.7, f**).

3.6.3 FZD₃ TMD point mutations increase receptor thermostability

In order to confirm that the improved expression of the constructs with point mutations corresponded to increased thermostability, samples of BRIL_FZD₃, BRIL_FZD₃ S215A, S334A and BRIL_FZD₃ S215A, S334A, G393V were incubated at temperatures between 4°C and 80°C for one hour and then analysed by SEC (**Figure 3.8, a**). The relative maximum peak intensities (i.e. the greatest A₂₈₀ observed for each sample, normalised to the same protein incubated at 4°C, **Figure 3.8, b**) and elution times of samples (i.e. the time at which the maximum peak intensity was observed, **Figure 3.8, c**) incubated at different temperatures were monitored to assess stability. A decrease in maximum peak intensity indicated degradation of the sample, whereas earlier elution of the peak indicated formation of higher molecular weight aggregates. For each sample, the relative maximum peak intensity and peak elution time was plotted against incubation temperature.

The maximum peak intensities and elution times for both BRIL_FZD₃ S215A, S334A and BRIL_FZD₃ S215A, S334A, G393V exhibited greater resistance to higher incubation temperature than BRIL_FZD₃. Trendlines were fitted to the plots of maximum peak intensity against temperature. The temperature at which the trendline dropped to 80% was used to compare each construct. For BRIL_FZD₃, the maximum peak intensity dropped to 80% at 40°C. On the other hand, the peak intensity of BRIL_FZD₃ S215A, S334A dropped to 80% at 44°C, while the intensity of BRIL_FZD₃ S215A, S334A, G393V did not drop to 80% until 58°C. Similarly, the peak retention

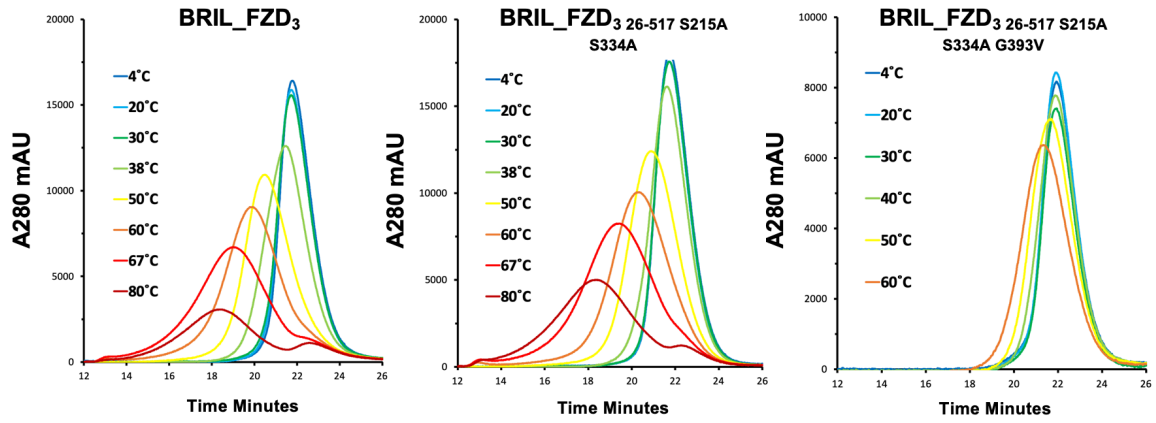
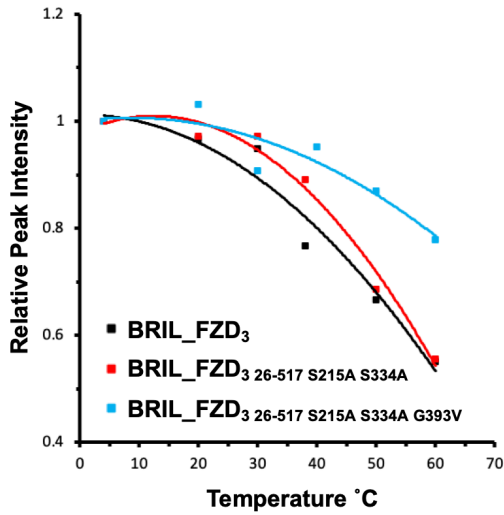
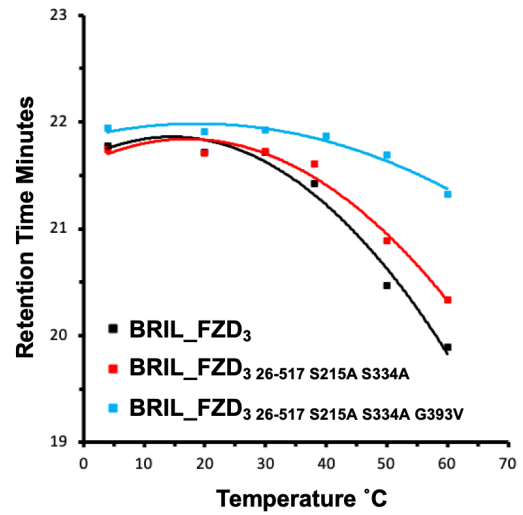
a**b****c**

Figure 3.8 TMD point mutations increase FZD₃ thermostability.

a, Samples of BRIL_FZD₃ (left), BRIL_FZD₃ 26-517 S215A S334A (middle) and BRIL_FZD₃ 26-517 S215A S334A G393V (right) were heated at temperatures between 4° C and 80° C for one hour and then analysed by SEC. **b** and **c**, The relative peak intensity (**b**) and peak retention time (**c**) are plotted against temperature for each construct. The peak intensity and retention time of BRIL_FZD₃ 26-517 S215A S334A G393V were more resistant to higher temperatures than BRIL_FZD₃ 26-517 S215A S334A, which in turn was more resistant than BRIL_FZD₃.

time of BRIL_FZD₃ S215A, S334A was less affected by higher incubation temperatures than BRIL_FZD₃, and that of BRIL_FZD₃ S215A, S334A, G393V exhibited even greater resistance.

Overall, these results show that BRIL_FZD₃ S215A, S334A demonstrates greater thermostability than BRIL_FZD₃, while BRIL_FZD₃ S215A, S334A, G393V exhibits greater thermostability still. This

confirms that the expression-boosting point mutations S215^{1.45f}A, S334^{4.48f}A and G393^{5.64f}V also increase the thermostability of FZD₃.

3.6.4 Theoretical basis for FZD₃ thermostabilisation by TMD point mutations

In order to investigate the mechanism by which the identified point mutations increased the thermostability of FZD₃, homology models were created using the FZD₄ TMD crystal structure (PDB 6BD4) as a template (**Figure 3.9**). The first homology model was produced using the wild-type (WT) FZD₃ TMD sequence, while the second contains the G393^{5.64f}V mutation.

When looking at the WT FZD₃ TMD homology model, it becomes apparent that both S215^{1.45f} and S334^{4.48f} are located on the outer surface of the TMD, on TM1 and TM4, respectively (**Figure 3.9, a and b**). Therefore, these polar serines would be exposed to the hydrophobic tails of the plasma membrane phospholipids, which would be energetically unfavourable, although permissible within the context of a large hydrophobic domain. Hence, when mutated to hydrophobic alanine, these residues are likely to be more energetically favourable, thereby increasing the stability of the receptor.

As for G393^{5.64f}, this residue is located buried within the 7TM bundle, towards the cytosolic end of TM5 (**Figure 3.9, c and d**). In the G393^{5.64f}V homology model, it can be seen that the hydrophobic valine side chain is located in close proximity to the side chains of W311^{3.50f} and A315^{3.55f} on TM3 and is likely to form hydrophobic contacts with these residues. Therefore, the G393^{5.64f}V mutation may increase the stability of the interaction between helices V and III by increasing the extent of hydrophobic contacts between them. The valine side chain also appears to be able to form contacts to the side chain of L397^{5.68f}, also located on TM5, which may increase the stability of this helix.

Interestingly, W311^{3.50f} and L397^{5.68f} are both perfectly conserved in all Class F GPCRs, while A315^{3.55f} is perfectly conserved in all FZDs (but not SMO, in which it is phenylalanine). Therefore, it is possible that the equivalent mutation to G393^{5.64f}V would have a similar

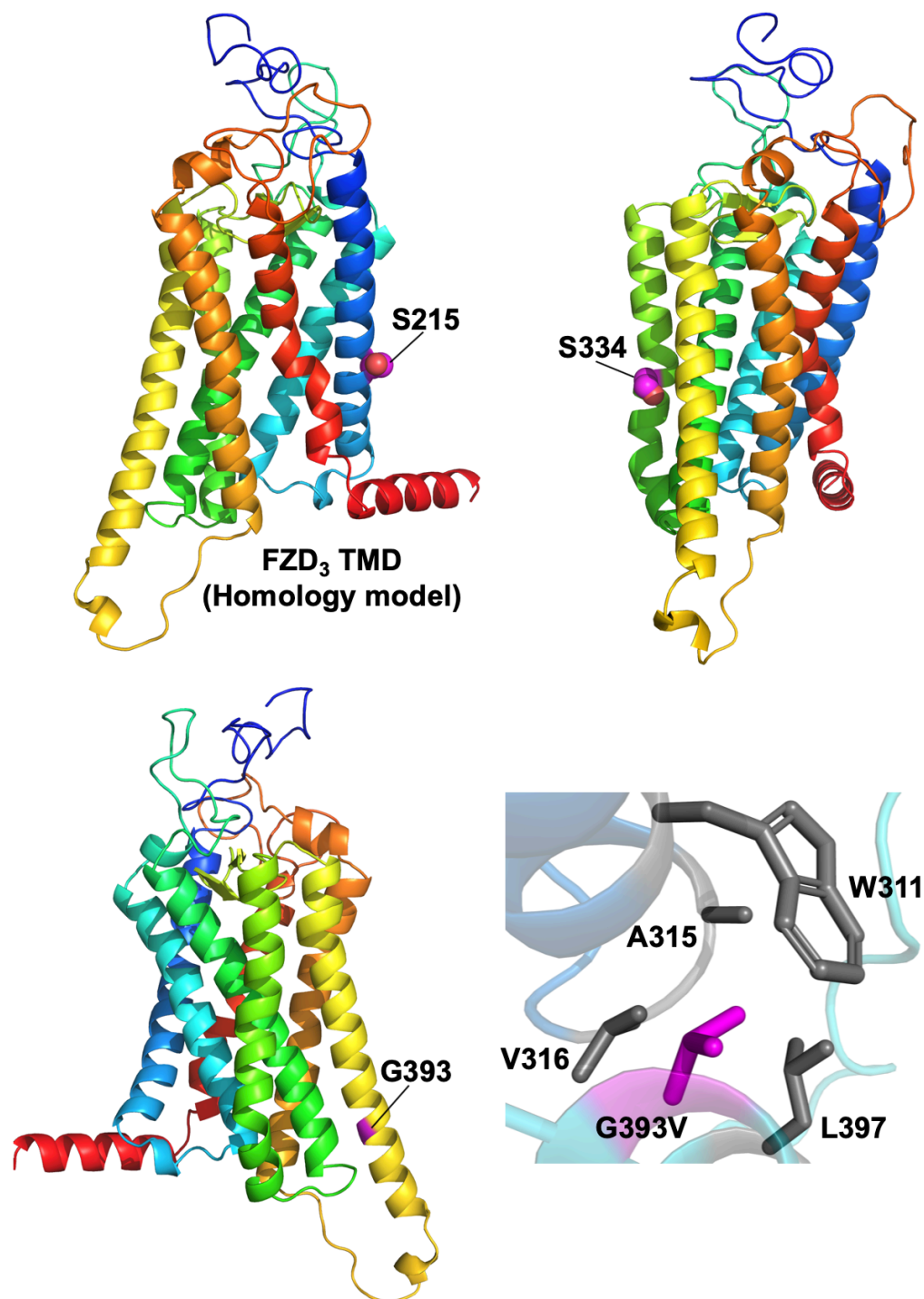


Figure 3.9 Potential basis for the stabilising effects of FZD₃ TMD point mutations.

Homology models of the FZD₃ TMD were generated using the crystal structure of the FZD₄ TMD (PDB 6BD4) as a template (cartoon and surface representation). **a**, **b**, The positions of residues S215^{1.45f} (**a**) and S334^{4.48f} (**b**) within the FZD₃ TMD are shown as magenta spheres. **c**, The position of G393^{5.64f} within the FZD₃ TMD is displayed in magenta. **d**, Detail of G393^{5.64f}V (magenta sticks) and residues that it is predicted to form hydrophobic contacts with (grey sticks).

stabilising effect on other FZDs as these hydrophobic contacts would be conserved. When considering the mechanism by which this mutation could inactivate FZD, it seems likely that this increased stabilisation could reduce the ability of the 7TM to undergo conformational changes required to facilitate binding of either DVL, G proteins, or other downstream effectors.

3.7 Large scale purification and crystallisation trials of optimised multidomain FZD₃ construct

Crystallisation trials were performed for the optimised multidomain FZD₃ construct (BRIL_FZD₃ 26-517, S215A, S334A, G393V) using the lipidic cubic phase method. Small birefringent crystals were observed (**Figure 3.5, e**), however these still did not diffract x-rays.

3.8 Conclusions and future perspectives

While efforts to determine a multidomain FZD structure have proved unsuccessful to date, the progress made in optimisation of expression constructs, expression strategy and purification strategy, provides a strong platform on which future efforts can build. In particular, the stability of the FZD₃ TMD has been greatly improved by the introduction of the BRIL fusion protein and thermostabilising point mutations. Consequently, strategies to stabilise the conformation of the FZD LD, for example by the use of antibodies, will hopefully provide the final ingredient to allow the determination of a multidomain FZD structure. Alternatively, the reported optimisations of the thermostability of the FZD₃ TMD could facilitate the crystallisation of the FZD₃ TMD alone. Although less informative than a multidomain structure, a FZD₃ TMD structure would allow useful comparisons with the FZD₄ TMD structure (Yang *et al.*, 2018), which remains the only example of a FZD TMD structure.

Chapter 4 Nanobodies against FZD₃

4.1 Introduction

Due to the unsuccessful attempts to produce multidomain FZD₃ crystals that diffracted x-rays, which were compounded by the lack of known stabilising small molecule ligands or natural ligands that can be produced in suitable quantities, anti-FZD₃ antibodies were considered a potential way forward for the project. Of particular interest were recombinant, antigen-specific, single-domain antibodies, derived from heavy chain-only antibodies found in camelids (Muyldermans, 2013). These so-called ‘nanobodies’ (Nbs) have numerous advantages that make them ideally suited to structural biology, including their small size, ease of expression, typically high binding affinities, excellent stability and ability to stabilise specific protein conformations (Muyldermans, 2013).

Nbs against FZD₃ were produced by the Steyaert laboratory in the Vlaams Instituut voor Biotechnologie at Vrije Universiteit Brussel. In order to generate FZD₃-binding Nbs, samples of FZD₃ 26-517 S215A S334A G393V were used for llama immunisation. These samples were produced without a BRIL fusion so that Nbs against the wild-type intracellular loops could be identified. For selections, the same construct was provided, along with a sample in which the CRD was replaced with BRIL. This sample could be used both to positively select Nbs against the FZD₃ TMD and negatively select Nbs against the FZD₃ CRD, however at the time of writing only Nbs selected with the original immunisation construct have been received.

4.2 Aims

The initial aim of the work described in this chapter was to identify Nbs that would stabilise the conformation of FZD₃ to produce a more suitable sample for crystallisation trials. However, due to time constraints upon receiving the Nbs, the aims were adjusted. As such, the primary aim became to identify a Nb against the FZD₃ CRD, which could be used to determine a FZD₃

CRD structure and potentially be useful for targeting FZD₃ to modulate WNT signalling. A further aim was to characterise such a Nb by quantifying the affinity of its interaction with FZD₃ and confirming the FZD binding specificity of the Nb. Finally, the Nb would be used in crystallisation trials with multidomain FZD₃.

4.3 Identification and characterisation of an anti-FZD3 CRD Nb

4.3.1 Nb14478 binds to the FZD₃ CRD

As the Nbs received had been selected only with multidomain FZD₃ protein, no prior knowledge about their epitopes was available. However, analysis of the Nb sequences showed that six of the ten Nbs belonged to a single family, while the other four were unique (**Appendix 3**). Nbs are considered to belong to a single family if they share an identical complementarity-determining region (CDR) 3 loop length and sequence identity of >80% within the CDR3 loop (Pardon *et al.*, 2014).

In order to identify Nbs that bind to the FZD₃ CRD, small quantities of Nbs and FZD₃ CRD protein were mixed and analysed by size-exclusion chromatography (SEC). When analysed separately using a Superdex 75 10/300 column (GE Healthcare), FZD₃ CRD elutes at a volume of approximately 12.6 mL, while Nbs were typically eluted at approximately 13.6 mL. Therefore, the appearance of a species eluting at a lower volume would indicate the formation of a larger species, i.e. a CRD:Nb complex. For Nbs that did not bind to the FZD₃ CRD, two separate peaks would be expected at 12.6 mL and 13.6 mL elution volume (**Figure 4.1, a**).

When Nb₁₄₄₇₈ was mixed with FZD₃ CRD, an earlier-eluting peak was observed at an elution volume of approximately 11.3 mL (**Figure 4.1, a**), indicating formation of a larger species. SDS-PAGE confirmed the presence of both FZD₃ CRD and Nb₁₄₄₇₈ in this peak (**Figure 4.1, c**) All other Nbs failed to produce a lower elution volume peak, indicating no strong interaction with FZD₃ CRD.

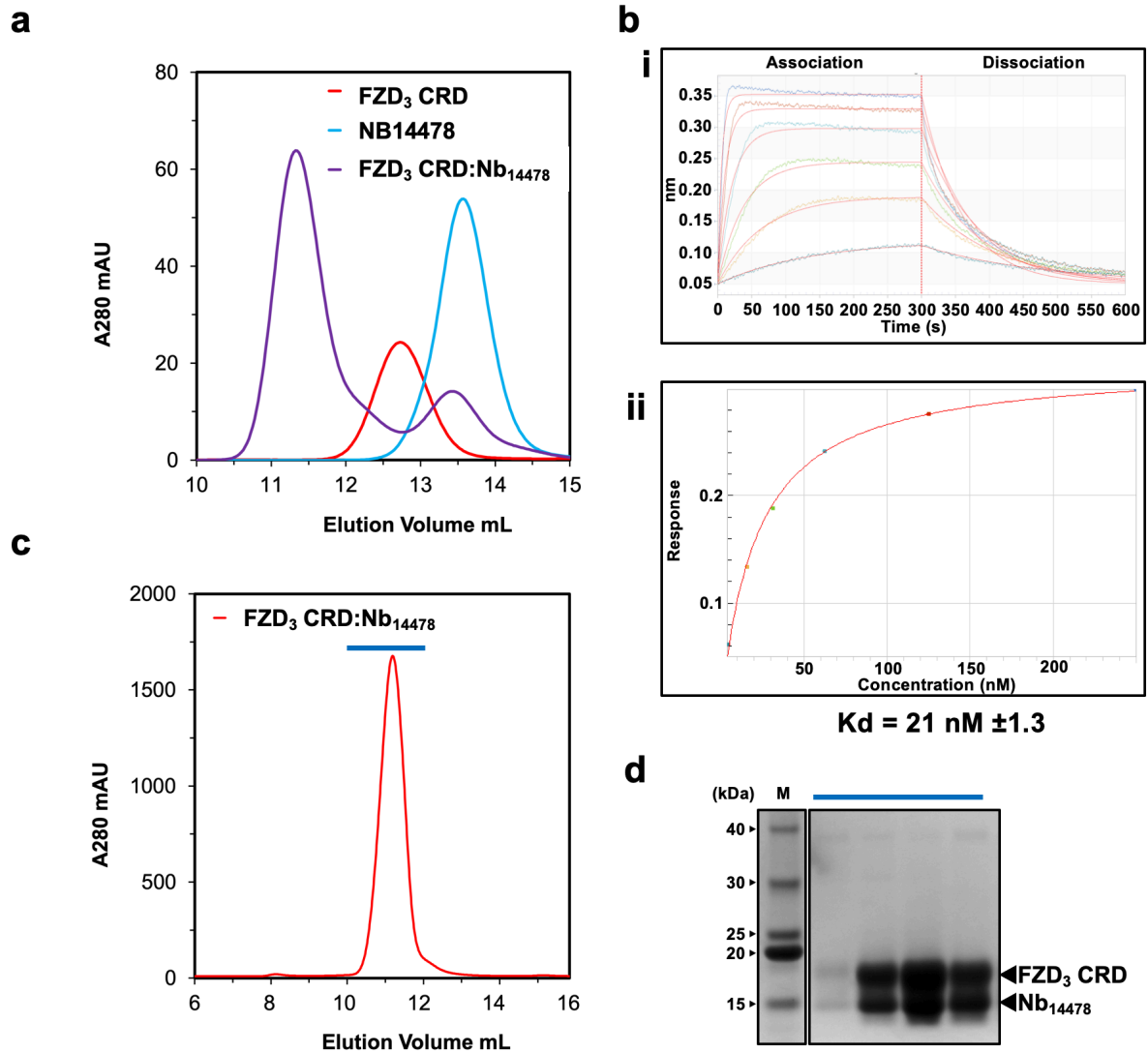


Figure 4.1 Nb₁₄₄₇₈ binds to FZD₃ CRD with high affinity

a, SEC analysis indicates complex formation upon mixing FZD₃ CRD with Nb₁₄₄₇₈. **b, i**, Binding of Nb₁₄₄₇₈ to FZD₃ CRD. The data was fit (red) using a 1:1 model. **ii**, Response curve for binding of Nb₁₄₄₇₈ to FZD₃. **c, d**, SEC (**c**) and SDS-PAGE (**d**) analysis of copurified FZD₃ CRD:Nb₁₄₄₇₈ complex. Blue bars represent fractions pooled and concentrated for crystallisation experiments.

4.3.2 Affinity of the FZD₃:Nb14478 interaction

Bio-layer interferometry (BLI) was performed to quantify the affinity of the interaction between Nb₁₄₄₇₈ and FZD₃ CRD (**Figure 4.1, b**). This showed that the Nb bound with a dissociation constant (K_d) of around 21 nM. This low-nanomolar affinity is in agreement with the observation that, when co-expressed and co-purified, FZD₃ CRD and Nb₁₄₄₇₈ form a highly

stable complex on gel filtration, with no evidence of unbound components (**Figure 4.1, c**), indicating that all of the FZD₃ proteins purified were bound to Nb₁₄₄₇₈. The high stability of the interaction meant that FZD₃ CRD could be coexpressed with untagged Nb₁₄₄₇₈ using lentiviral coinfection. The complex could then be copurified from conditioned medium, thereby reducing the number of flexible tag residues present in the sample, a potential advantage for crystallisation attempts. This allowed mg quantities of a highly pure sample of the FZD₃:Nb₁₄₄₇₈ complex to be purified for crystallisation trials (**Figure 4.1, d**).

4.3.3 Nb14478 binds with high specificity to FZD_{3/6}

BLI was also used to investigate the cross-reactivity of Nb₁₄₄₇₈ for all ten FZD CRDs. At a concentration of 5 µM, Nb₁₄₄₇₈ was only observed to bind to FZD₃ and FZD₆ CRDs (**Figure 4.2, a**), indicating a strong preference for binding to this particular subfamily of FZDs. As this concentration was 250 times higher than the experimentally determined K_d for the interaction between FZD₃ CRD and Nb₁₄₄₇₈, it is clear that Nb₁₄₄₇₈ is highly selective for the FZD_{3/6} subfamily, making it a useful tool for investigating these two receptors. This is in stark contrast to another anti-FZD antibody that binds to six receptors across every FZD subfamily excluding FZD_{3/6} (Pavlovic *et al.*, 2018). Expression and successful *in vivo* biotinylation of FZD proteins was confirmed by SDS-PAGE (**Figure 4.2, b**) and streptavidin blot (**Figure 4.2, c**), respectively.

4.4 Crystal structure of FZD₃ CRD:Nb₁₄₄₇₈ complex

4.4.1 Overview of the FZD₃ CRD:Nb₁₄₄₇₈ crystal structure

The crystal structure of the FZD₃ CRD:Nb₁₄₄₇₈ complex was determined at 1.8 Å resolution (**Figure 4.3, a; Figure 4.4**), in order to determine the epitope and to investigate the basis of its specificity for FZD_{3/6} (data collection and refinement statistics are reported in **Table 2**). The structure reveals that Nb₁₄₄₇₈ binds to an epitope proximal to the ‘bottom’ of the hydrophobic

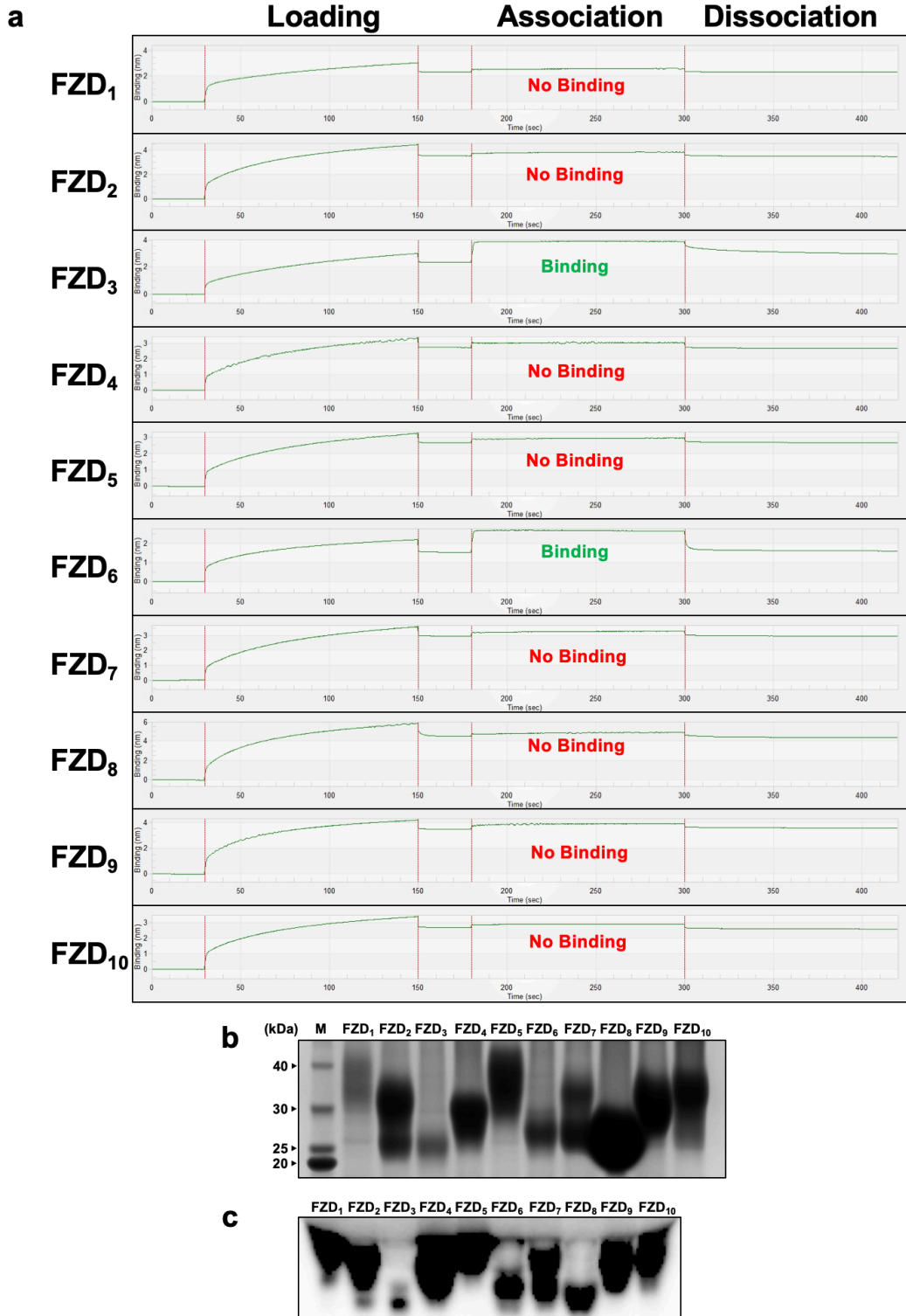


Figure 4.2 Nb₁₄₄₇₈ specifically binds to the CRDs of the FZD_{3/6} subfamily

a, Bio-layer interferometry (BLI) data for all 10 human FZD CRDs showing binding to Nb₁₄₄₇₈ at 5 μ M concentration. **b** and **c**, SDS-PAGE (**b**) and Streptavidin-HRP blot (**c**) of FZD₁₋₁₀ CRD samples to confirm expression and successful *in vivo* biotinylation, respectively. Differing apparent molecular weights are caused by variable length of CRD, specific construct boundaries and number of N-linked glycosylation sites.

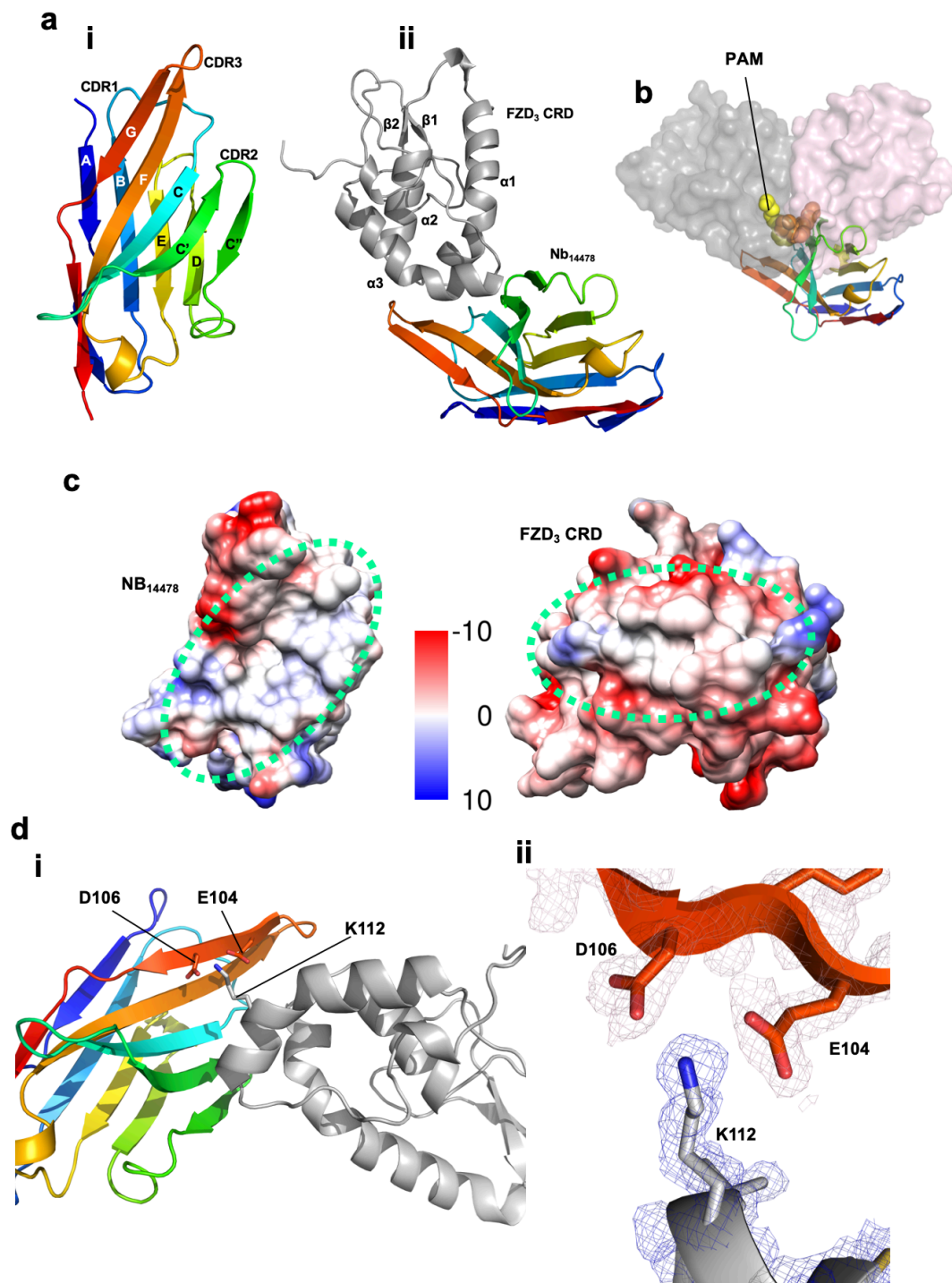


Figure 4.3 Crystal structure of Nb₁₄₄₇₈ in complex with FZD₃ CRD

a, Crystal structure of Nb₁₄₄₇₈ (rainbow, **i**) bound to FZD₃ CRD (grey, **ii**). **b**, Structure of dimeric WNT3-bound mFZD8 (PDB 6AHY, grey and pink) showing PAM (yellow and

orange spheres) bridging the dimer interface. NB₁₄₄₇₈ is superimposed at the equivalent binding site in the pink monomer, which overlaps with the dimerisation interface. **c**, Interaction surfaces (denoted by green dashed ovals) on NB₁₄₄₇₈ (left) and FZD₃ CRD (right). Positively charged, neutral and negatively charged surfaces are coloured blue, white and red, respectively. Units are kcal/(mol·e). **d, i**, Crystal structure of NB₁₄₄₇₈ bound to FZD₃ CRD with electrostatic interactions between NB₁₄₄₇₈ E105/D107 and FZD₃ CRD K112 (side chains displayed as sticks) indicated. **ii**, Electron density map (mesh, m2Fo-DFc, contoured at 1.0σ; Read, 1986) and model of electrostatic interaction (sticks).

Table 2 Crystallographic data collection and refinement statistics for FZD₃ CRD:Nb₁₄₄₇₈

FZD₃ CRD:Nb₁₄₄₇₈

Data collection	
Space group	<i>C</i> 222 ₁
Unit cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	54.0 146.4 229.5
<i>α</i> , <i>β</i> , <i>γ</i> (°)	90.0 90.0 90.0
Wavelength	0.98 Å
Resolution (Å)	114.76 - 1.77 (1.99 - 1.77)
<i>R</i> _{merge}	0.078 (>1)
<i>R</i> _{pim}	0.022 (0.387)
CC(1/2)	0.999 (0.797)
Mean ((<i>I</i>)/σ(<i>I</i>))	18.0 (1.8)
Completeness	93.9 (81.3)
Redundancy	13.8 (14.6)
Refinement	
Resolution (Å)	114.76 - 1.77
Total reflections	663243 (35278)
Unique reflections	48208 (2410)
<i>R</i> _{work}	0.1899
<i>R</i> _{free}	0.2226
Number of atoms	4270
Protein	3795
Ligand/ion	206
Water	269
B factors (Å²)	
Protein	44.51
Ligand/ion	101.31
Water	44.99
RMS deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.29
Ramachandran favoured	97.06
Ramachandran outliers	0.42

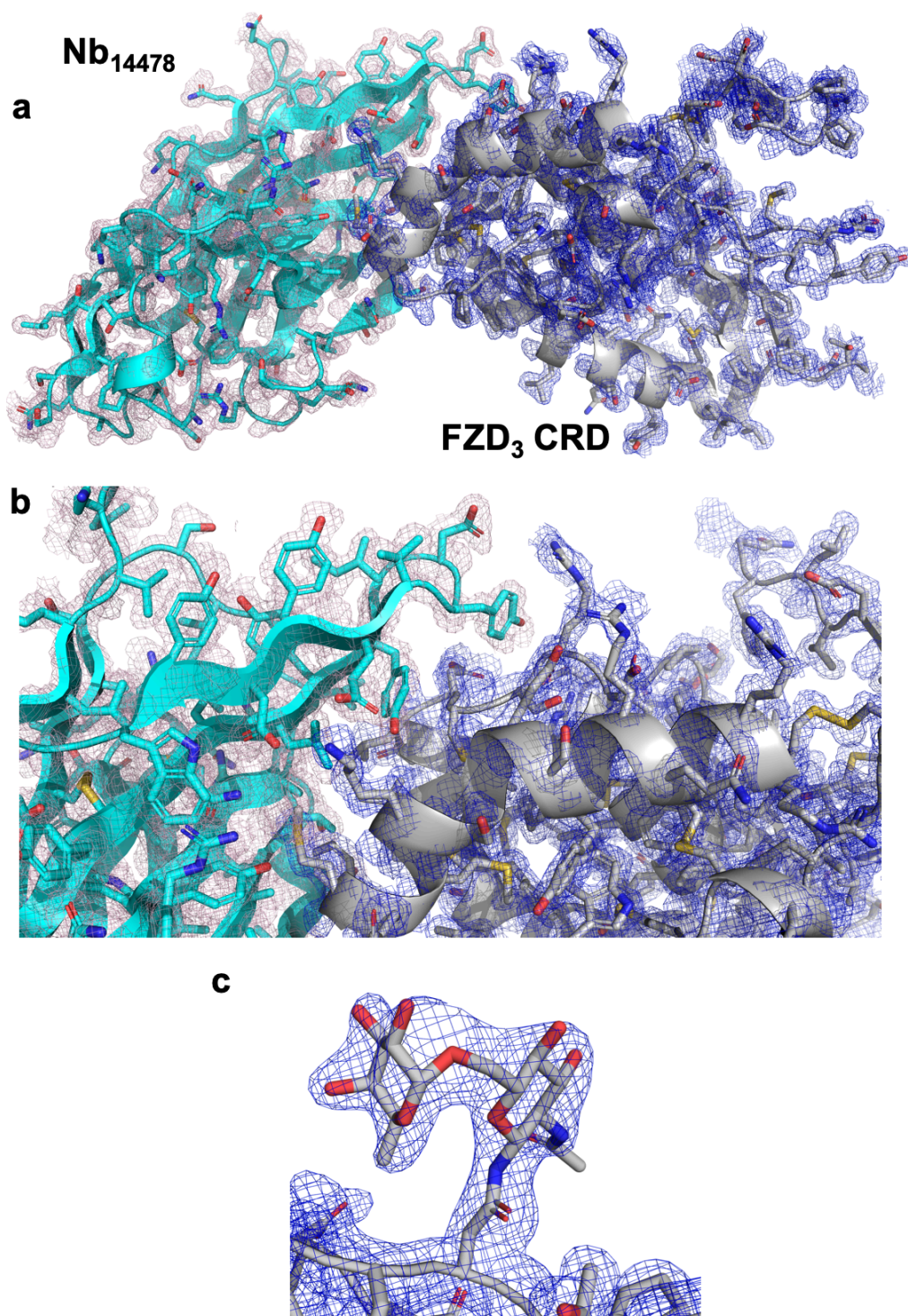


Figure 4.4 Electron density of FZD₃ CRD:Nb₁₄₄₇₈ structure

The final electron density was calculated using SigmaA m2Fo-DFc coefficients as calculated by the program SigmaA (Read, 1986) and contoured at $\sigma = 1.0$ (mesh) contoured at 1.0. FZD₃ CRD:Nb₁₄₄₇₈ model displayed as cartoons with side chains displayed as sticks. **a**, Overall structure; **b**, Electron density at FZD₃ CRD:Nb₁₄₄₇₈ interface; **c**, N-linked glycosylation site on FZD₃ CRD. Unexpectedly, fucose could be unambiguously built into the electron density connected to the glycan core.

lipid-binding groove on the FZD₃ CRD, around the putative dimerisation interface (**Figure 4.3, b**).

Intriguingly, Nb₁₄₄₇₈ is seen to bind the FZD CRD in a ‘sideways’ orientation, with the majority of the interaction formed from the concave surface of the five-stranded β -sheet, formed primarily from the variable regions, which is extended by the long CDR3 loop, and wraps around the rounded surface of the CRD (**Figure 4.6, a**). This is in contrast to the typical front-wise binding exhibited by Nbs, whereby the majority of the interaction is mediated by the CDR3 loop inserting into concave epitopes on the antigen (De Genst *et al.*, 2006).

The interaction between Nb₁₄₄₇₈ and FZD₃ is predominantly mediated through hydrophobic interactions between the concave surface of Nb₁₄₄₇₈ and the round surface of the FZD₃ CRD, which exhibit strong shape complementarity (**Figure 4.3, c; Figure 4.5**). However, the interaction also involves electrostatic interactions between the negatively-charged E105 and D107 residues of the Nb₁₄₄₇₈ CDR3 and the positively-charged K112 of the FZD₃ CRD (**Figure 4.3, d**). Other CDR3 residues that form contacts with the FZD₃ CRD include N96, which contacts M116; and L98, which also contacts M116, as well as L113 (**Figure 4.6, b**). The Y103 side chain of the CDR3 inserts into a small groove formed by the side chains of D72 and R75 of the FZD₃ CRD (**Figure 4.6, b**). Two interacting residues are contributed by the CDR1 of Nb₁₄₄₇₈, I31 and A33 (**Figure 4.6, c**), which contact N70 and L71; and L71 and F117, respectively. Three CDR2 side chains contact the FZD₃ CRD (**Figure 4.6, c**), including A50, which contacts M116 and F117. Meanwhile, T52 contacts F117 of the FZD₃ CRD, while N58 contacts M116 and F117. A total surface area of approximately 680 Å² is buried at the interface between FZD₃ CRD and Nb₁₄₄₇₈.

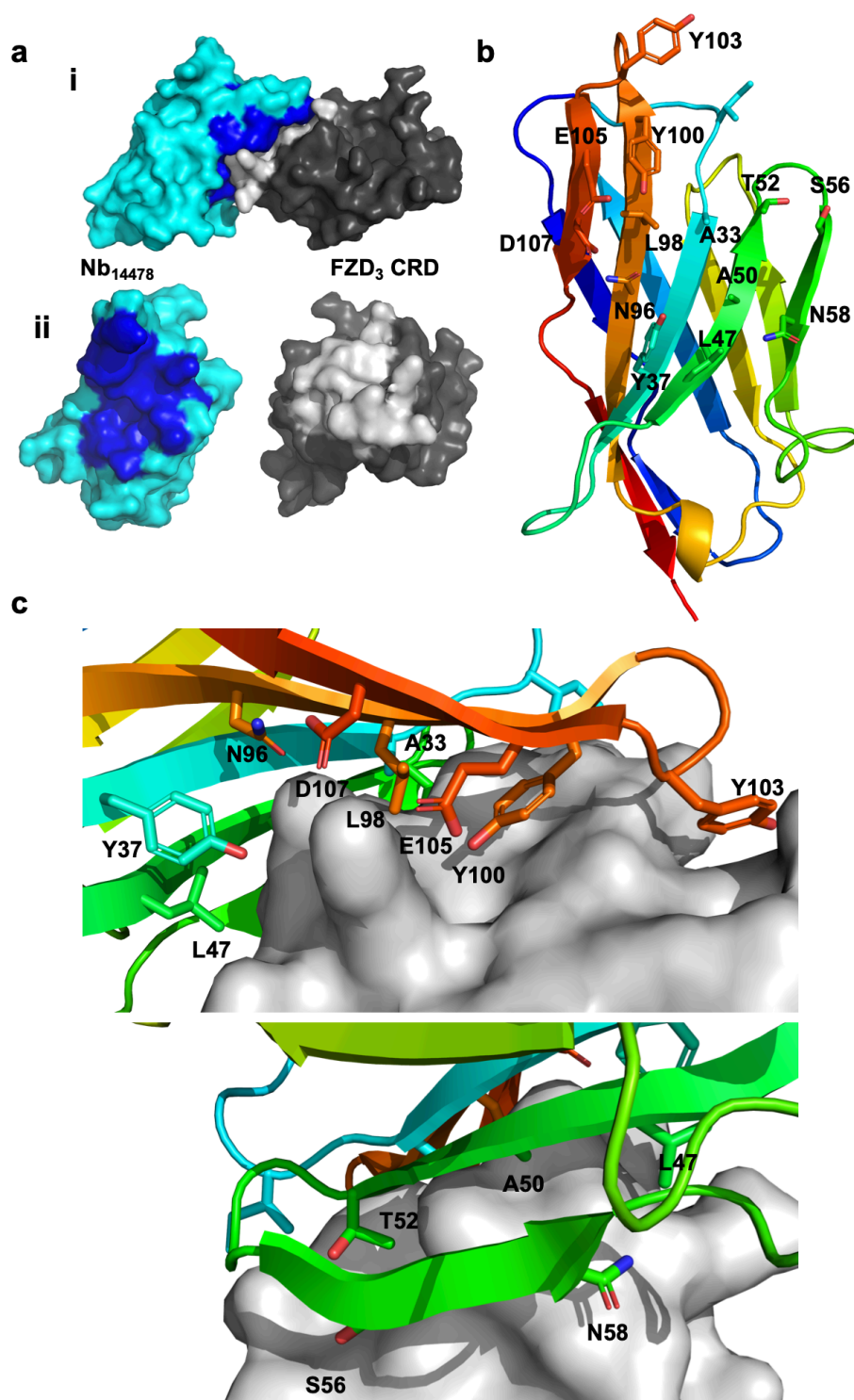


Figure 4.5 Detail of the interactions between Nb₁₄₄₇₈ and FZD₃ CRD

a, i, Crystal structure of Nb₁₄₄₇₈ (cyan) in complex with FZD₃ CRD (grey) with interaction surfaces coloured blue and light grey, respectively, **ii**, 'open book' view of the interaction interfaces. **b**, Crystal structure of Nb₁₄₄₇₈ with FZD₃ CRD-interacting residues labelled. **c**, Detailed view of the interactions between Nb₁₄₄₇₈ residues and the FZD₃ CRD.

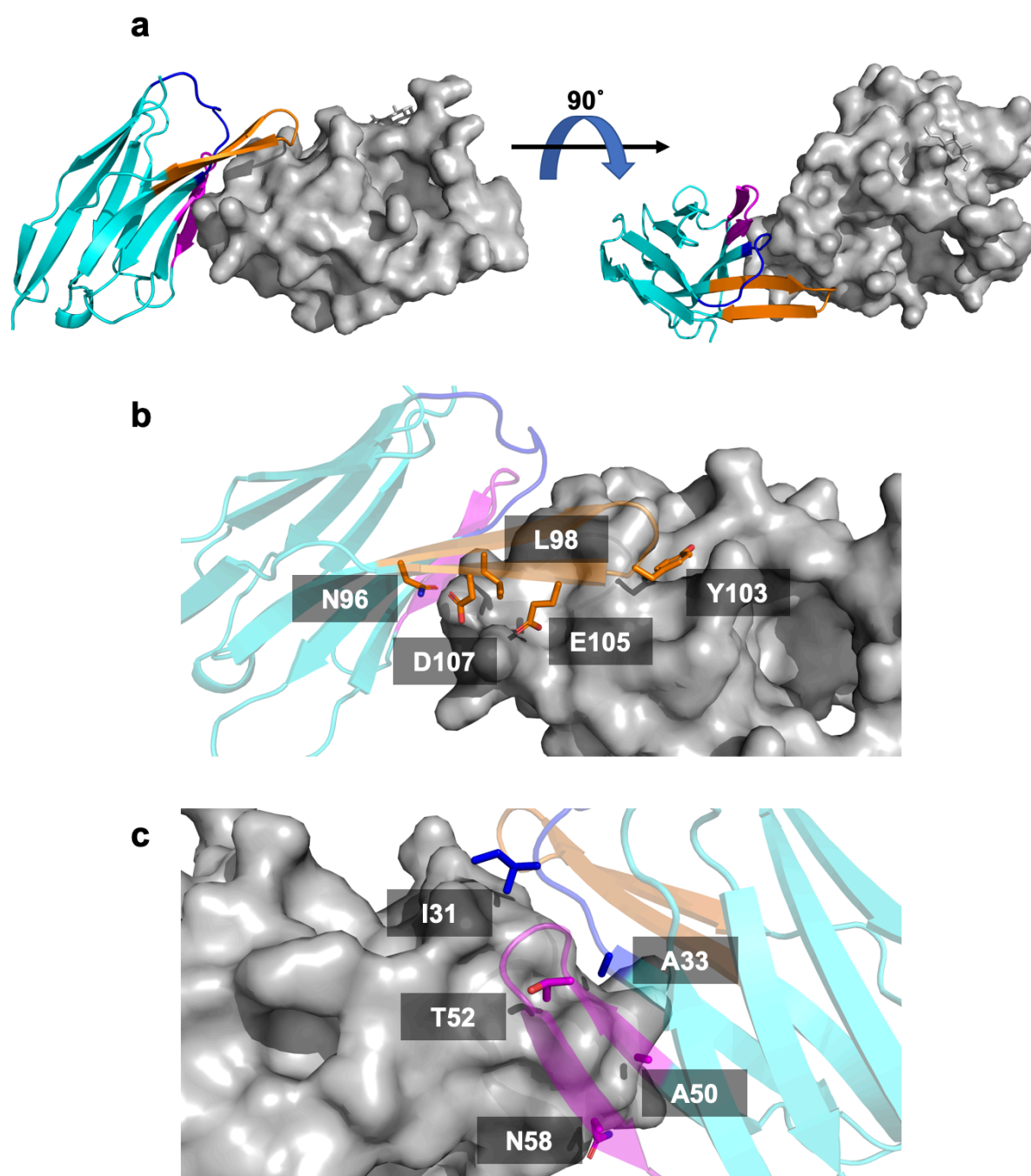


Figure 4.6 Contributions of the Nb₁₄₄₇₈ CDRs to the interaction with FZD₃

a, Interaction between Nb₁₄₄₇₈ (cyan) and FZD₃ CRD (grey). Nb₁₄₄₇₈ CDRs 1-3 coloured blue, magenta and orange, respectively. **b** and **c**, Nb₁₄₄₇₈ CDR3 (**b**) and CDR1/2 (**c**) residues that interact with FZD₃ CRD.

4.4.2 Potential implications of Nb₁₄₄₇₈ binding on FZD₃ signalling

Interestingly, the Nb₁₄₄₇₈ binding site overlaps with the putative FZD₃ CRD dimerisation interface (**Figure 4.3, b**), which is observed in structures of FZD_{4/5/7/8} CRD (DeBruine *et al.*, 2017; Nile *et al.*, 2017). This suggests that if FZD₃ dimerises via this interface, Nb₁₄₄₇₈ binding would abolish this ability to dimerise. Preventing the dimerisation of FZD₃ CRDs at the cell surface could have implications for signalling. For example, although the effect of CRD dimerisation on FZD activation remains obscure, if it is important for maximal WNT signalling then Nb₁₄₄₇₈ may inhibit FZD₃ activation by disrupting homodimerisation.

Nb₁₄₄₇₈ binds to a site on the CRD that is opposite to the putative WNT binding site (**Figure 4.7, a**). Interestingly, despite this, when the FZD₃:Nb₁₄₄₇₈ structure is aligned to the structure of WNT3 in complex with FZD₈ CRD, threonine 52 of Nb₁₄₄₇₈ is situated at the bottom of the hydrophobic groove and clashes with the tip of the palmitoleic acid moiety on WNT3 (**Figure 4.7, b**) that inserts into the hydrophobic groove on FZD. This suggests that Nb₁₄₄₇₈ could possibly block WNT binding to FZD₃ CRD by preventing the palmitoleic acid modification from entering into the hydrophobic groove, which is a requirement for WNT-FZD signalling, and would certainly prevent palmitoleic acid-dependent CRD dimerisation.

The hydrophobic groove on the FZD₃ CRD is markedly narrower than that of other published FZD CRD structures (**Figure 4.8**, Nile *et al.*, 2017). It cannot be ruled out that this is an intrinsic property of the FZD₃ CRD, however it appears more likely that the narrowing is induced by Nb₁₄₄₇₈ binding, as such a narrow groove would seemingly prevent FZD₃ from binding to WNTs *in vivo*. Furthermore, the groove of FZD CRDs is known to exhibit flexibility as the groove of palmitoleic acid-bound FZD CRDs adopts a more ‘open’ conformation compared to the ‘closed’ conformation observed in apo structures (Nile *et al.*, 2017).

Assuming that the observed narrowing of the lipid-binding groove is indeed caused by Nb₁₄₄₇₈ binding, it is possible that Nb₁₄₄₇₈ could also prevent WNT binding to FZD₃ by restricting the

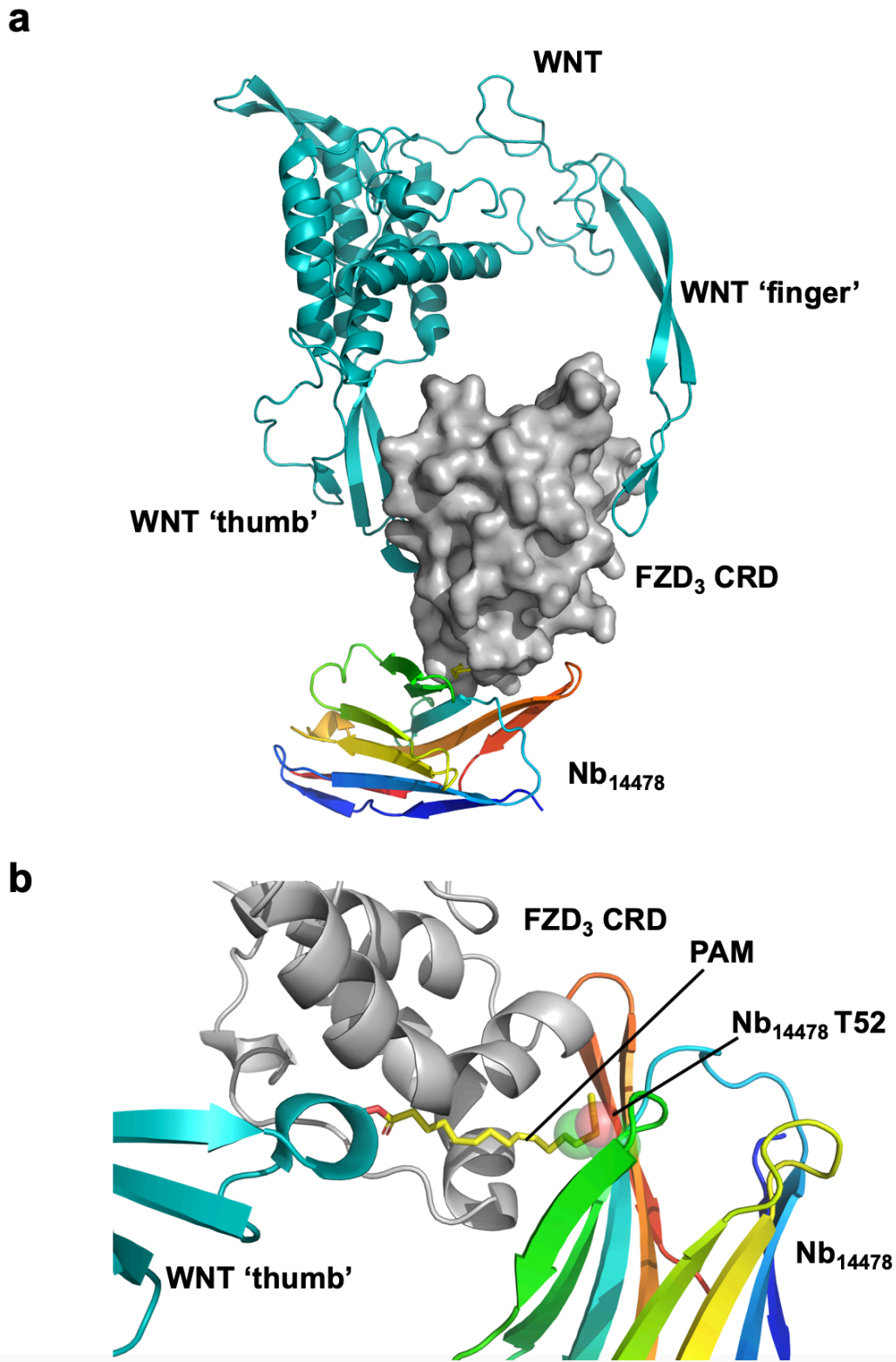


Figure 4.7 Nb₁₄₄₇₈ binding to FZD₃

a Structure of FZD₃ CRD (grey) in complex with NB₁₄₄₇₈ (rainbow) with WNT (teal) superimposed at predicted binding site. NB₁₄₄₇₈ binds away from the WNT binding site. **b**, The palmitoleic acid (yellow sticks) attached to WNT clashes with T52 of Nb₁₄₄₇₈ (spheres).

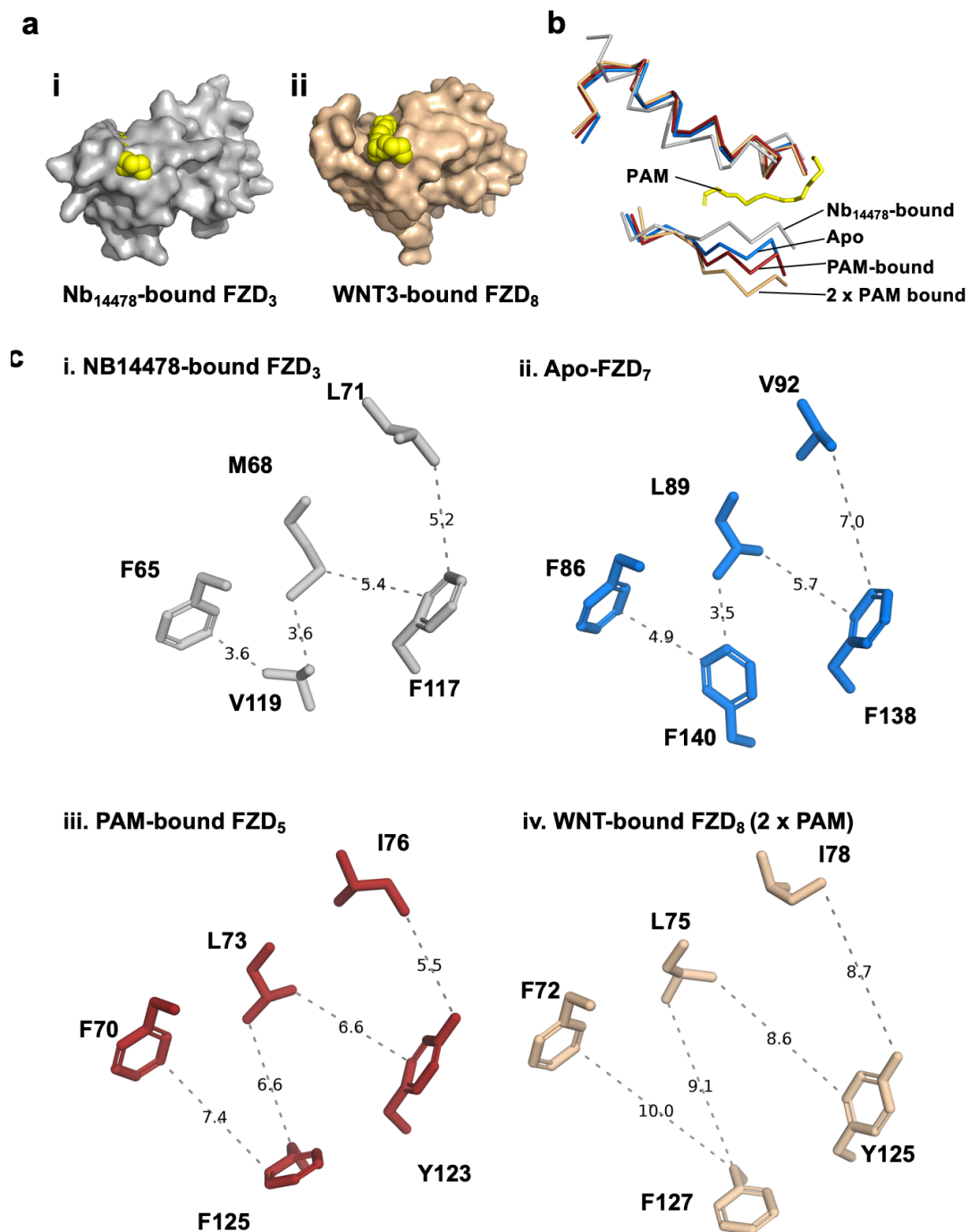


Figure 4.8 Nb₁₄₄₇₈ appears to cause narrowing of hydrophobic groove on FZD₃ CRD

a, Comparison between **(i)** the ‘closed’ groove observed in Nb₁₄₄₇₈-bound FZD₃ (grey) and **(ii)** the ‘open’ groove observed in WNT-bound FZD₈ CRD (PDB 6AHY, wheat). Palmitoleic acid indicated by yellow spheres. **b**, Comparison of the hydrophobic grooves of Nb14478-bound FZD₃ CRD (grey), **(ii)** apo-FZD₇ CRD (PDB 5T44, blue), **(iii)** palmitoleic acid-bound

FZD₅ CRD (PDB 5URY, burnt red) and WNT-bound FZD₈ (wheat). The width of the groove decreases relative to the apo state upon Nb₁₄₄₇₈ binding but increases when palmitoleic acid is bound. The groove expands further in the WNT-bound structure, as it contains two copies of palmitoleic acid as this structure is a 2:2 WNT:FZD complex. **c**, Relative orientations of, and distances between, key residues lining the hydrophobic grooves of Nb₁₄₄₇₈-bound (**i**), apo (**ii**), palmitoleic acid-bound (**iii**) and 2 x palmitoleic acid (WNT)-bound (**iv**) FZD CRDs. Distance units are Å.

size of its hydrophobic groove, rendering it inaccessible to the WNT palmitoleic acid modification. In order to test this hypothesis, crystals of the FZD₃ CRD:Nb₁₄₄₇₈ complex were soaked in reservoir solution containing high concentrations of palmitoleic acid (10 mg mL⁻¹) prior to freezing, to try to determine if palmitoleic acid could occupy the Nb-bound hydrophobic groove. No widening of the groove or extra electron density within the groove was observed, indicating that palmitoleic acid could not be accommodated in the Nb₁₄₄₇₈-bound CRD. This suggests that Nb₁₄₄₇₈ may indeed prevent FZD_{3/6} activation by constricting the hydrophobic groove, rendering it unable to accommodate the WNT palmitoleic acid moiety. Co-crystallisation trials of the FZD₃ CRD in complex with Nb₁₄₄₇₈ in the presence of a molar excess of palmitoleic acid also did not yield any crystals, again hinting that Nb₁₄₄₇₈ and palmitoleic acid binding are mutually exclusive.

4.4.3 Basis for the specificity of Nb₁₄₄₇₈ binding to FZD_{3/6}

By aligning the amino acid sequences of the regions of the FZD₃ CRD that interact with Nb₁₄₄₇₈ to the equivalent regions on the 9 other FZD CRDs, it is possible to see which of these amino acids are unique to FZD₃ and FZD₆ and may therefore contribute to the specificity of FZD_{3/6}:Nb₁₄₄₇₈ binding. The eight amino acids, sidechains of which interact with Nb₁₄₄₇₈, exhibit different levels of conservation between the FZDs.

For example, K112 of the FZD₃ CRD is conserved in FZD₆ (K108) but is replaced by alanine in the FZD_{1/2/7} subfamily, and proline in the remaining FZDs (**Figure 4.9, a**). As this large, positively-charged residue interacts with two negatively-charged residues on Nb₁₄₄₇₈, it is

likely that this is a major source of specificity for FZD_{3/6}:Nb₁₄₄₇₈ binding, as this interaction could not happen with the smaller, hydrophobic amino acids at this location on other FZDs, leading to weaker binding. In stark contrast, FZD₃ L113 is conserved in all FZDs, except for the FZD_{4/9/10} subfamily, although in these it is substituted by another large, hydrophobic amino acid (isoleucine, valine in FZD₄). Therefore, L113 does not appear to contribute to the specificity of Nb₁₄₄₇₈ to bind FZD_{3/6}. D72 and R75, which sandwich Y103 of Nb₁₄₄₇₈ are not conserved in other FZDs (**Figure 4.9, b**). In FZD₆, D72 becomes glutamate and therefore maintains the negative charge, while in the FZD_{1/2/7} and FZD_{5/8} subfamilies, it is polar glutamine. As Y103 of the Nb appears to interact mainly with the hydrophobic stalk of this residue, it seems likely that either substitution would maintain the interaction. In the FZD_{4/9/10} subfamily, however, this residue is replaced with glycine, so the interaction would be lost. R75 becomes proline in FZD_{2, 5, 6, 7 and 8}; alanine in FZD₁; serine in FZD_{4 and 9}; and glycine in FZD₁₀. As Nb₁₄₄₇₈ still binds to FZD₆, it is apparent that proline is tolerated at this position and could participate in stacking with Nb₁₄₄₇₈ Y103. Serine on the other hand would likely lead to a weaker interaction at this location, while the interaction would not occur with the glycine of FZD₁₀. Therefore, while the interaction of Nb₁₄₄₇₈ Y103 with FZD₃ D72 and R75 does not appear to be a major source of specificity, the fact that these residues are replaced with small amino acids in FZDs _{1, 4, 9 and 10}, resulting in a substantially smaller cleft for Nb₁₄₄₇₈ Y103 to bind, may help to explain why the Nb does not bind detectably to these FZDs.

M116 may also make a contribution to the specificity of Nb₁₄₄₇₈ for FZD_{3/6}. M116 becomes positively-charged lysine in FZD_{1/2/7}, negatively-charged glutamate in FZD₄ and polar glutamine in all other FZDs (**Figure 4.9, c**). All of these substitutions change the nature of the hydrophobic amino acid at this position in FZD₃, which presumably leads to unfavourable interactions with the hydrophobic amino acids, A50 and L98, of Nb₁₄₄₇₈, which interact with M116. Although this residue is replaced by threonine in FZD₆, this is a smaller amino acid, so

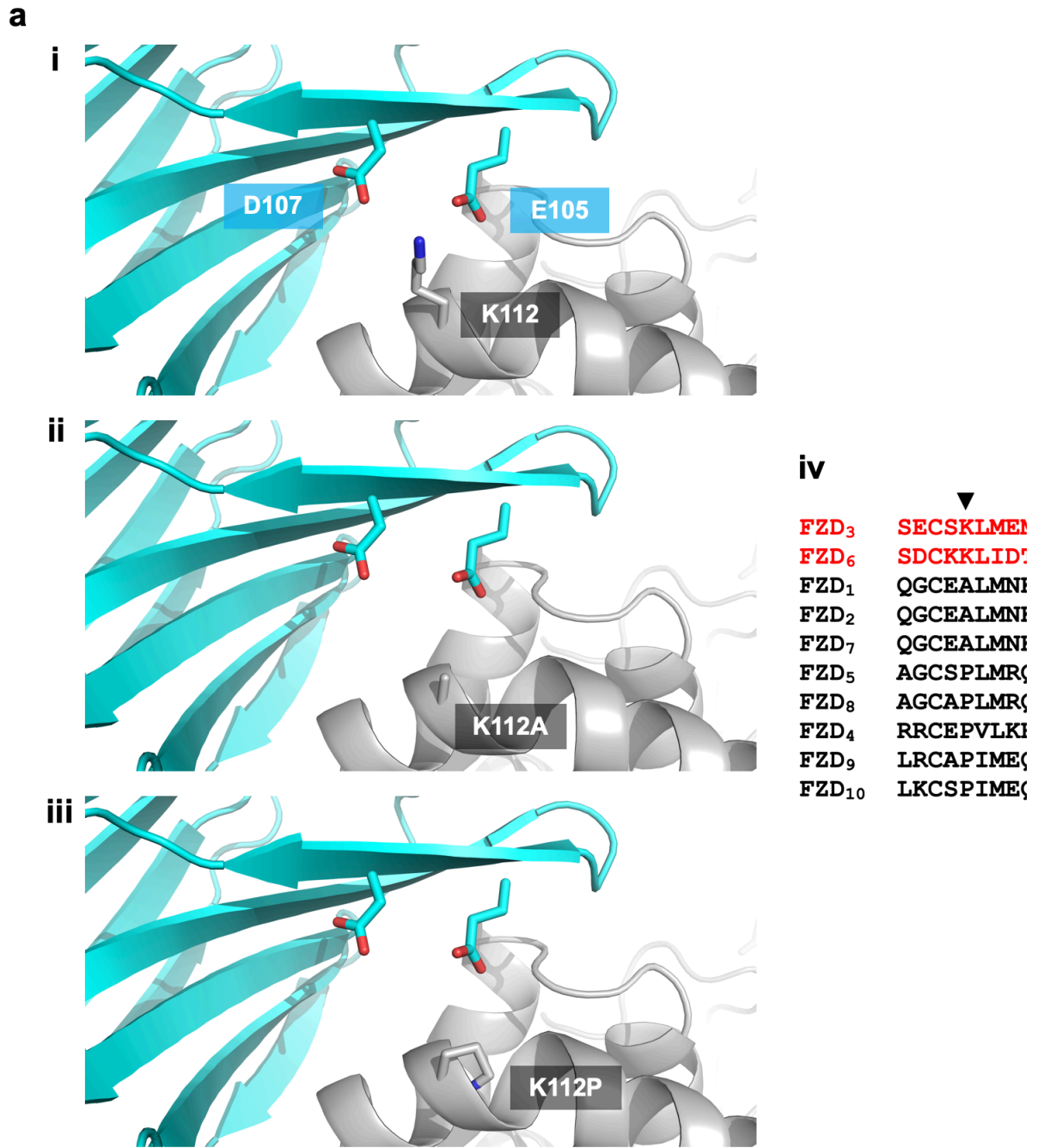
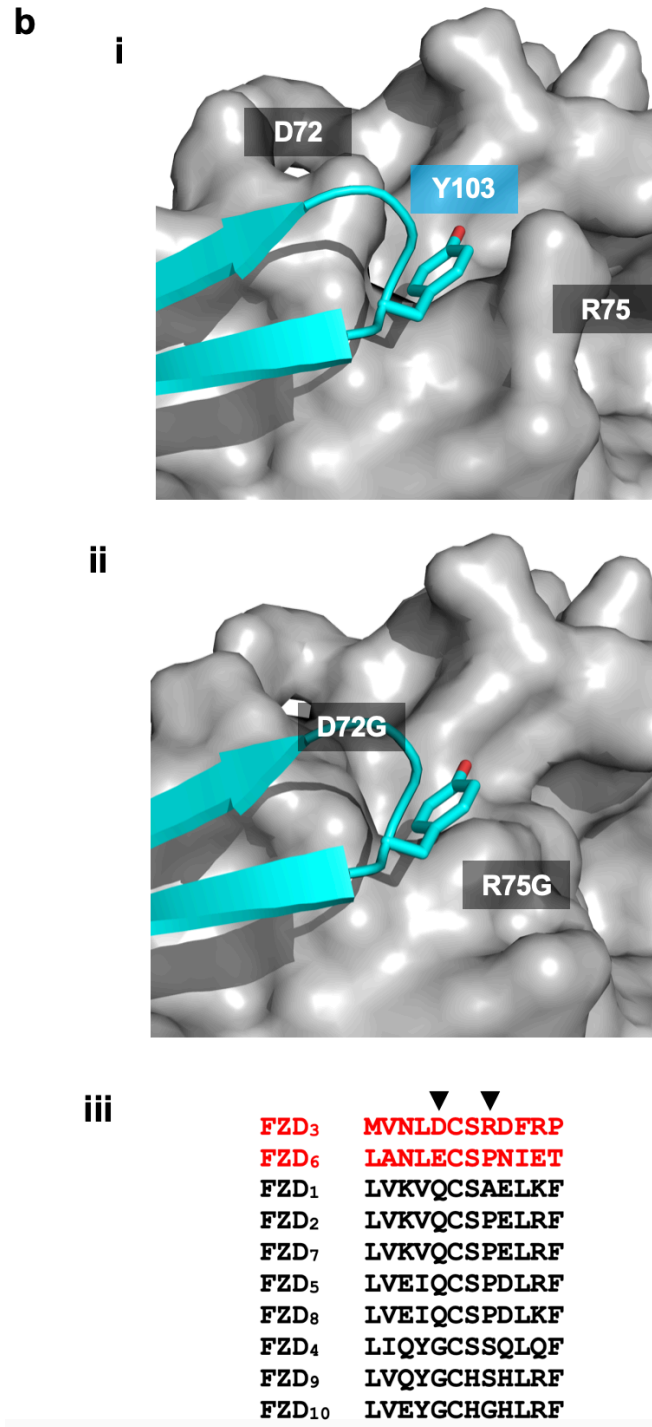


Figure 4.9 Basis of the specificity of Nb₁₄₄₇₈ for FZD_{3/6}

a, i, K112 (FZD₆ K108) interacts with E105 and D107 of Nb₁₄₄₇₈. **ii** and **iii**, K112 is replaced with a hydrophobic amino acid in all other FZDs; (**ii**) alanine in FZD_{1/2/7} and (**iii**) proline in FZD_{4, 5, 8, 9} and ₁₀. The electrostatic interactions are thereby lost. **iv**, MSA of human FZDs with position of FZD₃ K112 indicated by arrow.

**Figure 4.8 continued**

b, i, Nb₁₄₄₇₈ Y103 inserts into a notch formed by D72 and R75 of FZD₃ (E68 and P71 in FZD₆). **ii**, In FZD_{4/9/10} D72 becomes glycine, leading to a large decrease in the size of this groove, which is further decreased in FZD₁₀, where R75 is also replaced by glycine (shown). This likely causes a reduced affinity between these FZDs and Nb₁₄₄₇₈. **iii**, MSA of human FZDs showing position of FZD₃ D72 and R75, indicated by arrows.

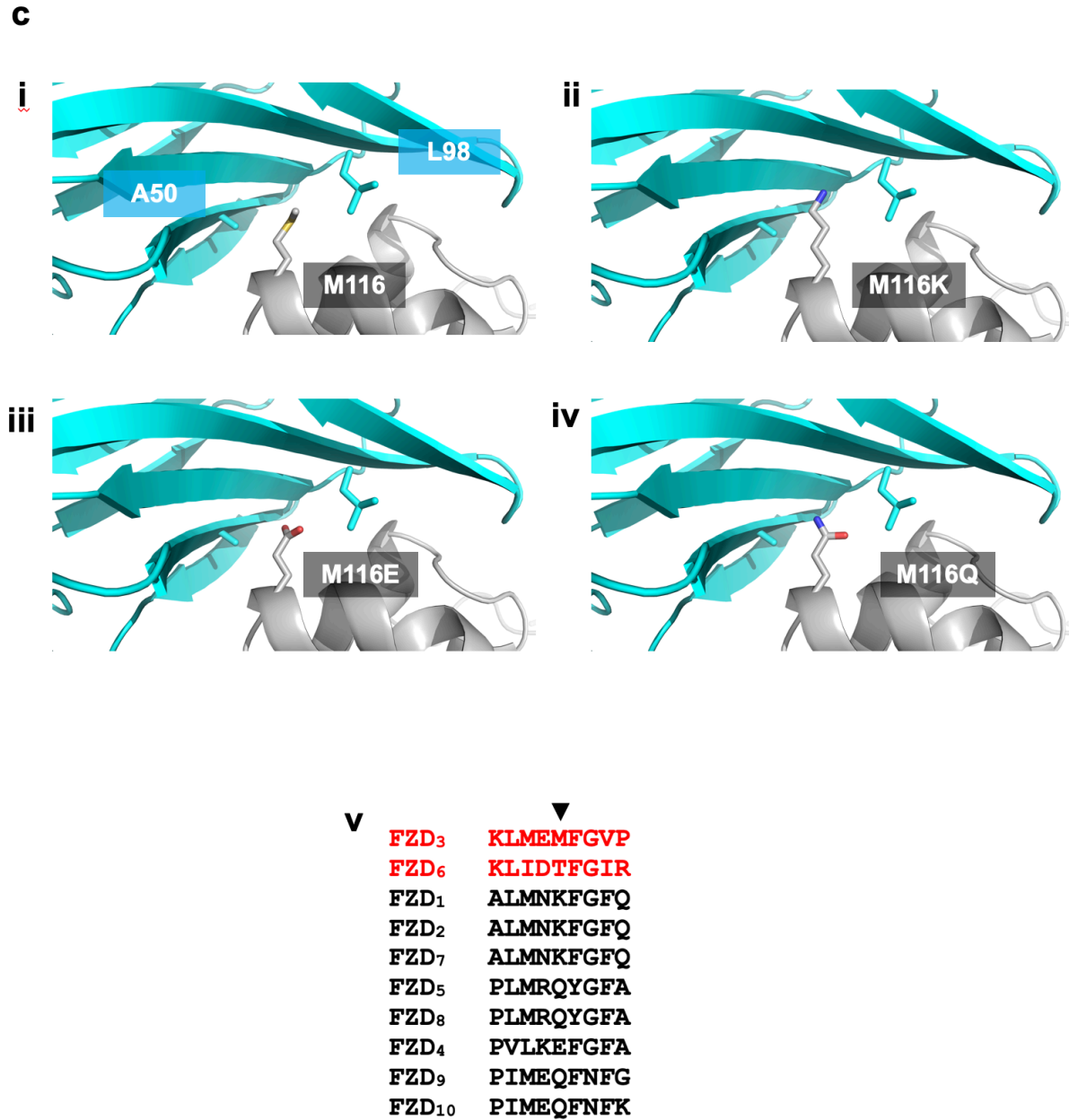


Figure 4.8 continued

c, i, FZD₃ CRD M116 interacts with the hydrophobic residues, A50 and L98, of Nb₁₄₄₇₈. ii-iv, M116 is replaced by (ii) positively-charged lysine in FZD_{1/2/7}; (iii) negatively-charged glutamate in FZD₄; (iv) and polar glutamine in FZD_{5, 8, 9} and ₁₀. All of these substitutions lead to less favourable interactions with Nb₁₄₄₇₈ A50 and L98. Although this amino acid is replaced with polar threonine in FZD₆, its smaller size is likely to mitigate these negative impacts on binding. v, MSA of human FZDs, with the position of FZD₃ M116 indicated by arrow.

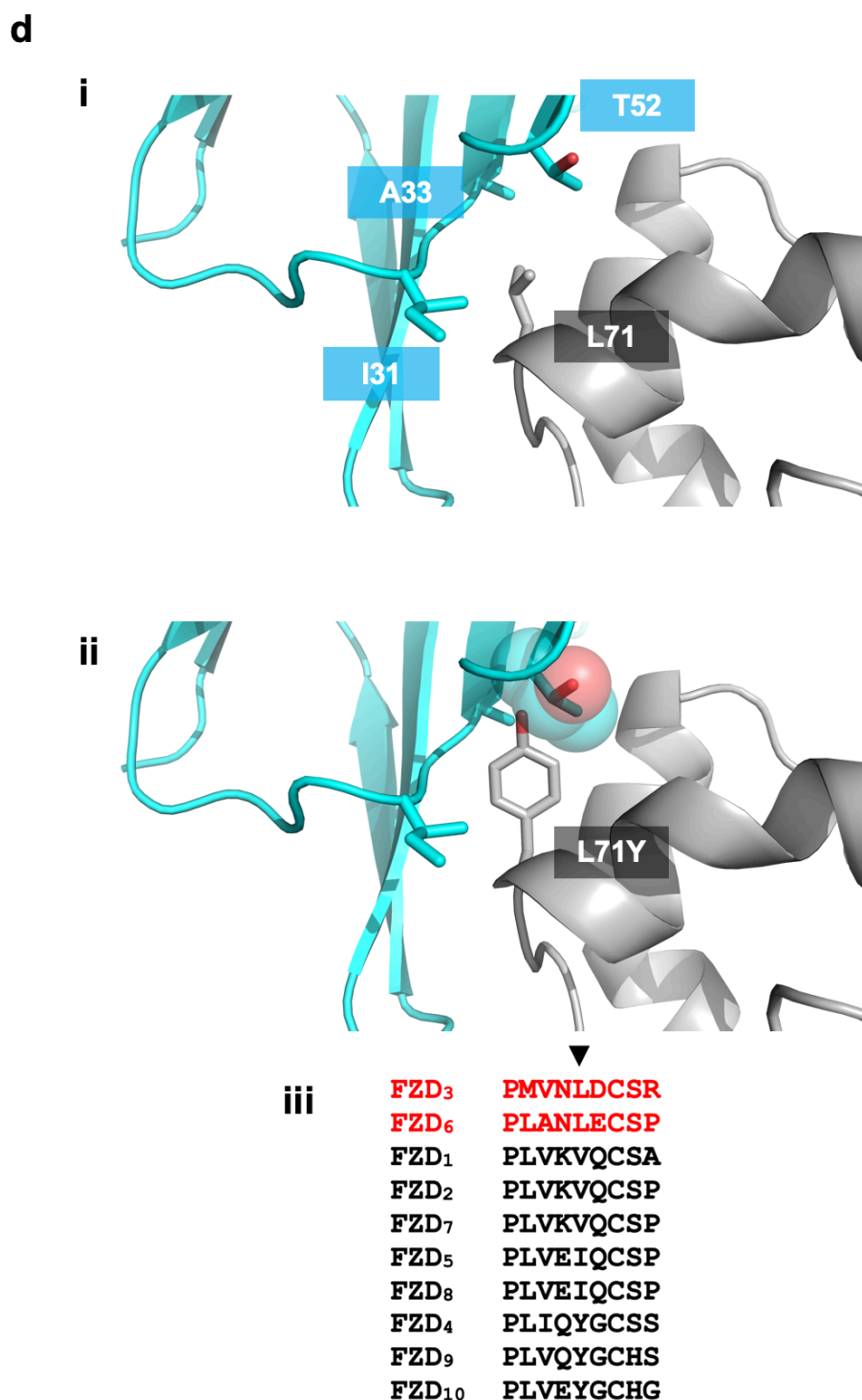
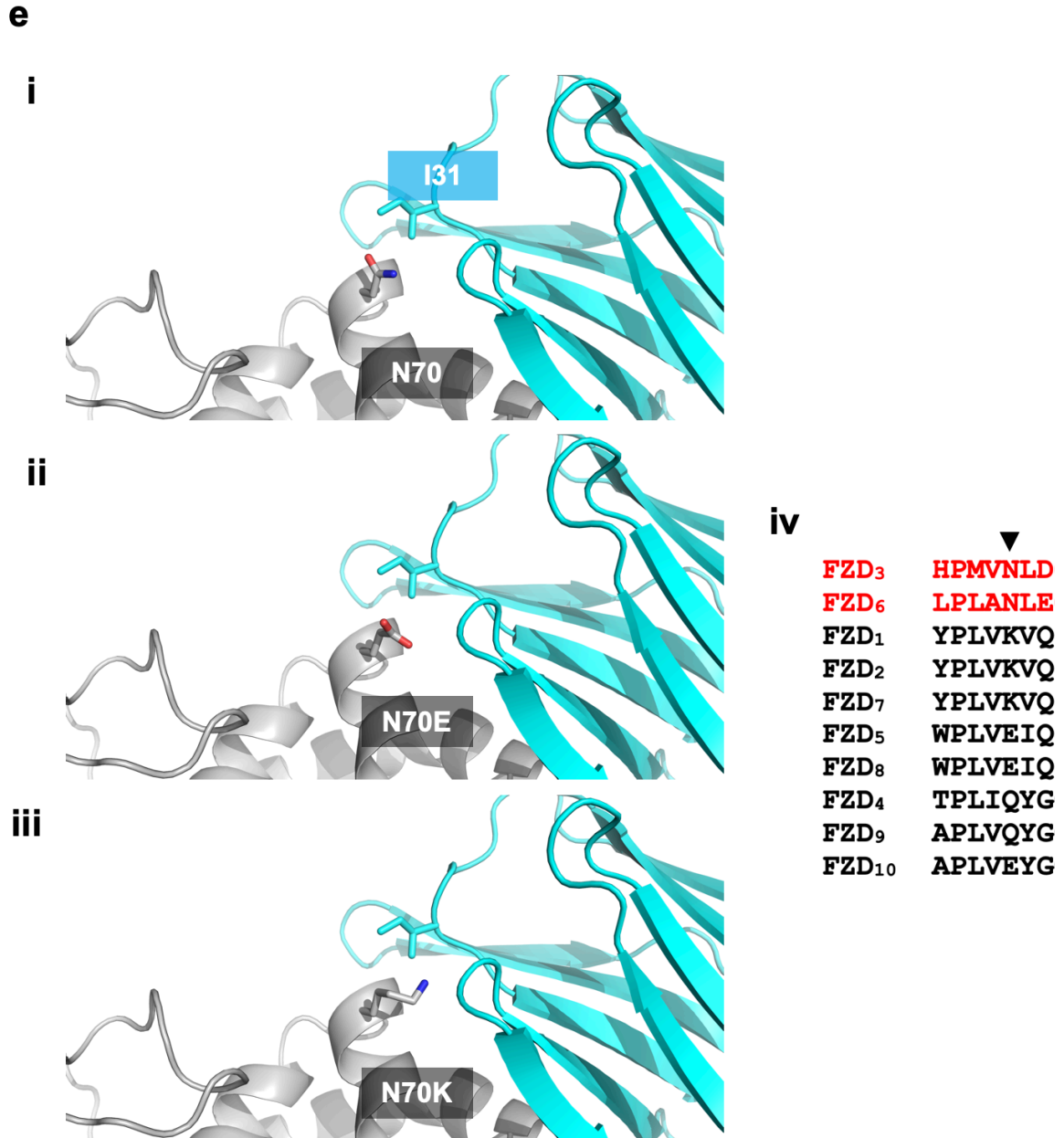


Figure 4.8 continued

d, i, FZD₃ L71 interacts with L31 and A33 of Nb₁₄₄₇₈. ii, In the FZD_{4/9/10} subgroup, substitution of the equivalent residue to FZD₃ L71 with tyrosine leads to a clash with threonine 52 (spheres) of Nb₁₄₄₇₈, which could prevent Nb₁₄₄₇₈ from binding these FZDs. iii, MSA of all human FZDs with position of FZD₃ L71 marked by arrow.

**Figure 4.8 continued**

e, i, FZD₃ N70 interacts with I31 of Nb₁₄₄₇₈. ii and iii, N70 is substituted with (ii) negatively-charged glutamate in FZD_{5, 8} and ₁₀, (iii) while in the FZD_{1/2/7} subfamily this residue becomes positively-charged lysine. In either case, these substitutions are likely to reduce the favourability of the interaction of this residue with the hydrophobic I31 of Nb₁₄₄₇₈. iv, MSA of human FZD sequences with position of FZD₃ N70 indicated by arrow.

while the hydrophobic interactions may be lost, it is less likely that they will be replaced by unfavourable interactions.

Similarly to K112, N70 and L71 of FZD₃ are both conserved in FZD₆ but not found in other FZDs (**Figure 4.9, d and e**). The equivalent residue to N70 is positively-charged (lysine) in the FZD_{1/2/7} subfamily; negatively-charged (glutamate) in the FZD_{5/8} subfamily, as well as FZD₁₀; and polar (glutamine) in FZD₄ and ₉. Therefore, while the interaction between FZD₃ N70 (a polar residue) and Nb₁₄₄₇₈ I31 (hydrophobic) would logically not contribute to the affinity between the two proteins, a larger, charged amino acid at this position would likely lead to clashes or otherwise weaken the interaction. On the other hand, the equivalent to L71 remains a bulky, hydrophobic or aromatic residue in all other FZDs (valine in FZD_{1/2/7}, isoleucine in FZD_{5/8} and tyrosine in FZD_{4/9/10}) and is therefore unlikely to contribute much to the specificity of the interaction. However, it should be noted that tyrosine appears to be too big to be accommodated by Nb₁₄₄₇₈, which may explain why it does not bind to FZD_{4/9/10}.

In summary, the specificity of the Nb₁₄₄₇₈:FZD_{3/6} interaction seems to be determined by surprisingly few amino acids. In particular, the electrostatic interactions between Nb₁₄₄₇₈ E105 and D107; and FZD_{3/6} K112 appears to be a major source of this specificity, as no other FZD could participate in this interaction. The affinity of Nb₁₄₄₇₈ for FZD_{1/2/7} and FZD_{5/8} would be weakened relative to FZD₃ by the substitution of FZD₃ N70 with a charged amino acid, whereas the ability of Nb₁₄₄₇₈ to bind FZD_{4/9/10} would be hindered by the presence of the bulky tyrosine side chain that replaces FZD₃ L71, as well as the loss of the cleft formed by FZD₃ D72 and R75. The substitution of M116 for a large, charged amino acid in FZD_{1, 2, 4} and ₇ is also likely to reduce the ability of Nb₁₄₄₇₈ to bind these FZDs.

4.4.4 Nb₁₄₄₇₈ binding in the context of full-length FZD₃

Homology models of FZD₃ were produced using cholesterol-bound inactive SMO (**Figure 4.10, a, i**, PDB 5LZD, model 1) and cholesterol-bound ‘active’ SMO (**Figure 4.10, a, ii**, PDB

6D35, model 2), to analyse the CRD:Nb interaction in the context of the full-length receptor. Surprisingly, when the FZD CRD:Nb₁₄₄₇₈ structure is aligned to model 1, the majority of the Nb clashes with the LD and extracellular loops of the TMD, indicating that it would not be possible for the Nb-bound receptor to adopt this conformation. However, in model 2, the CRD is rotated away from the linker domain and extracellular loops, meaning that when the FZD₃ CRD:Nb₁₄₄₇₈ structure is aligned, there are no major clashes with the receptor.

This implies that binding of the Nb to full-length FZD₃ would restrict the conformations that the extracellular domain would be able to adopt, as the Nb-bound CRD would be incapable of folding down against the LD. More pertinently, when aligned to model 2, the CDR loops of the Nb, which are only partially engaged by the FZD₃ CRD, are positioned in close proximity to the FZD₃ LD (**Figure 4.10, a, iii**). Many residues on Nb₁₄₄₇₈ that are not involved in the interaction with the CRD are exposed on the surface facing the LD. This raised the possibility that the interaction observed between the FZD₃ CRD and Nb₁₄₄₇₈ in the crystal structure was only part of the whole FZD₃:Nb₁₄₄₇₈ interaction, and that the Nb may simultaneously be able to bind the CRD and LD. This would explain the unusual binding mode observed for Nb₁₄₄₇₈ to the FZD₃ CRD; although it binds to the convex surface of the CRD, binding a concave epitope between the CRD and LD would be more typical for a Nb. If this were the case, the Nb could essentially lock the conformation of the CRD relative to the LD, thereby overcoming one of the main hurdles that has prevented the crystallisation of a multidomain FZD receptor to date.

In order to test this hypothesis, the affinity of the interaction between Nb₁₄₄₇₈ and a FZD₃ CRD construct containing the LD (FZD₃ CRD_LINKER) was measured so it could be compared to the affinity of the interaction of Nb₁₄₄₇₈ with the FZD₃ CRD alone. The resulting K_d was almost identical in the presence of the LD (**Figure 4.10, b**), which appears to exclude the possibility of Nb₁₄₄₇₈ binding to this region. It therefore appears that Nb₁₄₄₇₈ binds exclusively to the

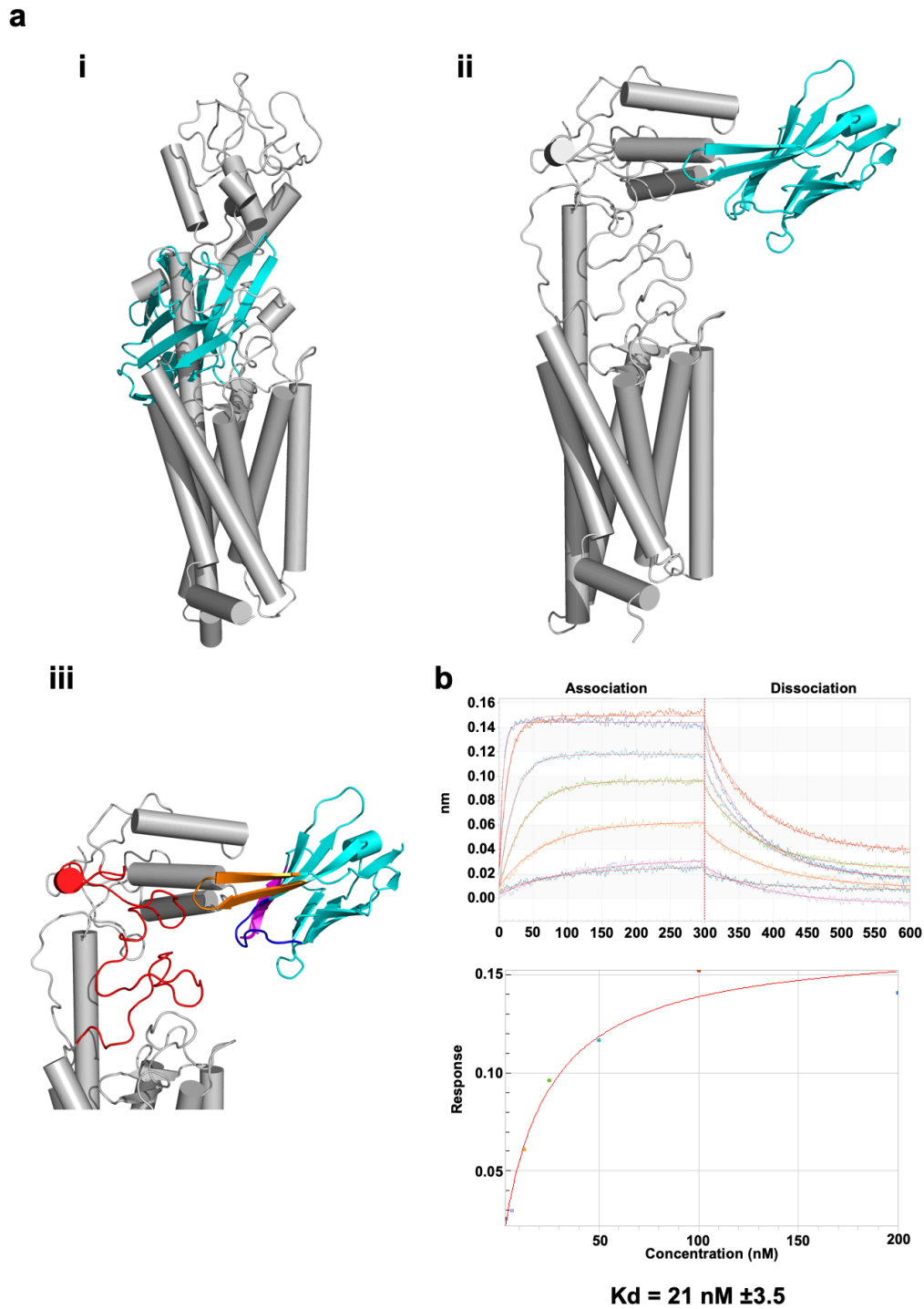


Figure 4.10 The CDRs of Nb₁₄₄₇₈ do not interact with FZD₃ LD

a, Homology models of FZD₃ (grey) using i, inactive, cholesterol-bound SMO (PDB 5L7D) and ii, 'active' SMO (PDB 6D35) as templates. In each case, Nb₁₄₄₇₈ (cyan) is superimposed at its interaction site. iii, in the homology model based on 'active' SMO, the rotation of the CRD leads to the CDRs of Nb₁₄₄₇₈ (blue, magenta and orange) being proximal to the FZD₃ LD (red), raising the possibility of an interaction between the two. b, BLI reveals a near-identical K_d for Nb₁₄₄₇₈ binding to FZD₃ CRD_LINKER compared to FZD₃ CRD alone, ruling out an interaction between Nb₁₄₄₇₈ and the FZD₃ LD.

FZD_{3/6} CRD. Initial attempts to crystallise a multidomain FZD₃:Nb₁₄₄₇₈ complex did not yield any crystals.

4.5 Conclusions and future perspectives

This work lead to the successful identification of a Nb that binds with high affinity and specificity to the FZD_{3/6} CRD. Furthermore, the crystal structure of the FZD₃:Nb₁₄₄₇₈ complex suggests that this Nb will be able to inhibit FZD₃ activation by WNT. Based on this observation, experiments were designed to assess whether Nb₁₄₄₇₈ was indeed capable of modulating FZD₃ activity, which are described in the following chapter. Separately, a collaboration is being arranged to assess the efficacy of Nb₁₄₄₇₈ against melanoma cells (Li *et al.*, 2019), as well as colorectal cancer cells that are known to overexpress FZD₃ (Wong *et al.*, 2013). No therapies targeting FZD₃ have been reported to date, so a Nb that specifically blocks FZD₃ could be a useful tool against these tumours.

Chapter 5 Functional investigations of Class F GPCRs

5.1 Introduction

Functional data regarding the role of FZD₃ in β -catenin-dependent WNT signalling remains limited. Expression of *Xenopus* FZD₃ was shown to induce expression of the β -catenin target gene, *Siamois*, in *Xenopus* embryos (Umbhauer *et al.*, 2000). However, FZD₃ was excluded from two studies that sought to dissect specific WNT-FZD interactions in mammalian cells (Ring *et al.*, 2014; Voloshanenko *et al.*, 2017). It is therefore unclear whether human FZD₃ can activate β -catenin-dependent WNT signalling.

Previous research has highlighted the potential of FZD-targeting antibodies in the treatment of cancer (Gurney *et al.*, 2012; Raman *et al.*, 2019), as well as in stimulating WNT signalling (Janda *et al.*, 2017; Tao *et al.*, 2019). However, none of these antibodies target FZD_{3/6}. Interestingly, FZD₃ was recently suggested to be a critical gene in melanoma, which highlights the need for FZD₃-blocking reagents (Li *et al.*, 2019). Following structural and biophysical quantification of Nb₁₄₄₇₈ and its interaction with FZD₃, attention therefore turned to whether this Nb could be used to modulate FZD₃ activity.

5.2 Aims

This chapter describes work that was undertaken to functionally characterise FZD₃ and Nb₁₄₄₇₈. The first aim of this section of the project was to identify a WNT that could activate ‘canonical’ WNT signalling via FZD₃ in cellular signalling assays *in vitro*. This was a prerequisite for investigating the ability of Nb₁₄₄₇₈ to inhibit WNT signalling via FZD₃.

The other aim of this work was to determine whether Nb₁₄₄₇₈ could be developed into a synthetic WNT agonist that activated β -catenin-dependent signalling via FZD₃. If a FZD₃-activating WNT could not be identified, then a synthetic agonist would provide a convenient way of investigating the ability of variant FZD₃ proteins to activate ‘canonical’ signalling.

5.3 Activity of FZD₃ in response to WNTs

5.3.1 WNT3A does not activate β -catenin-dependent signalling in HEK293T cells expressing FZD₃

WNT3A is a ‘promiscuous’ WNT that was shown to activate β -catenin-dependent signalling (**Figure 5.1, a, i**) via any human FZD ectopically expressed in HEK293T cells lacking endogenous FZD_{1/2/7} (HEK293T FZD_{1/2/7}⁻, which are WNT3A insensitive), except for FZD₆ and FZD₉ (Voloshanenko *et al.*, 2017). However, FZD₃ was not investigated in this experiment. FZD₃ is closely related to FZD₆, and these receptors exhibit a level of functional redundancy (Hua *et al.*, 2014; Dong *et al.*, 2018), so it was expected that FZD₃ would also be unable to activate β -catenin-dependent signalling in HEK293T FZD_{1/2/7}⁻ cells. Nb₁₄₄₇₈ was predicted to inhibit WNT signalling via FZD₃ based on the substantial narrowing of the hydrophobic groove observed in the crystal structure of the FZD₃ CRD:Nb₁₄₄₇₈ complex (see section 4.4) and the high affinity of the interaction (see section 4.3.2).

As expected, HEK293T FZD_{1/2/7}⁻ cells did not exhibit β -catenin stabilisation-dependent luciferase activity in response to WNT3A in cell-based assays (**Figure 5.1, b**), even in cells expressing FZD₃, suggesting that FZD₃ is not a receptor for WNT3A. Nb₁₄₄₇₈ was added in the absence of WNT to confirm that it did not activate signalling and indeed, no luciferase activity was observed, despite Nb₁₄₄₇₈ being present at a concentration of 10 μ M. Unmodified HEK293T cells, which endogenously express FZD_{1/2/7}, were sensitive to WNT3A and exhibited luciferase activity upon stimulation. This activity was slightly diminished upon treatment with 10 μ M Nb₁₄₄₇₈, which is in agreement with the observation that this Nb does not exhibit high-affinity binding to any FZD other than FZD_{3/6} (see section 4.3.3). However, the fact that signalling activity was slightly reduced in the presence of a high concentration of Nb₁₄₄₇₈ suggests that Nb₁₄₄₇₈ binding may indeed suppress FZD activity, although the high concentrations required to inhibit these other FZDs also supports the notion that, if Nb₁₄₄₇₈

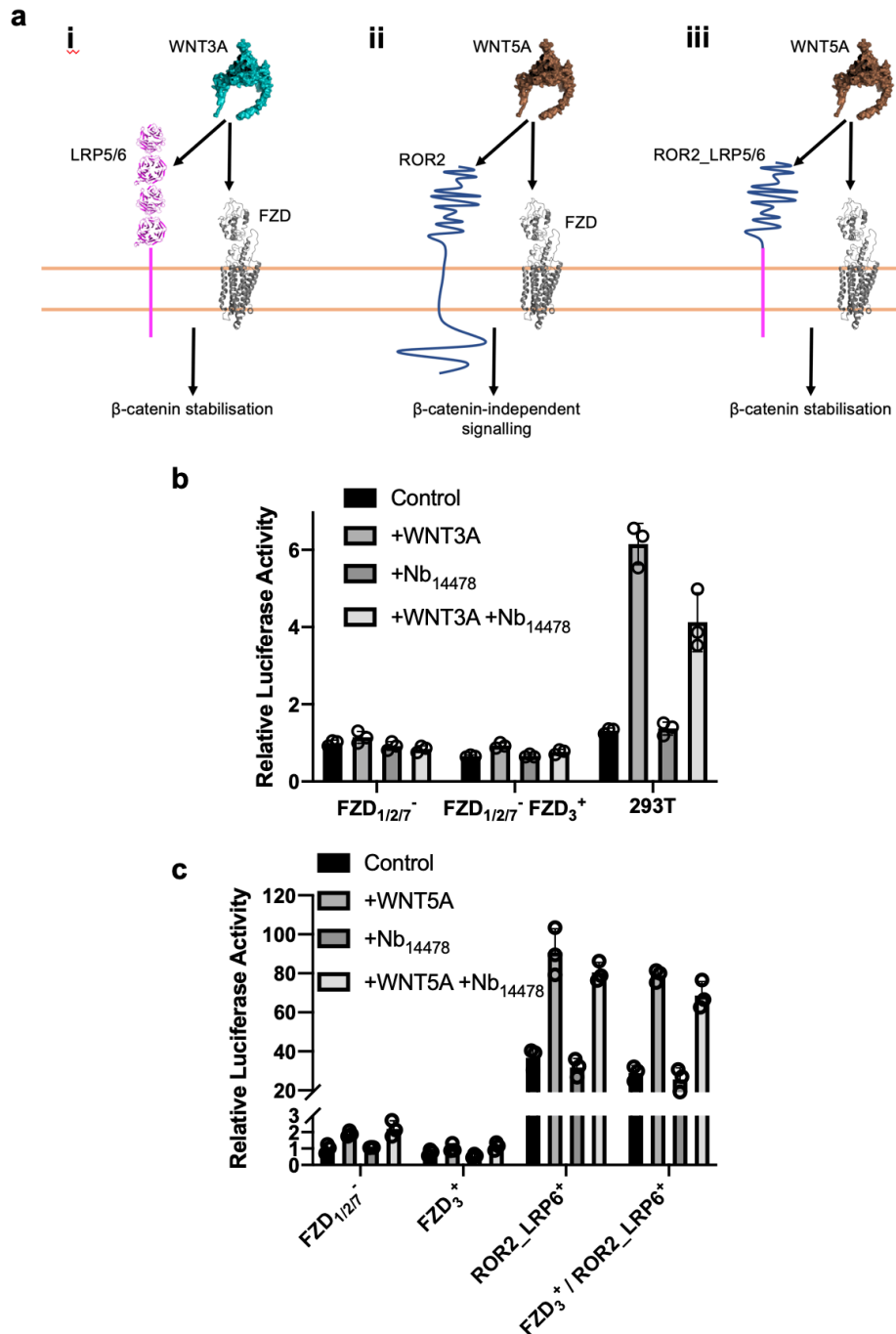


Figure 5.1 FZD₃ is not a receptor for WNT3A or WNT5A in HEK293 cells.

a, WNT3A triggers β -catenin stabilisation by heterodimerising a FZD with an LRP5/6 coreceptor. WNT5A can stimulate β -catenin-independent signalling using ROR2 as a coreceptor. By replacing the extracellular domain of LRP5/6 with the ROR2 extracellular domain, a coreceptor is created that can respond to a noncanonical WNT/FZD pair but that activates β -catenin-dependent signalling via its intracellular LRP5/6 domain. **b**, Luciferase activity relative to unstimulated HEK293T FZD_{1/2/7}⁻ cells. WNT3A stimulates β -catenin-dependent luciferase activity in HEK293T cells but cannot do so in HEK293T FZD_{1/2/7}⁻ cells, including those expressing FZD₃. Nb₁₄₄₇₈ does not stimulate luciferase activity but has a very weak inhibitory effect in HEK293T cells, likely due to weak binding to FZD_{1/2/7}.

were to be used clinically to target FZD₃, it would be unlikely to cause side effects through the inhibition of other FZDs at effective concentrations.

5.3.2 WNT5A activates a chimeric ROR2_LRP6 receptor independently of FZD₃ in HEK293T cells

As FZD₃ did not respond to WNT3A, assays were performed to determine whether this receptor can be activated by WNT5A. Although WNT5A can activate β -catenin-dependent signalling via FZD₄/LRP6 in HEK293 cells, it cannot do so via FZD_{1/2/5/6/7/8/9/10} (FZD₃ not investigated, Ring *et al.*, 2014) and is primarily associated with β -catenin-independent signalling (**Figure 5.1, a, ii**; Sato *et al.*, 2010). However, it has been demonstrated that WNT5A can bind to the ROR2 extracellular domain and, by fusing this domain to the transmembrane and intracellular regions of LRP6, a fusion protein, ROR2_LRP6, can be created that allows ROR2-binding ‘non-canonical’ WNTs to activate β -catenin-dependent signalling (**Figure 5.1, a, iii**; Grumolato *et al.*, 2010). Such a fusion protein allows the formation of a WNT5A-FZD-ROR2 complex outside the cell that resembles a ‘canonical’ FZD-LRP6 complex within the cell, and can therefore stimulate β -catenin stabilisation.

WNT5A was observed to stimulate β -catenin-dependent luciferase activity in HEK293T FZD_{1/2/7}⁻ cells transfected with ROR2_LRP6 (**Figure 5.1, c**), however this effect was unchanged in cells expressing FZD₃, suggesting that this activity could be due to another FZD endogenously expressed in these cells. A likely candidate is FZD₄, which is known to be sensitive to WNT5A (Ring *et al.*, 2014), and is expressed at a low level in HEK293T cells (Voloshanenko *et al.*, 2017). Unsurprisingly, WNT5A did not strongly stimulate β -catenin-dependent luciferase activity in cells lacking ROR2_LRP6. Nb₁₄₄₇₈ treatment only slightly inhibited WNT5A-dependent luciferase activity in cells expressing ROR2_LRP6, which is again unsurprising as this Nb does not bind strongly to FZDs other than FZD_{3/6} (see section 4.3.3). It is noteworthy that very high basal activity was observed in cells expressing

ROR2_LRP6, which may be due to constitutive activation of LRP6 in the absence of its normal extracellular domain (Mao *et al.*, 2001), or the endogenous expression of a WNT that can couple FZD to ROR2.

5.4 A synthetic agonist of β -catenin signalling permits functional dissection of FZD3

5.4.1 A fusion protein comprising Nb₁₄₄₇₈ fused to the DKK1 C-terminal domain functions as a FZD₃-dependent agonist of β -catenin-dependent signalling

The above results suggested, but did not prove, that FZD₃ is not a receptor for WNT3A or WNT5A. An alternative explanation is that FZD₃ may be capable of binding these WNTs but is intrinsically incapable of activating β -catenin-dependent signalling, even when brought into proximity with LRP5/6. Therefore, in order to test whether human FZD₃ is intrinsically able to activate β -catenin-dependent WNT signalling in concert with LRP5/6, a chimeric protein was created comprising the C-terminal LRP6-binding domain of DKK1 fused downstream of Nb₁₄₄₇₈ via a GSGSG repeat (**Figure 5.2, a**). This chimera, known as Nb₁₄₄₇₈-DKK1C, is theoretically capable of binding both FZD₃ and LRP6, thereby bringing the two proteins into proximity (**Figure 5.2, b**). While the C-terminal domain of DKK1 is usually a potent inhibitor of β -catenin-dependent WNT signalling (Brott and Sokol, 2002), similar chimeric proteins comprising the DKK1 C-terminal domain fused to other FZD-binding proteins have previously been used to demonstrate that bringing certain FZDs into close proximity with LRP6 is sufficient to activate β -catenin-dependent WNT signalling (Janda *et al.*, 2017).

Nb₁₄₄₇₈-DKK1C and FZD₃ were co-transfected into HEK293T cells lacking expression of FZD_{1/2/7} (HEK293T FZD_{1/2/7}⁻, Voloshanenko *et al.*, 2017). Nb₁₄₄₇₈-DKK1C was observed to strongly activate β -catenin-dependent luciferase activity in the presence of exogenous FZD₃ (**Figure 5.2, c**), confirming that it can indeed function as a ‘canonical’ WNT agonist.

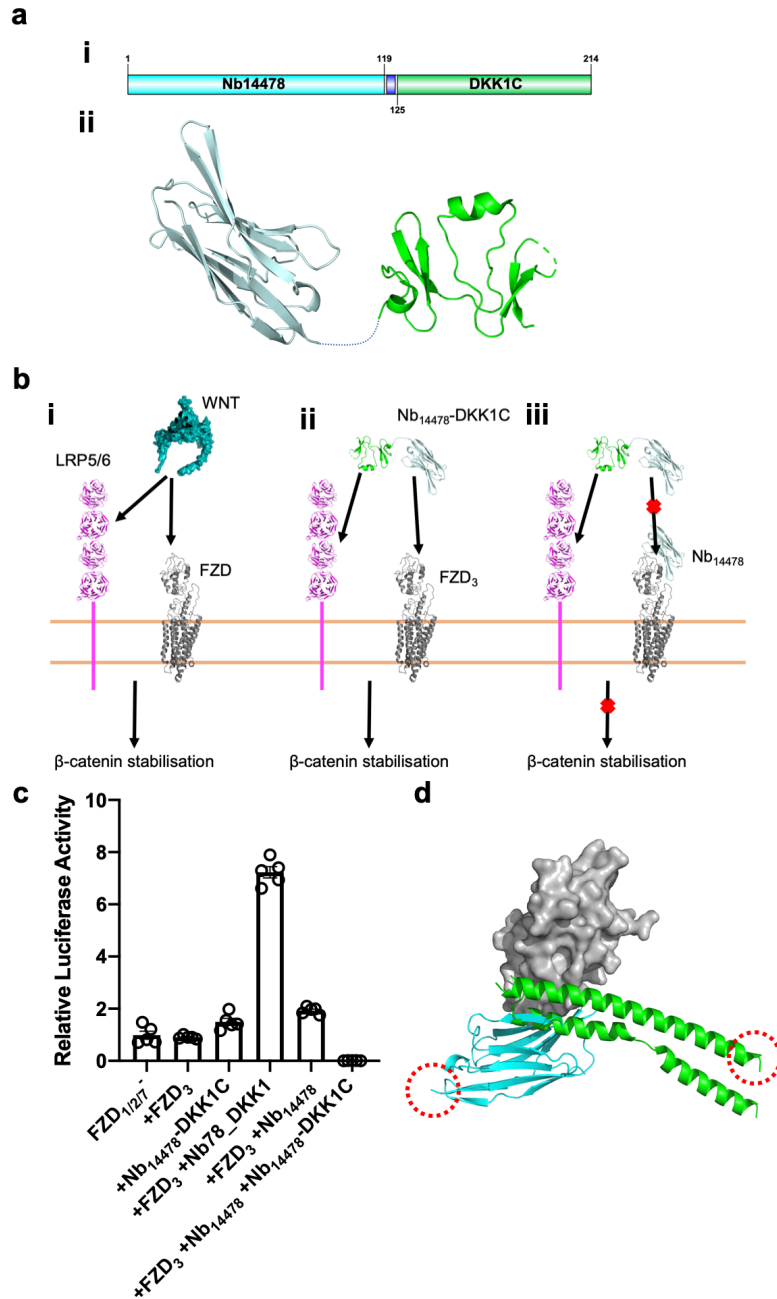


Figure 5.2 Nb₁₄₄₇₈_DKK1C is a synthetic WNT agonist.

a, i, Schematic diagram of the bivalent Nb₁₄₄₇₈_DKK1C fusion protein. **ii**, Known structural information of the components of the fusion protein. **b**, Model for β -catenin-dependent signalling activation by **i**, WNT and **ii**, Nb₁₄₄₇₈_DKK1C. **iii**, Nb₁₄₄₇₈ competes with Nb₁₄₄₇₈_DKK1C for binding to FZD₃, resulting in attenuated β -catenin stabilisation. **c**, Relative firefly luciferase activity in untransfected HEK293T (FZD_{1/2/7}^{-/-}) cells and HEK293T (FZD_{1/2/7}^{-/-}) transfected with FZD₃ or Nb₁₄₄₇₈_DKK1C alone; FZD₃ and Nb₁₄₄₇₈_DKK1; or FZD₃ in combination with Nb₁₄₄₇₈_DKK1C and Nb₁₄₄₇₈. **d**, Crystal structure of B12 (green, PDB 5UN5) superimposed onto crystal structure of FZD₃ (grey) in complex with Nb₁₄₄₇₈ (cyan). The positions of the C-termini of B12 and Nb₁₄₄₇₈ are denoted by dashed red circles.

Co-transfection of Nb₁₄₄₇₈ with FZD₃ modestly increased luciferase activity. One possible explanation for this is that, as it can bind to both FZD₃ and FZD₆, Nb₁₄₄₇₈ may block an interaction between endogenous FZD₆ and an endogenous β -catenin-inhibitory WNT in these cells, as FZD₆ has been demonstrated to inhibit β -catenin-dependent signalling in certain contexts (Golan *et al.*, 2004). However, when Nb₁₄₄₇₈ was co-transfected alongside FZD₃ and Nb₁₄₄₇₈-DKK1C, even basal luciferase activity was abolished. This is likely because Nb₁₄₄₇₈ can compete with Nb₁₄₄₇₈-DKK1C for FZD₃ binding and, when not bound to FZD₃, the DKK1C domain of Nb₁₄₄₇₈-DKK1C can bind to LRP5/6 in an inhibitory manner. Nb₁₄₄₇₈-DKK1C was observed to only weakly activate β -catenin-dependent signalling without co-transfection with FZD₃, which is likely due to the relatively low endogenous expression level of FZD₃ in HEK293T cells (Voloshanenko *et al.*, 2017).

Therefore, while it is unclear which WNT, if any, is capable of activating FZD₃- β -catenin signalling *in vivo*, it is clear that FZD₃ is indeed intrinsically capable of activating β -catenin-dependent signalling when heterodimerised with LRP6. The intrinsic ability of FZD₃ to activate ‘canonical’ signalling suggests that this receptor is indeed incapable of binding to WNT3A or WNT5A. FZD₃ had been predicted to bind to both of these WNTs with high affinity by a structure-based model (Agostino, Pohl and Dharmarajan, 2017), although an updated model predicted substantially weaker binding between FZD₃ and these ligands (Agostino, Öther-Gee Pohl and Dharmarajan, 2019), which is in agreement with the results presented here. However, the possibility that WNT3A/FZD₃ or WNT5A/FZD₃ signalling is dependent on an unknown auxiliary coreceptor cannot be excluded, as recent studies have identified such factors in other WNT/FZD contexts (Eubelen *et al.*, 2018; Grainger *et al.*, 2019). It will therefore be necessary to assess the affinity of FZD₃ for WNT3A and WNT5A biophysically to be able to say with confidence whether FZD₃ is a ligand for either of these WNTs.

5.4.2 The FZD₃ C-terminal tail is dispensable for activation of the ‘canonical’ pathway

Having been confirmed as an agonist of β -catenin-dependent WNT signalling, Nb₁₄₄₇₈-DKK1C can now be used as a tool to investigate the ability of variant FZD₃ proteins to activate the pathway. As such, a truncated version of FZD₃, FZD₃ 1-517, which lacks the long C-terminal tail (**Figure 5.3, a, ii**), was investigated for β -catenin stabilising activity in the presence of Nb₁₄₄₇₈-DKK1C. This protein maintains the KTXXXW motif in H8 required for DVL recruitment (Umbhauer *et al.*, 2000). This construct was able to stimulate luciferase activity at a similar level to wild-type FZD₃ (**Figure 5.3, c**), which is in agreement with a previous experiment, whereby *Xenopus* FZD₃ lacking the terminal 153 amino acids (but retaining the KTXXXW motif) maintained its ability to stimulate ‘canonical’ WNT signalling (Umbhauer *et al.*, 2000). It can therefore be concluded that the long C-terminal tail of FZD₃ is not required for β -catenin-dependent signalling activity but may be involved in a distinct function.

5.4.3 Replacement of the FZD₃ ICL3 with a BRIL fusion abolishes signalling activity

Nb₁₄₄₇₈-DKK1C was also used to probe the activity of FZD₃ with the ICL3 (residues 404-411) replaced with the BRIL apocytochrome (**Figure 5.3, a, iii**; see section 3.3). Mutations in certain ICL1/2/3 residues in the rat FZD₁ and human FZD₅ were previously shown to severely impair β -catenin stabilisation activity by preventing DVL recruitment (Cong, Schweizer and Varmus, 2004), showing the importance of this region for DVL binding. It was therefore expected that BRIL insertion into ICL3, while not replacing any of the critical residues, would abolish the β -catenin-dependent signalling activity of FZD₃ by blocking DVL binding. Indeed, BRIL_FZD₃ was unable to activate ‘canonical’ signalling when stimulated with Nb₁₄₄₇₈-DKK1C.

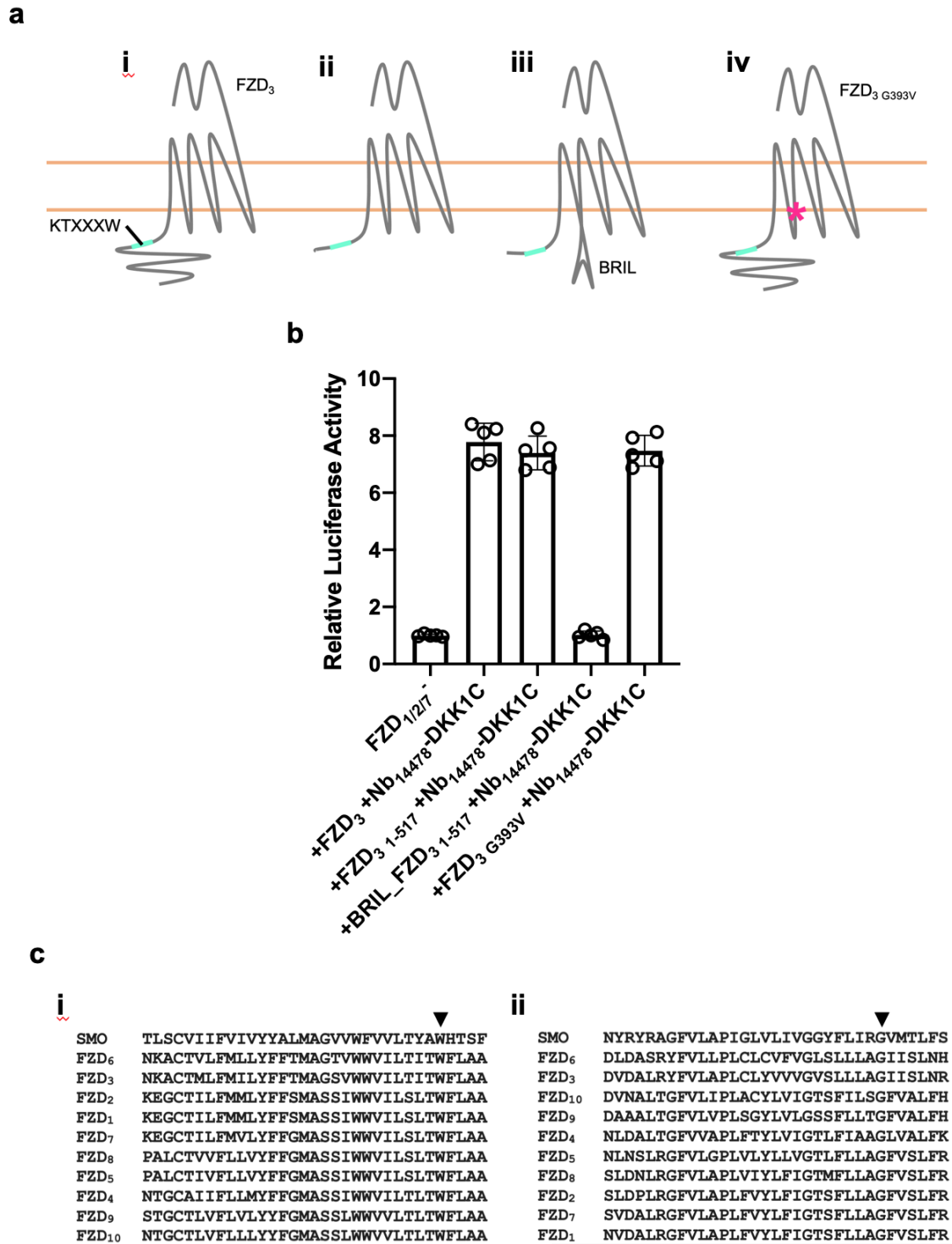


Figure 5.3 Functional dissection of FZD₃

a, Schematic diagrams of FZD₃ constructs used in functional experiments. **i**, wild-type FZD₃; **ii**, FZD₃ 1-517; **iii**, BRIL_FZD₃ 1-517; **iv**, FZD₃ G393V. **b**, Relative luciferase activity in HEK293T FZD_{1/2/7} cells transfected with NB₁₄₄₇₈_DKK1C and different FZD₃ variant proteins. **c**, Multiple sequence alignment of (i) TM3 and (ii) TM5 from all Class F GPCRs with positions of SMO W339^{3.50f} (i, and equivalent residues) and G442 (ii, and equivalent residues) indicated by arrows.

5.4.4 The mutation of a conserved glycine does not affect FZD₃ β -catenin-dependent signalling activity in response to Nb₁₄₄₇₈-DKK1C

The mutation of a conserved glycine residue in the FZD₂ TM5 to valine (**Figure 5.3, a, iv, and c, ii**) has previously been implicated in the genetic disease OMOD2 (see section 3.6.1). Nb₁₄₄₇₈-DKK1C was therefore used to investigate what effect the equivalent mutation (G393^{5.64f}V) would have on the ability of FZD₃ to activate ‘canonical’ signalling. As this mutation phenocopies other mutations in FZD₂ that result in the truncation of the H8 KTXXXW DVL-binding motif (Umbhauer *et al.*, 2000; Türkmen *et al.*, 2017; Nagasaki *et al.*, 2018), which abrogate ‘canonical’ WNT signalling, it was expected that the G393^{5.64f}V mutation would also prevent FZD₃ from triggering β -catenin stabilisation. However, when transfected alongside Nb₁₄₄₇₈-DKK1C, FZD₃ G393V was able to stimulate luciferase activity to the same extent as wild-type FZD₃, showing that mutation of this residue does not interfere with ‘canonical’ signalling, at least in FZD₃.

5.4.5 OMOD2 may be caused by aberrations in a ‘non-canonical’ FZD G protein signalling activity conserved in SMO

The observation that OMOD2 can be caused by mutations in the FZD₂ KTXXXW motif, which prevent β -catenin-dependent WNT signalling (Saal *et al.*, 2015; Nagasaki *et al.*, 2018) suggested that this disease is caused by a loss of ‘canonical’ signalling activity. However, as the glycine to valine mutation that causes the same disease (Türkmen *et al.*, 2017) did not affect the ability of FZD₃ to activate β -catenin-dependent signalling activity, it seems that the cause of autosomal dominant OMOD2 is likely to be the misregulation of a ‘non-canonical’ FZD signalling function. This suggests that the loss of β -catenin stabilisation in other FZD₂ mutants implicated in OMOD2 (Saal *et al.*, 2015; Nagasaki *et al.*, 2018) is not the cause of the disease. Indeed, FZD₂ is known to mediate both β -catenin-dependent and -independent signalling (Gujral *et al.*, 2014), and the functional redundancy between the FZD_{1/2/7} subfamily members

in terms of ‘canonical’ signalling (Voloshanenko *et al.*, 2017) may mitigate the impact of these mutations on β -catenin-dependent signalling.

The theory that OMOD2 is not caused by the loss of β -catenin stabilising activity in FZD₂ is also supported by the fact that FZD₂ G434^{5.64f} is conserved in SMO, which does not recruit DVL (Gammons *et al.*, 2016) or activate canonical WNT signalling. Were this residue important for canonical signalling, one might expect it to be variable in SMO as it would presumably be dispensable for its function. Furthermore, the stabilising effect of mutation of this residue on FZD₃ suggests this glycine is also not critical for the structural integrity of Class F GPCRs. Rather, the severe phenotype of the mutation, and conservation of this residue in SMO, hints that it is required for a shared function between SMO and FZD. As FZD and SMO both appear to signal through G proteins (Dijksterhuis, Petersen and Schulte, 2014), it is possible that the conserved glycine residue is important for G protein signalling.

5.4.6 FZD₂ G434^{5.64f}V (and equivalent mutations) may stabilise a previously unidentified, Class F GPCR-specific ‘switch’ in an ‘off’ position, preventing interaction with G proteins

In order to investigate how the FZD₂ G434^{5.64f}V mutation could affect G protein signalling, the crystal structure of the ‘inactive’ wild-type SMO TMD bound to the small molecule antagonist, LY2940680 (Wang *et al.*, 2013), was compared to the cryogenic electron microscopy (cryo-EM) structure of ‘active’ oxysterol-bound SMO in complex with a heterotrimeric G_i protein (Qi *et al.*, 2019).

In the inactive SMO structure, the equivalent residue to FZD₂ G434^{5.64f}, SMO G422^{5.64f}, is located proximal to a tryptophan, W339^{3.50f}, present on TM3, the side chain of which points towards G422^{5.64f} in a ‘horizontal’ manner in relation to the cell membrane (**Figure 5.4, a, i and ii**). This tryptophan, perfectly conserved in all Class F GPCRs (**Figure 5.3, c, i**), adopts a

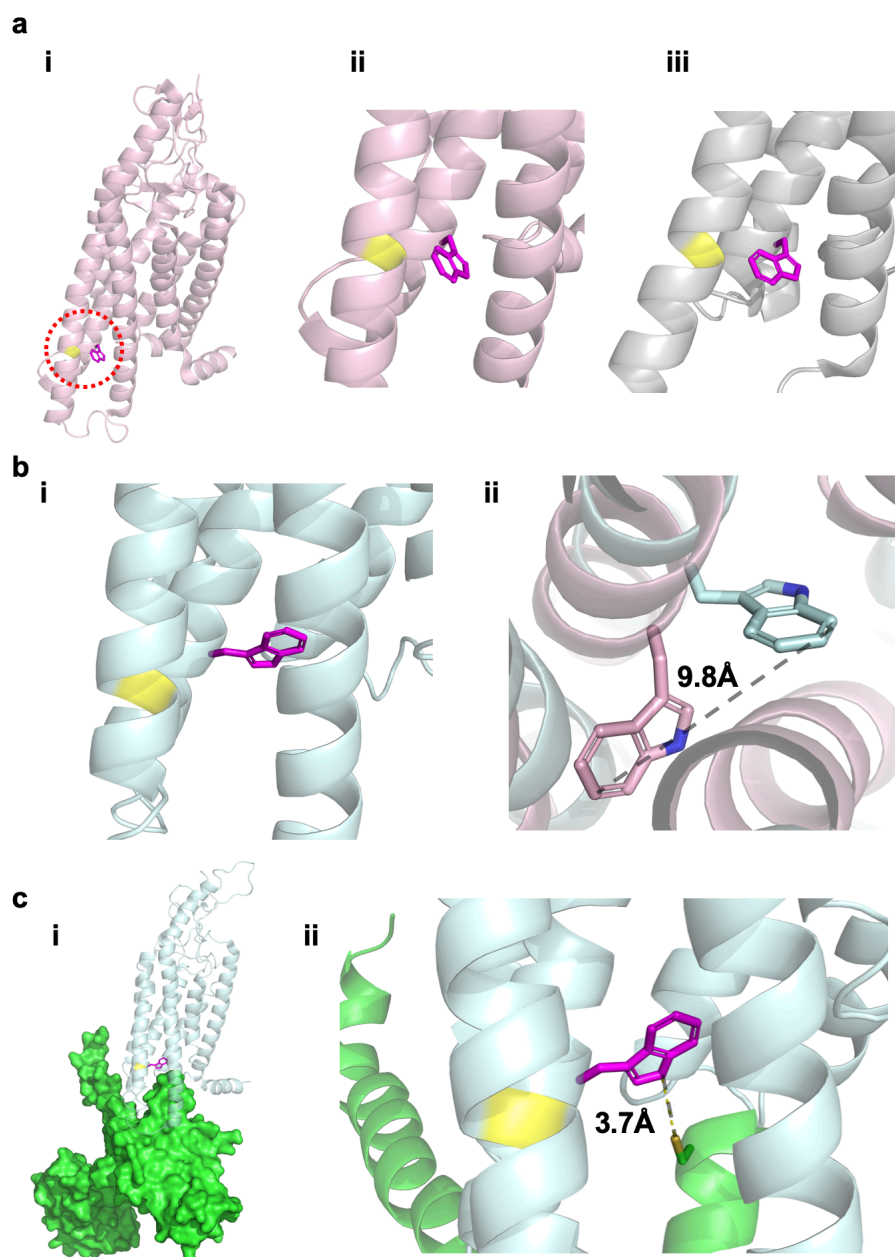


Figure 5.4 SMO G422^{5.64f} is located proximal to a conserved tryptophan, W339^{3.50f}, that appears to be involved in G protein signalling.

a, i, Crystal structure of the inactive SMO receptor TMD (PDB 4JKV), with the location of G422^{5.64f} (yellow) and W339^{3.50f} (magenta) indicated by red dashed circle. **ii**, Closeup of G422^{5.64f} and W339^{3.50f}, showing the tryptophan sidechain pointing towards the glycine. **iii**, The equivalent residues in FZD4, G409 and W327, adopt an arrangement that is highly similar to that observed in inactive SMO. **b, i**, In active SMO, W339^{3.50f} rotates upwards and away from G422^{5.64f}, into the centre of the TMD bundle, **ii**, position 6 of the W339^{3.50f} indole ring is displaced by 9.8Å in the active structure. **c, i**, Cryo-EM structure of SMO in an 'active' conformation, bound to a G α_i protein, with W339^{3.50f} (magenta) indicated. **ii**, Detail of W339^{3.50f} in active SMO, with a contact with C351 of the G α_i α 5 helix.

highly similar conformation in the FZD₄ TMD structure (**Figure 5.4 a, iii**). However, in the active SMO structure, the W339^{3.50f} sidechain undergoes a dramatic rearrangement, rotating ~90° upwards and away from G422^{5.64f}, into the centre of the TMD (**Figure 5.4, b, i**), which is in a more ‘open’ conformation upon G protein binding. Furthermore, the carbon atom at position 6 of the tryptophan indole ring is displaced by approximately 10Å (**Figure 5.4, b, ii**). In the ‘active’ structure, W339^{3.50f} also appears to form a contact with C351 of the Gα_i α5 helix that inserts into the intracellular interface of the SMO TMD (**Figure 5.4, c**). This demonstrates a direct link between G422^{5.64f} and a highly conserved amino acid involved in G protein binding in SMO.

Based on these observations, two potential theories present themselves. On the one hand, substitution of G442^{5.64f} with valine may lead to a stronger interaction between this residue and W339^{3.50f}. In this case, the inactive conformation would be stabilised and the glycine to valine mutation would inactivate G protein signalling. Alternatively, a valine at this position might prevent W339^{3.50f} from adopting the conformation observed in the inactive TMD structure, forcing it away and towards the conformation observed in the active structure. In this theory, the mutation would prevent W339^{3.50f} from adopting the inactive conformation, instead forcing it into a conformation that constitutively binds G proteins, possibly leading to higher basal activity of the receptor in the absence of a stimulus. In either case, the increased thermostability conferred on FZD₃ by the G393^{5.64f}V mutation (see section 3.6.3) indicates that conformational stabilisation of FZD₂ G434^{5.64f}V is likely to be the mechanism behind in OMOD2.

While a constitutively activating mutation could explain why FZD₂ G434^{5.64f}V is a dominant mutation (Türkmen *et al.*, 2017), one argument against this explanation is that the mutation phenocopies the truncation of the KTXXXW motif required for DVL recruitment (Umbhauer *et al.*, 2000; Saal *et al.*, 2015; Türkmen *et al.*, 2017). As DVL appears to be essential for FZD coupling to G proteins (Kilander *et al.*, 2014), the fact that these two mutations cause the same

disease suggests that the FZD₂ G434^{5.64f}V mutation may prevent FZD₂ from signalling through G proteins. A possible explanation of how a single mutated copy of FZD₂ can cause OMOD2 is that the mutated receptor would essentially function as a decoy receptor, binding to ligands but not activating G protein signalling, thereby leading to insufficiency of FZD₂ signalling activity.

Together, these observations suggest that SMO W339^{3.50f} may represent a ‘switch,’ which must be allowed to flip from an ‘off’ to an ‘on’ position to allow a larger concerted conformational change during activation of G protein signalling in Class F GPCRs. In this model, mutation of SMO G422^{5.64f} to valine leads to hydrophobic interactions between this amino acid and W339^{3.50f} that stabilise the switch in the off position. As such, the conformational rearrangement required for G protein binding is prevented and the receptor is locked in an inactive conformation.

5.5 Further evidence that the FZD₂ G434^{5.64f}V mutation affects G protein signalling

Intriguingly, the FZD₂ G434^{5.64f}V mutation implicated in OMOD2 has recently also been identified in individuals with Robinow syndrome, a disease with an overlapping phenotype (White *et al.*, 2018). This disease is also associated with mutations in other WNT signalling components, including WNT5A, ROR2 and DVL, which are all involved in WNT/PCP signalling (Seifert and Mlodzik, 2007; Yang and Mlodzik, 2015). Importantly, normal WNT/PCP signalling relies on G proteins (Katanaev *et al.*, 2005). Mutation of the same amino acid to serine (FZD₂ G434^{5.64f}S) is also implicated in Robinow syndrome (White *et al.*, 2018). In this case, it is possible that the serine could form a hydrogen bond with the indole nitrogen of W^{3.50f}, which would also result in stabilisation of FZD₂ in a conformation that does not bind G proteins. Collectively, these lines of evidence suggest that mutation of FZD₂ G434^{5.64} may cause OMOD2 or Robinow syndrome by disrupting the PCP pathway by preventing the transduction of a WNT signal to G proteins.

5.6 Conclusions and future perspectives

While a WNT that can activate β -catenin-dependent signalling via FZD₃ remains to be identified, this work proves the intrinsic ability of FZD₃ to activate this branch of WNT signalling. Nb₁₄₄₇₈-DKK1C adds to a growing list of synthetic WNT agonists (Janda *et al.*, 2017; Tao *et al.*, 2019) and becomes the first such example that acts upon the FZD_{3/6} subfamily. While this agonist can now be used in more extensive functional dissection of FZD₃, it is possible that it may find use in the derivatisation and maintenance of tissue organoids, and eventually as a therapeutic in degenerative disease, (Janda *et al.*, 2017; Tao *et al.*, 2019).

Due to time constraints, Nb₁₄₄₇₈-DKK1C was only assessed in transfection-based assays and as such, it is unclear what concentrations are required for robust β -catenin-dependent signalling. It will therefore be interesting to investigate the ability of different concentrations of purified Nb₁₄₄₇₈-DKK1C to activate signalling. However, as the affinity of Nb₁₄₄₇₈ for FZD₃ (21 nM) is highly similar to the affinity of the B12 FZD_{5/8}-binding module of a previously-described WNT agonist, B12_DKK1C for FZD₈ (18 nM), which activates β -catenin-dependent signalling at concentrations as low as 5 nM (Janda *et al.*, 2017), it is possible that Nb₁₄₄₇₈-DKK1C will exhibit similar potency.

It may also be possible to enhance the efficacy of Nb₁₄₄₇₈_DKK1C by optimising the length of the linker between the two domains, as different linker lengths between B12 and DKK1C were found to affect the activity of B12_DKK1C (Janda *et al.*, 2017). Although a five amino acid linker was chosen for Nb₁₄₄₇₈_DKK1C, as this was found to be the optimum length for B12_DKK1C activity (Janda *et al.*, 2017), B12 and Nb₁₄₄₇₈ are structurally unrelated. Furthermore, despite binding to an epitope on FZD₈ that overlaps with the equivalent epitope for Nb₁₄₄₇₈ on FZD₃, B12 possesses a long morphology that results in the C-terminus extending away from the CRD to a greater extent and in different direction relative to the C-terminus of Nb₁₄₄₇₈ (**Figure 5.3, d**). Indeed, the C-terminus of B12 resides approximately 50Å away from

the FZD₈ CRD in the crystal structure (**Figure 5.3, d**). By comparison, the C-terminus of Nb₁₄₄₇₈ is located only approximately 25Å away from the FZD₃ CRD. It is therefore unlikely that the optimal linker length connecting B12 to DKK1C would also be the optimal linker length for Nb₁₄₄₇₈_DKK1C, as this would likely result in different distances between FZD and LRP6.

It will also be interesting to determine whether co-treatment with RSPO2 or fusion of RSPO2 to the C-terminus of Nb₁₄₄₇₈-DKK1C will enhance the activity of this WNT agonist in the same manner as observed for the vantiectumab-based synthetic WNT agonist described previously (Janda *et al.*, 2017). Given that FZD₃ is strongly downregulated by RNF43, which in turn is inactivated by RSPO proteins (Koo *et al.*, 2012), it is likely that this will be the case.

This work also appears to rule out a role for ‘canonical’ WNT signalling in the pathogenesis of OMOD2, instead supporting a role for G protein signalling in this disease. Moreover, a potential molecular switch has been identified in Class F GPCRs.

Several experiments should be performed in order to test the molecular switch hypothesis. First of all, it would be very interesting to determine the effect of mutating FZD₂ G434^{5,64f} to valine, or the equivalent mutation in another Class F GPCR, on G protein signalling. It would also be interesting to determine the effect of mutating SMO W339^{3,50f} or the same residue in a FZD receptor, and assess the impact of this mutation on G protein signalling or the ability of a Class F GPCR to recruit G proteins. Furthermore, experiments should be performed to determine whether mutation of the KXXXXW motif required for DVL recruitment also prevents FZD G protein signalling. Unfortunately, assessing the activation of Class F GPCR-G protein signalling is significantly more challenging than measuring ‘canonical’ WNT signalling activity. Therefore, a Class F GPCR structure bearing the glycine to valine mutation would be incredibly valuable, as the interaction between this amino acid and the associated tryptophan would provide insights into Class F GPCR function and the mechanism underlying OMOD2.

Chapter 6 Comparisons of FZD CRDs

6.1 Introduction

At first glance, the role of the FZD CRD in WNT signalling appears very straightforward; to bind specific WNT ligands and, in the case of FZD₄, Norrin. While the basis of the specificity of Norrin binding to FZD₄ seems relatively clear, for example the positively-charged K109 in FZD₄ that forms an electrostatic interaction with Norrin is absent in all other FZDs (Chang *et al.*, 2015), how FZDs discriminate between different WNTs remains poorly understood. As the site 2 interaction between the WNT ‘finger’ and the FZD CRD has been suggested as a source of WNT-FZD interaction specificity (Janda *et al.*, 2012), comparative analyses of the nineteen human WNT fingers were performed to investigate this theory.

The role of dimerisation in FZD activation also remains unclear. Structural evidence points towards the existence of a conserved FZD CRD dimerisation mode with the two lipid-binding grooves connecting to form an extended U-shaped cavity that can be bridged by palmitoleic acid (Nile *et al.*, 2017). Furthermore, a 2:2 WNT:FZD complex structure has been published (Hirai *et al.*, 2019), indicating that such an arrangement is possible, however whether this WNT-induced dimerisation is required for signalling is unclear. While this interface has been observed for members of the FZD_{1/2/7} and FZD_{5/8} subfamilies (Nile *et al.*, 2017), structures of the FZD₄ CRD have revealed a dramatically different dimerisation mode (DeBruine *et al.*, 2017). Structural studies were therefore undertaken to determine whether the typical FZD CRD dimerisation interface is conserved in the FZD_{3/6} subfamily and to find out whether FZD_{9/10}, from the same subfamily as FZD₄, adopt the dimerisation interface observed for FZD₄ or the more typical helical interface observed in other FZDs. A structure of FZD_{9/10} was also sought in order to improve understanding of the basis of the specificity of the FZD₄:Norrin interaction.

6.2 WNTs exhibit high levels of conservation in amino acids that interact with FZD CRDs

The structures of *Xenopus* WNT8 and human WNT3 in complex with the mouse FZD₈ CRD (PDB 4F0A and 6AHY, respectively) were analysed to determine which WNT finger residues were involved in the site 2 interaction in each case. Unexpectedly, in each structure, the interacting residues are identical (**Figure 6.1, a**). The sequences of the nineteen human WNT proteins were then aligned to investigate the diversity in these residues across different WNT families (**Figure 6.1, c, i**). Surprisingly, the FZD-interacting residues are perfectly conserved in all human WNTs, with the exception of a phenylalanine, which is mutated to tyrosine in WNT11 and to valine in WNT9A/B; **Figure 6.1, c, i; Appendix 4**). Therefore, while the site 2 interaction has been suggested as a source of WNT-FZD specificity (Janda *et al.*, 2012), it is difficult to see how WNTs can differentiate between different FZD CRDs at this site. Furthermore, a positively-charged residue on the *Xenopus* WNT8 ‘thumb’ (that bears the palmitoleic acid modification that interacts with the FZD CRD hydrophobic groove), which interacts with a negatively-charged residue found on all FZDs other than FZD_{3/6} (**Figure 6.1, b**), is also found in every WNT (**Figure 6.1, c, ii**). As such, this is also unlikely to be a source of specificity.

As Norrin appears to bind to the FZD₄ CRD with approximately ten-times greater affinity in the presence of the LD (but not when the LD is replaced by the FZD₅ LD, Bang *et al.*, 2018), it is possible that the poorly-conserved linker between the FZD CRD and the TMD may be involved in the interaction with WNT. As structures of the WNT:FZD CRD complex show that the C-terminus of the FZD CRD is located close to the WNT finger, it is tempting to speculate that the LD may indeed confer WNT-binding specificity on FZDs. Outside of the CFXWCCXVXC motif (*F* is mutated in three WNTs), which interacts with the FZD CRD site 2, the WNT finger sequence is far less conserved. The FZD LD could therefore feasibly

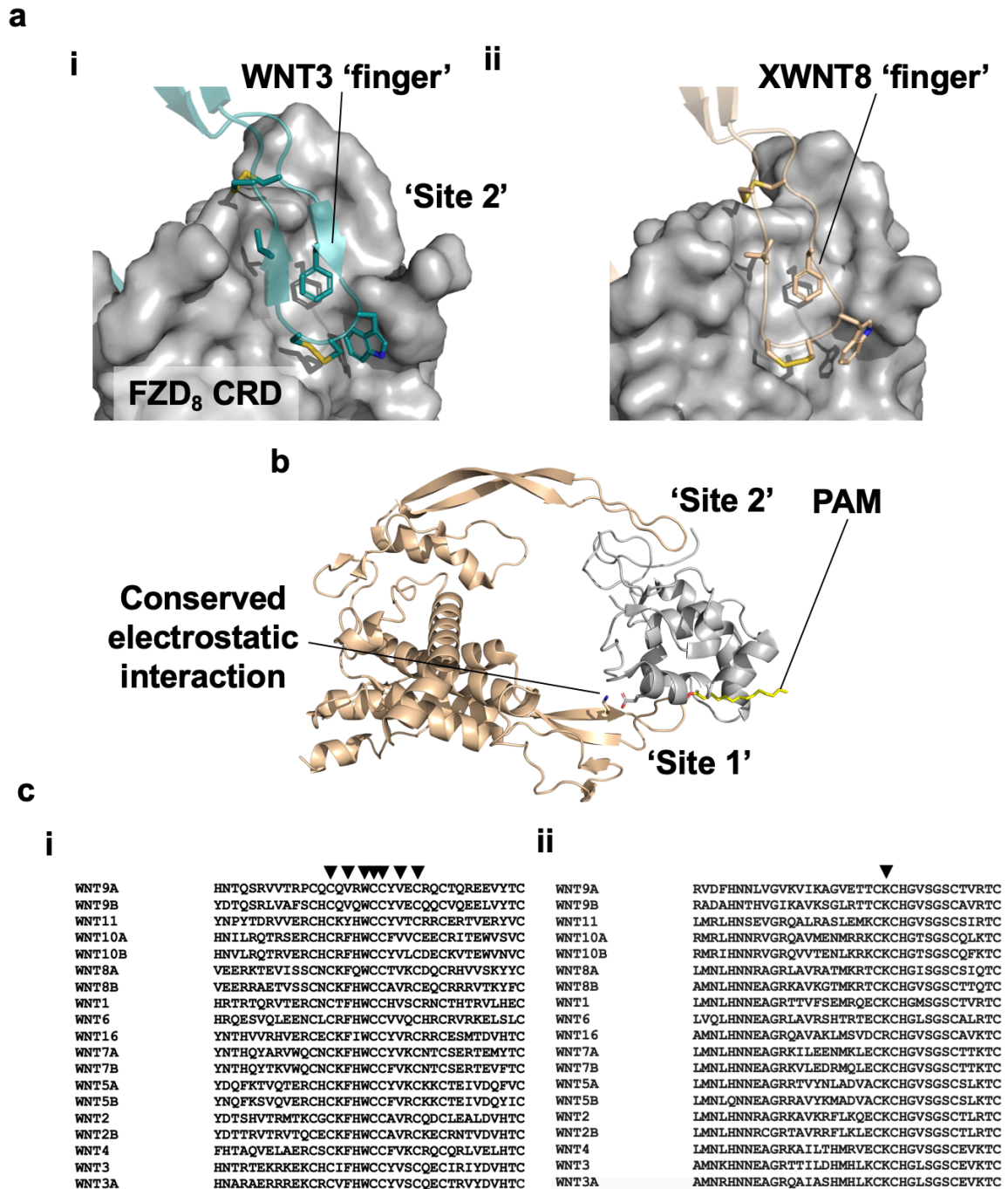


Figure 6.1 WNT-FZD interactions

a, Crystal structures of **(i)** human WNT3 (teal) and **(ii)** *Xenopus* WNT8 (wheat) in complex with the mouse FZD₈ CRD (grey, PDB 6AHY and 4F0A, respectively), focusing on the WNT finger. Residues that interact with FZD are shown as sticks. **b**, Crystal structure of *Xenopus* WNT8 in complex with the mouse FZD₈ CRD (PDB 4F0A), with XWNT8 K182 and mFZD₈ E64, which form a conserved electrostatic interaction, represented by sticks. Palmitoleic acid represented as yellow sticks **c**, **i**, Multiple sequence alignment of all human WNT ‘fingers’ with positions of amino acids that interact with FZD denoted by arrows. **ii**, Alignments of all human WNT ‘thumb’ sequences with position of the conserved positively-charged amino acid indicated by arrow.

interact with either residues in the WNT finger site 2-interacting motif (but that do not actually contact the CRD) or to residues elsewhere on the WNT finger. This is supported by the observation that a FZD missing its CRD can restore WNT signalling activity in *Drosophila* lacking any FZD expression (Chen *et al.*, 2004), indicating that WNT can interact functionally with another part of FZD, such as the LD.

6.3 Comparisons between the FZD₃ CRD and other FZD CRD structures

The FZD_{3/6} subfamily is the only group of FZDs for which no structural information has been published. As such, the FZD₃:Nb₁₄₄₇₈ structure (**Chapter 1**) completes the list of FZD subfamilies for which a CRD structure has been determined. However, as Nb₁₄₄₇₈ binds an epitope that overlaps the typical α -helical dimerisation interface, this structure cannot confirm or disprove the existence of this interface in FZD_{3/6}. As such, attempts were made to crystallise the FZD₃ CRD either alone or in complex with palmitoleic acid or the detergent n-Octyl- β -D-glucoside, which has previously crystallised with the FZD₅ CRD (Nile *et al.*, 2017). Disappointingly, no crystals grew for any of these samples. Nevertheless, the monomeric Nb-bound FZD₃ CRD structure facilitates the first structural comparisons between FZD_{3/6} and other FZDs.

6.3.1 FZD₃ possesses an usual site 2 architecture

While the overall fold of the FZD₃ CRD observed in the FZD₃ CRD:Nb₁₄₄₇₈ structure is highly similar to other FZD CRDs as expected, there are some key differences. The most notable of these is the predicted WNT ‘finger’ interaction site (site 2). The sequence of this site, between cysteines 9 and 10 of the CRD fold, diverges substantially between FZD_{3/6}, and other family members, with these two FZDs containing several amino acids fewer (**Figure 6.2, a**). Between the two cysteines, FZD₃ and FZD₆ contain only 5 residues, compared to between 11 and 13 for

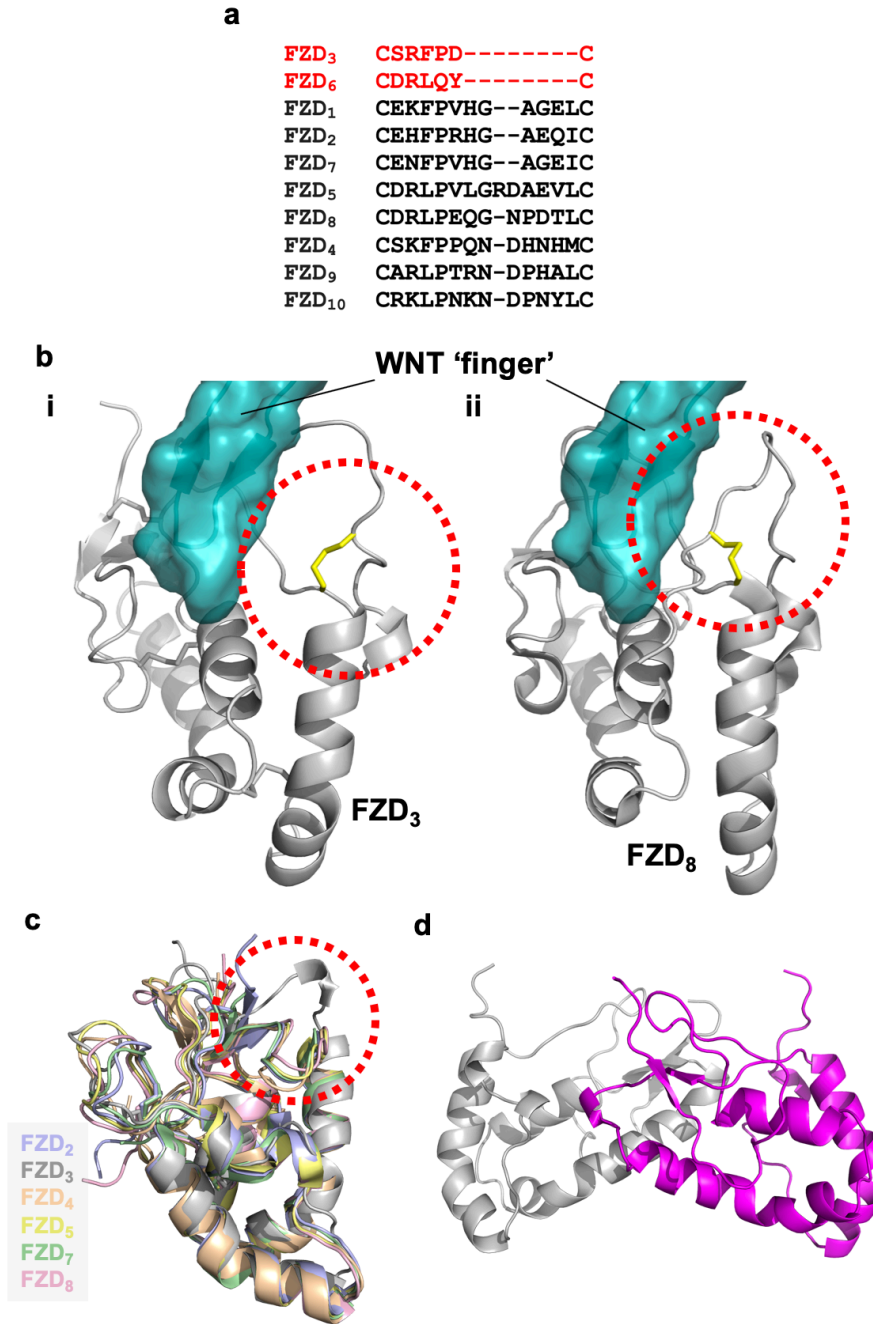


Figure 6.2 Differences between the CRDs of FZD_{3/6} and other FZDs

a, FZD_{3/6} contain a fewer amino acids between the final two cysteines of the CRD. **b**, **i**, Whereas all other FZD CRDs harbour a loop of several amino acids between the terminal two cysteines, **ii**, FZD_{3/6} have a much shorter loop, resulting in a smaller putative WNT-interacting interface. Furthermore, the sequence immediately after the CRD is consequently oriented differently in FZD_{3/6} than in other FZDs. **c**, The FZD₃ CRD crystal structure reveals an extended helix from residues that fold down against the CRD in all other FZDs. FZD₂, lavender, FZD₃, grey; PDB 6C0B; FZD₄, wheat, PDB 5CM4; FZD₅, yellow, PDB 5URY; FZD₇, pale green, PDB 5T44; FZD₈, pink, PDB 1IJY. **d**, The FZD₃ CRD interface observed in the FZD₃:Nb₁₄₄₇₈ crystal packing.

all other FZDs. This leads to the loss of a short loop in FZD₃ (and presumably FZD₆; **Figure 6.2, b, i**), which is present in all other FZD CRD structures (**Figure 6.2, b, ii**). As a result of this, the sequence following the final cysteine of the CRD in FZD₃ continues away from the CRD, whereas in the other CRD structures, the extra loop reverses the direction of this sequence, positioning it in such a way that it folds against the second and third α -helices of the CRD.

The absence of this loop in FZD_{3/6} has implications for the interaction of these CRDs with WNT. The loop, as well as the sequence immediately after the final cysteine of the CRD, forms a major part of the interaction interface between FZD₈ and WNT3/XWNT8. It is unclear whether the sequence immediately after the final cysteine of the FZD₃ CRD forms a similar structure to the loop seen in other FZDs, which could compensate for the absence of the loop, as this region was not included in the expression construct. Furthermore, this sequence could be intrinsically disordered and only become ordered upon interaction with WNT. Alternatively, this site might not be involved in the binding of WNTs to FZD₃. In either case, it is remarkable how much FZD_{3/6} differ from other FZDs at this location and this difference would likely affect the specificity and affinity of FZD₃ for different WNTs.

6.3.2 The FZD₃ CRD crystal structure reveals an extended helix not seen in other FZDs

The first α -helix of the FZD₃ CRD is significantly longer than the equivalent helix in all other reported FZD structures. The ‘extra’ residues in this helix in other FZDs comprise a loop that folds down against the helix (**Figure 6.2, c**). It is unclear whether this is an intrinsic feature of FZD₃, as this region is involved in a crystal contact in the crystal structure (**Figure 6.2, d**). Therefore, the loop that typically folds down in other FZD CRD structures may simply become ordered upon interacting with another copy of the CRD in the formation of this crystal contact. It is also unclear whether this difference in the length of the helix would contribute to the

function of FZD₃, as the equivalent residues in FZD₈ are not involved in the interaction with WNT3 or XWNT8. Similarly, when the FZD₃ CRD is aligned to the structure of the FZD₈-WNT3 complex, the extension of the helix does not come into close proximity with the nearby WNT3 ‘thumb’ (**Figure 6.3, a**) It is however worth noting that this interface buries a similar surface area ($\sim 730 \text{ \AA}^2$) to other typical FZD CRD dimerisation interfaces (e.g. $\sim 760 \text{ \AA}^2$ in FZD₇, PDB 5T44).

This crystal contact observed in the FZD₃:Nb₁₄₄₇₈ crystal structure resembles a typical FZD CRD dimerisation interface that has been pulled apart at the ‘bottom’ near the hydrophobic groove, leading to a crossing over at the ‘top.’ This is presumably a consequence of Nb₁₄₄₇₈ binding to a site that overlaps with the typical dimerisation interface, preventing the CRD from dimerising in the typical manner. This interface is not likely to be relevant to FZD activation as the extended U-shaped cavity that forms from two connected FZD hydrophobic grooves (Nile *et al.*, 2017) is lost in this dimer configuration, and WNT binding would not be possible due to clashes (**Figure 6.3, b**).

6.4 The crystal structure of the FZD₁₀ CRD compared to other FZD CRD structures

6.4.1 Crystal structure of the FZD₁₀ CRD

FZD₄ is the only FZD_{4/9/10} subfamily member for which CRD structures are published (Chang *et al.*, 2015; Shen *et al.*, 2015; DeBruine *et al.*, 2017). Furthermore, FZD₄ is the only FZD for which Norrin is known to be a ligand (Smallwood *et al.*, 2007), and has a seemingly unique dimerisation interface (DeBruine *et al.*, 2017). Efforts were therefore undertaken to determine a crystal structure of the FZD_{9/10} CRD, to confirm the basis of specificity for Norrin binding and to determine whether the unusual dimer interface of FZD₄ is specific to this FZD or shared with other subfamily members.

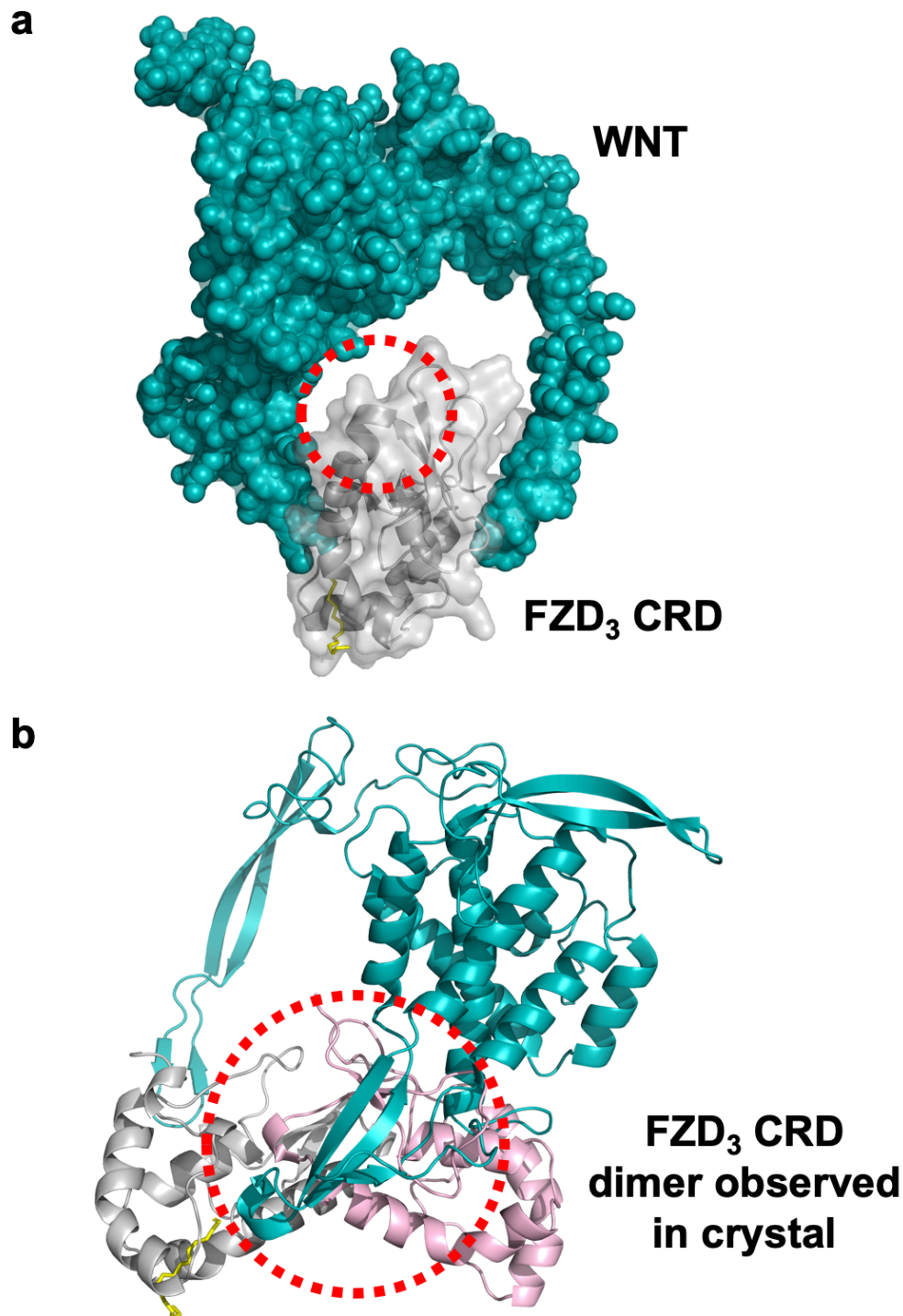


Figure 6.3 The implications of the FZD₃:FZD₃ interface observed in the FZD₃:Nb₁₄₄₇₈ crystal packing on WNT binding

a, When superimposed on the crystal structure of mouse FZD₈ CRD (hidden) in complex with WNT3 (teal, PDB 6AHY), the extended helix observed in the FZD₃ CRD crystal structure does not appear to be capable of interacting with WNT. **b**, When the same structure is superimposed on the FZD₃ dimer configuration observed in the crystal packing, WNT3 clashes extensively with the second FZD₃ CRD monomer (light pink).

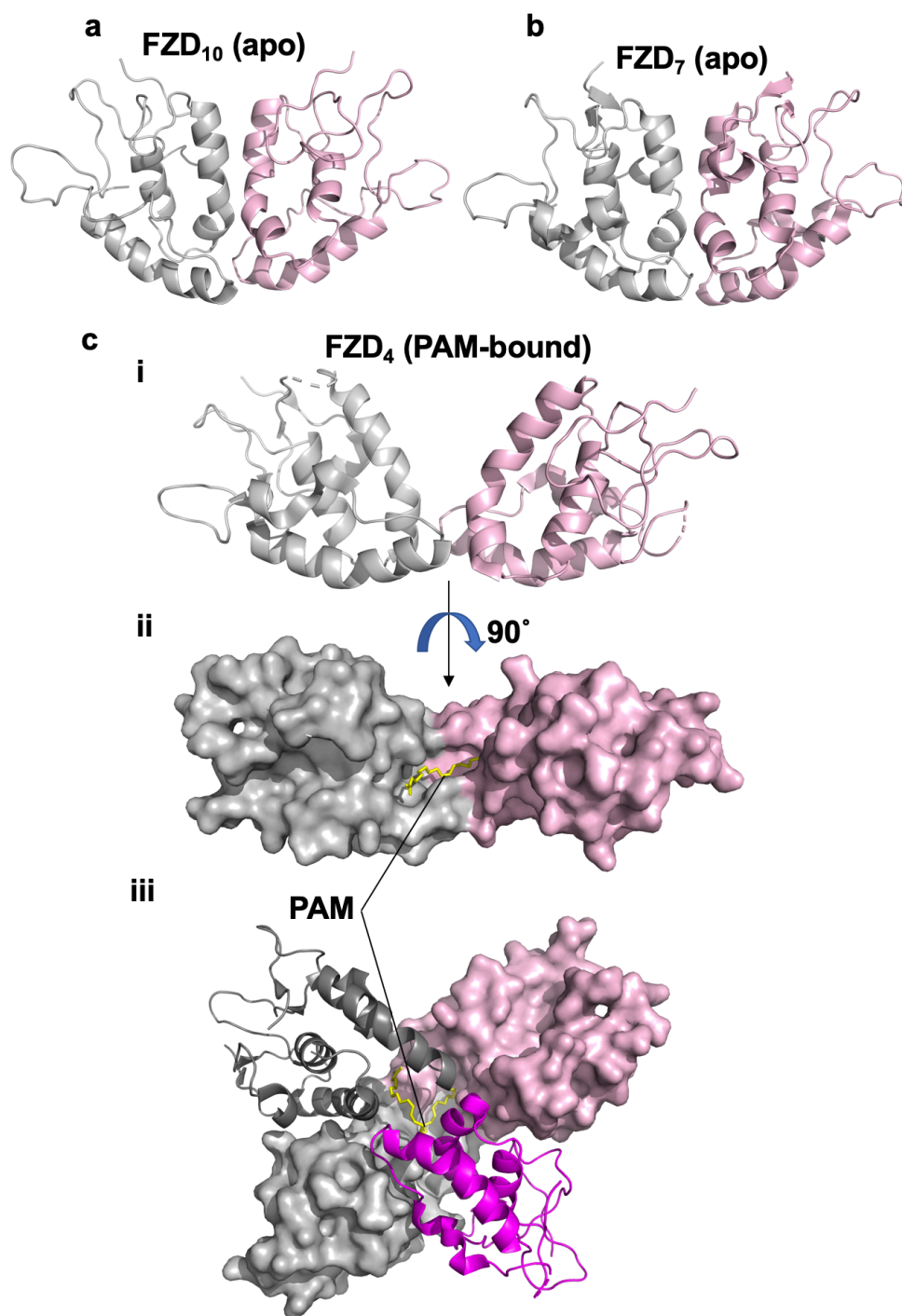


Figure 6.4 Different FZD CRD dimer configurations

a, Dimerisation interface observed in apo FZD₁₀ CRD structure, **b**, apo FZD₇ dimerisation interface (PDB 5T44). **c**, **i**, palmitoleic acid-bound FZD₄ dimerisation interface (PDB 5UWG), **ii**, the unusual dimerisation mode observed in FZD₄ (surface depiction) leaves palmitoleic acid (yellow sticks) substantially exposed to solvent, **iii**, FZD₄ CRD tetramerisation interface observed in crystal packing.

Table 3 Crystallographic data collection and refinement statistics for FZD₁₀ CRD

	FZD₁₀ CRD
Data collection	
Space group	<i>P</i> 6 ₅
Unit cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	66.3 66.3 321.2
<i>α</i> , <i>β</i> , <i>γ</i> (°)	90.0 90.0 120.0
Wavelength	0.98 Å
Resolution (Å)	56.54 - 2.77 (2.87 - 2.77)
R _{merge}	0.130 (>1)
R _{pim}	0.042(0.761)
CC(1/2)	0.999 (0.885)
Mean ((I)/σ(I))	8.7 (1.0)
Completeness	99.8 (99.0)
Redundancy	10.4 (10.7)
Refinement	
Resolution (Å)	56.54 - 2.77 (2.87 - 2.77)
Total reflections	212321 (10857)
Unique reflections	20338 (1010)
R _{work}	0.2369
R _{free}	0.2943
Number of atoms	5881
Protein	5880
Ligand/ion	0
Water	0
B factors (Å ²)	
Protein	122.37
Ligand/ion	N/A
Water	N/A
RMS deviations	
Bond lengths (Å)	0.014
Bond angles (°)	1.93
Ramachandran favoured	87.50
Ramachandran outliers	1.65

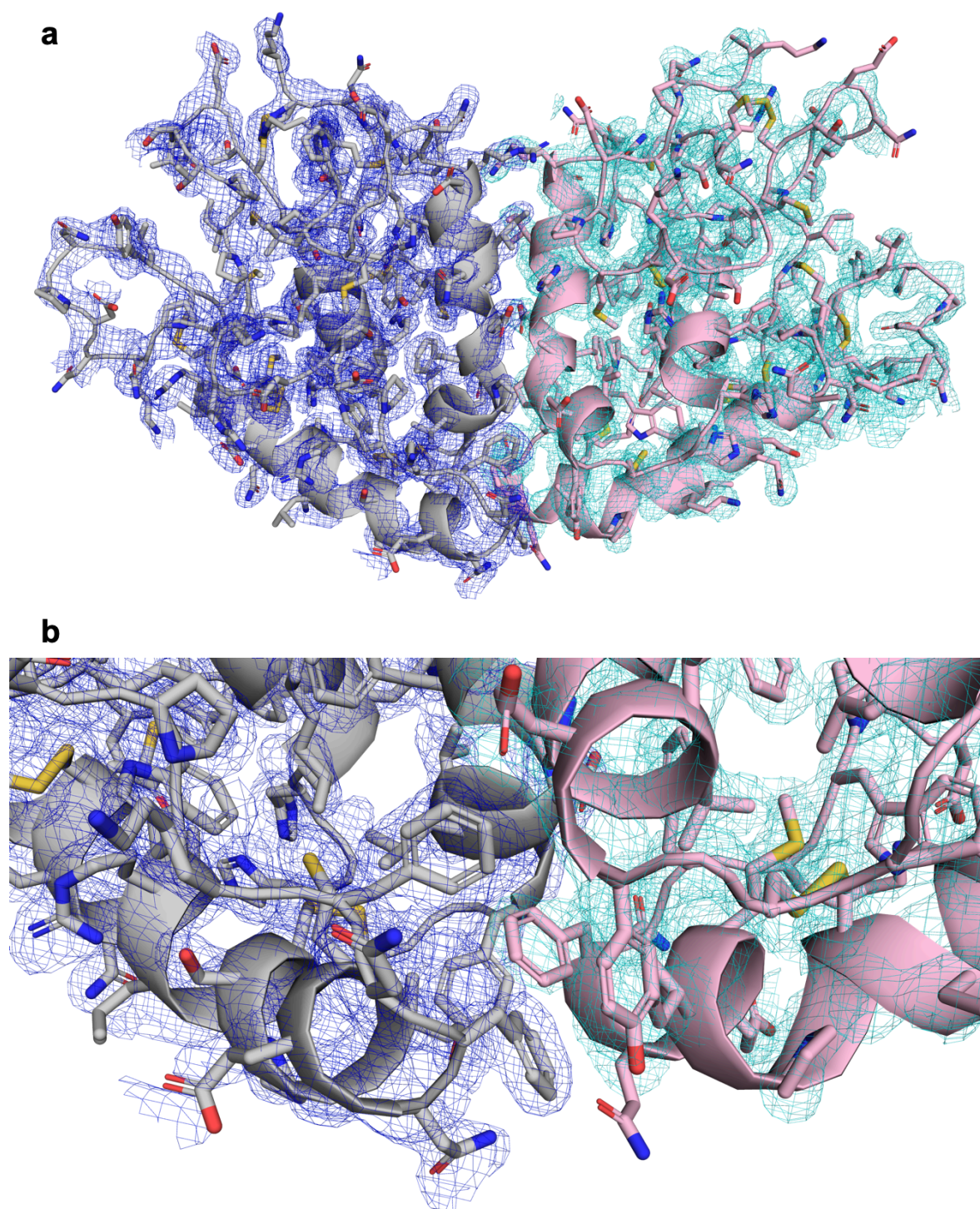


Figure 6.5 Electron density map of FZD₁₀ CRD structure

The final electron density was calculated using SigmaA m2Fo-DFc coefficients as calculated by the program SigmaA (Read, 1986) and contoured at $\sigma = 1.0$. FZD₁₀ CRD structure displayed as cartoon with side chains shown as sticks. **a**, Overall structure; **b**, Detail of dimerisation interface.

The structure of the CRD of FZD10 was solved to 2.4 Å resolution (**Figure 6.4, a; Figure 6.5**), revealing the characteristic CRD fold, stabilised by five intramolecular disulphide bonds (data collection and refinement statistics are reported in Table 3). Despite the addition of a molar excess of palmitoleic acid to the sample, no electron density is observed within the lipid-binding groove and the structure appears to be in an apo state. The α -helical dimerisation interface, which can be bridged by palmitoleic acid, is apparent. The loop-loop interface is also visible. Overall, the FZD₁₀ CRD is highly similar to other published FZD CRD structures, for example when the apo FZD₇ CRD structure is superimposed, the average root mean square deviation (RMSD) over 198 equivalent C α atoms is 1.32 Å.

6.4.2 The FZD₁₀ dimerisation interface is more similar to FZD_{5, 7}, and 8 than FZD₄

Despite being more closely related to FZD₄, the FZD₁₀ CRD adopts the typical α -helical dimerisation mode observed in FZD_{5, 7} and 8 (**Figure 6.4, a and b**), which gives rise to a U-shaped cavity formed from the two lipid-binding grooves, rather than the FZD₄ dimer configuration (**Figure 6.4, c, i**), which lacks the helical interface and U-shaped cavity, instead forming a relatively straight trough shape, which is occupied by palmitoleic acid (**Figure 6.4, c, ii**). The FZD₄ CRD dimer configuration leaves the palmitoleic acid somewhat exposed to solvent, which does not occur in the more typical helical dimerisation mode.

The FZD₄ CRD consistently crystallises in a tetrameric ‘dimer of dimers’ configuration, with the two dimers crossing over one another (**Figure 6.4, c, iii**; DeBruine *et al.*, 2017), although the biological relevance of this is controversial given the apparent clashes that would occur upon WNT binding to this assembly (DeBruine *et al.*, 2017; Nile and Hannoush, 2019). Furthermore, in this configuration, one dimer is ‘upside down’ relative to the other and it is therefore unclear whether this arrangement could actually be adopted by four FZDs within the same plasma membrane. Therefore, it cannot be excluded that the unusual dimer configuration

observed in FZD₄ is an artefact of crystallisation, caused by a tetramerisation that may not occur *in vivo*, which pushes the two monomers of each dimer apart.

It is also unclear why palmitoleic acid is not observed in the FZD₁₀ CRD structure, despite being included at a 1.5 times molar excess. Furthermore, HEK293S cells were used for production of both FZD₄ and FZD₁₀ CRDs for crystallisation with palmitoleic acid and several of the amino acids involved in the interaction between the FZD₄ CRD and palmitoleic acid, such as F82, F135 and F137, are conserved in FZD₁₀. As there is evidence for fatty acid binding in all published FZD₄ CRD structures (excluding Norrin complexes, DeBruine *et al.*, 2017; Nile *et al.*, 2017), including those where exogenous palmitoleic acid was not added to the protein during sample preparation (Chang *et al.*, 2015; Shen *et al.*, 2015), it is surprising that FZD₁₀ crystallised in an apo form even in the presence of palmitoleic acid. This raises the possibility that even closely related FZD CRDs may vary in their propensity to bind palmitoleic acid.

6.4.3 Why does Norrin bind to FZD₄ but not FZD₁₀?

The FZD₁₀ CRD shares 50% sequence identity with that of FZD₄, however, only the latter is a receptor for Norrin (**Figure 6.6, a**; Smallwood *et al.*, 2007). The specificity of Norrin binding to FZD₄ has previously been attributed to electrostatic interactions between FZD₄ K109 and Norrin D46 (**Figure 6.6, b i**); and FZD₄ E160 and Norrin K102 and K104 (**Figure 6.6, c, i**; Chang *et al.*, 2015). FZD₄ K109 is not found in other FZDs and superimposition of the FZD₁₀ CRD onto the FZD₄ CRD:Norrin structure confirms that this electrostatic interaction would not occur with FZD₁₀ (**Figure 6.6, b, ii**). On the other hand, FZD₄ E160, which forms an electrostatic interaction with Norrin K102, is conserved within the FZD_{4/9/10} (**Figure 6.6, c, ii**). This amino acid is also found in FZD_{3/6} and is replaced by another negatively-charged amino acid, aspartate, in FZD_{5/8}. The FZD₁₀ CRD crystal structure confirms this residue occupies a

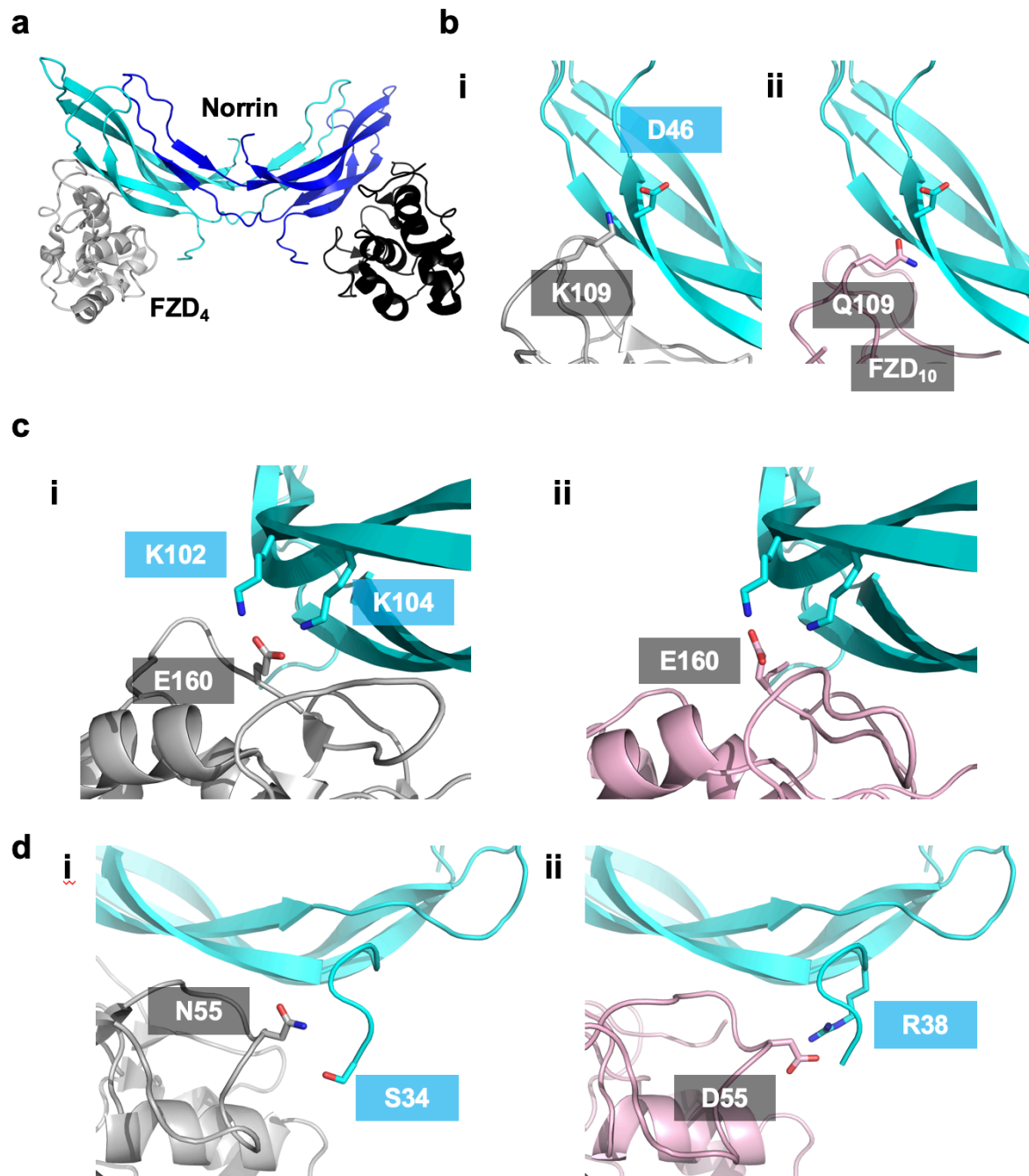


Figure 6.6 The FZD₁₀ CRD structure strengthens understanding of the basis of Norrin:FZD₄ specificity

a, Dimeric Norrin binds to the FZD₄ CRD in a 2:2 complex (PDB 5BQC). b, i, FZD₄ K109 forms an electrostatic interaction with Norrin D46. ii, when the FZD₁₀ CRD structure is superimposed onto the FZD₄:Norrin structure, K109 is replaced by glutamine, so the electrostatic interaction is lost. c, i, FZD₄ E160 forms an electrostatic interaction with Norrin K102. ii, E160 is conserved in FZD₁₀, meaning this interaction would be maintained. d, i, FZD₄ N55 forms a hydrogen bond with Norrin S34. ii, N55 is replaced by aspartic acid in FZD₁₀. However, while the hydrogen bond is lost, FZD₁₀ D55 appears perfectly placed to form an electrostatic interaction with Norrin R38.

similar position in both FZD₄ and FZD₁₀ and is therefore less likely to be a key determinant of the specificity of Norrin binding to FZD₄. A hydrogen bond between FZD₄ N55 and Norrin S34 has also been suggested to contribute to the specificity of this interaction (**Figure 6.6, d, i**; Chang *et al.*, 2015). As with K109, N55 is seen only in FZD₄. The FZD₁₀ CRD indeed lacks an amino acid that would form a hydrogen bond with Norrin S34 at this position. However, N55 is replaced in FZD₁₀ by D55, which appears to be well-placed to form an electrostatic interaction with Norrin R38 (**Figure 6.6, d, ii**). This aspartic acid is also found in FZD_{1/2/7}, as well as FZD₃. It is therefore a possibility that FZD₄ N55 makes a lesser contribution to FZD₄-Norrin binding specificity, as another interaction at this position could compensate for the lack of N55 in some FZDs.

6.5 Conclusions and future perspectives

The extremely high level of conservation of the WNT finger amino acids that are involved in contacting FZD CRDs at site 2 appears to rule out a role for this interaction in achieving WNT-FZD specificity. As the site 1 interaction is predominantly mediated by the invariant WNT palmitoleic acid, this is also unlikely to be a source of specificity. This suggests that interactions between WNT and the FZD CRD alone are unlikely to determine the specificity of WNT-FZD binding. As the FZD₄ LD appears to contribute to FZD₄-Norrin binding specificity (Bang *et al.*, 2018), it is possible that interactions between the FZD LD and the WNT finger may contribute to WNT-FZD specificity, which could explain why FZDs possess longer LDs than SMO, which does not bind WNTs (**Appendix 1**). In order to assess whether the FZD LD contributes to WNT binding, experiments should be performed to measure the affinity for specific WNT-FZD CRD pairs in the presence and absence of the LD. If the presence of the LD affects the affinities, then cellular assays can be performed to determine whether swapping the LDs of different FZDs can change which FZDs are activated by a particular WNT.

The observation that synthetic WNT agonists, such as Nb₁₄₄₇₈_DKK1C, which theoretically function by bringing together a single FZD with a single copy of LRP5/6 to trigger β -catenin stabilisation (Janda *et al.*, 2017, see section 1.11), suggests dimerisation of the FZD CRD is dispensable, at least for ‘canonical’ signalling. Furthermore, Nb₁₄₄₇₈ and the FZD-binding B12 component of the B12_DKK1C both bind to epitopes that overlap with the CRD dimerisation interface and therefore their agonistic activities are mutually exclusive with dimerisation of the FZD CRD at the helical interface. However, it is possible that FZDs may exist in the cell membrane as inactive dimers connected via an interface within the TMD (Petersen *et al.*, 2017), meaning WNTs (or synthetic agonists) could bring a FZD pair dimerised via the TMD into proximity with LRP5/6 independently of CRD dimerisation. If this is the case, then it raises the question of why a CRD dimerisation interface exists. One possible explanation is that WNT binding to a dimeric CRD is more favourable as the hydrophobic palmitoleic is better shielded from solution by a CRD dimer than a CRD monomer, as without a second CRD groove to occupy the terminal carbons of the hydrophobic lipid extend into solution (**Figure 6.1, a, 2**; Janda *et al.*, 2012).

The structures of the FZD₃ and FZD₁₀ CRDs presented here complement previously published structures to provide a more complete structural understanding of FZD CRDs. The first structure of a FZD_{3/6} protein reveals an unusual architecture of the site 2 WNT binding site. Meanwhile, the FZD₁₀ CRD is structurally highly similar to the FZD_{1/2/7} and _{5/8} CRDs, exhibiting the conserved helical dimerisation interface, rather than the unusual interface observed in the more closely-related FZD₄ CRD. This structure also provides a clearer insight into the basis of the specificity of the interaction between Norrin and FZD₄, with the electrostatic interaction between FZD₄ K109 and Norrin D46 now seeming to be the primary determinant.

Work should be undertaken to determine which WNTs can bind to the FZD_{3/6} CRD, as this information can be used in conjunction with the FZD₃ CRD structure to provide insights into how the structural differences in these FZDs affect WNT binding. It would also be very interesting to mutate residues at the helical dimerisation interface in a FZD CRD to prevent dimerisation and observe what effect this has on β -catenin-dependent WNT signalling.

Conclusion

This project sought to determine the first multidomain FZD receptor structure. While this still remains to be achieved, many important advances have been made towards this goal. Firstly, a range of construct boundaries were screened and replacement of ICL3 with BRIL was found to boost FZD expression in most cases. After identifying FZD₃ as a suitable candidate for structural studies, FZD₃ expression and thermostability was further boosted via the introduction of point mutations into the TMD. FZD₃ production was further optimised by the generation of lentivirus-transduced cell lines and sample heterogeneity was overcome with the use of an N-terminal affinity tag. Collectively, these optimisations allowed crystals of multidomain FZD₃ to be grown, although these did not diffract x-rays. Despite this, the ability to produce milligram quantities of a homogeneous, thermostabilised FZD₃ sample will be vital for future efforts to determine a multidomain FZD structure. Moreover, the discovery that the mutation of the perfectly conserved FZD₃ G393^{5.64f} to valine confers improved expression and thermostability on FZD₃ raises the possibility that this is a strategy that can potentially be applied to any Class F GPCR to increase its expression and stability.

Production of large quantities of homogeneous FZD₃ protein also facilitated the generation of anti-FZD₃ Nbs, leading to the identification of a Nb, Nb₁₄₄₇₈, which binds with high affinity and specificity to the FZD_{3/6} CRD. Although initial attempts to crystallise multidomain FZD₃ with this Nb were unsuccessful, a high resolution FZD₃ CRD:Nb₁₄₄₇₈ structure was determined, facilitating structural analyses of a FZD from this subfamily for the first time. This structure also hints at a likely role of Nb₁₄₄₇₈ in inhibiting FZD₃ by preventing WNT binding via narrowing and blocking the lipid-binding groove. Conversely, fusion of Nb₁₄₄₇₈ to the LRP-binding domain of DKK1 transforms this Nb into a potent agonist of FZD₃- β -catenin signalling, which was successfully exploited to assess the ability of different FZD₃ variants to

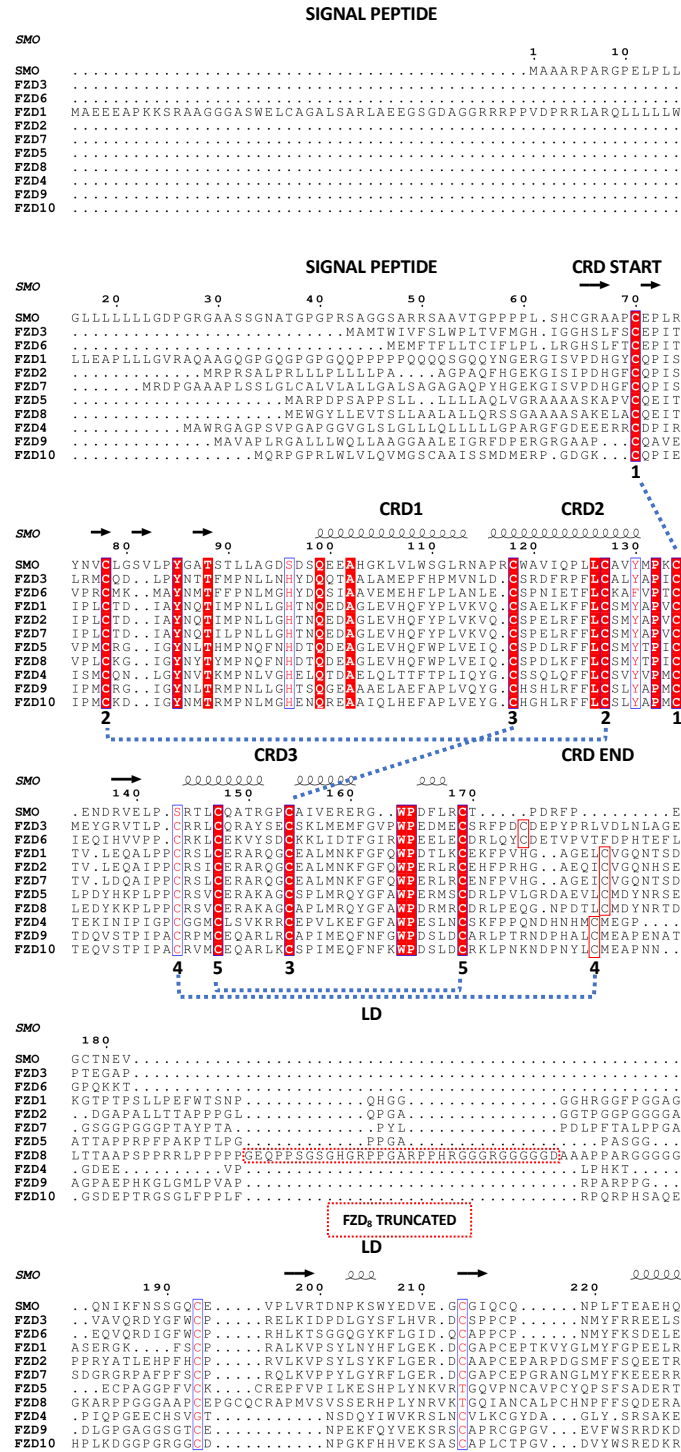
activate signalling. Nb₁₄₄₇₈ may therefore find use as both a modulator of WNT signalling in cells expressing FZD₃, as well as a tool to investigate FZD function. Meanwhile, FZD₃ was found to not be a receptor for WNT3A or WNT5A in HEK cells. Identifying the physiological ligands of this receptor will be important for improving the understanding of its function *in vivo*.

Finally, in addition to the FZD₃ CRD structure completing the list of FZD subfamilies for which CRD structures have been solved, the FZD₁₀ CRD structure represents the first CRD structure for a FZD_{4/9/10} subfamily member that does not bind Norrin. The FZD₃ CRD structure revealed a WNT binding site 2 architecture that has not been observed for any other FZD CRD, which may explain why it does not bind to either WNT3A or WNT5A, two so-called ‘promiscuous’ FZD ligands. On the other hand, the FZD₁₀ CRD structure confirmed that the typical FZD CRD dimerisation interface is conserved in the FZD_{4/9/10} family and the FZD_{3/6} subfamily is now the only group of FZDs that this interface is yet to be observed in. This structure also strengthened the understanding of why Norrin binds specifically to FZD₄ and no other FZD receptor.

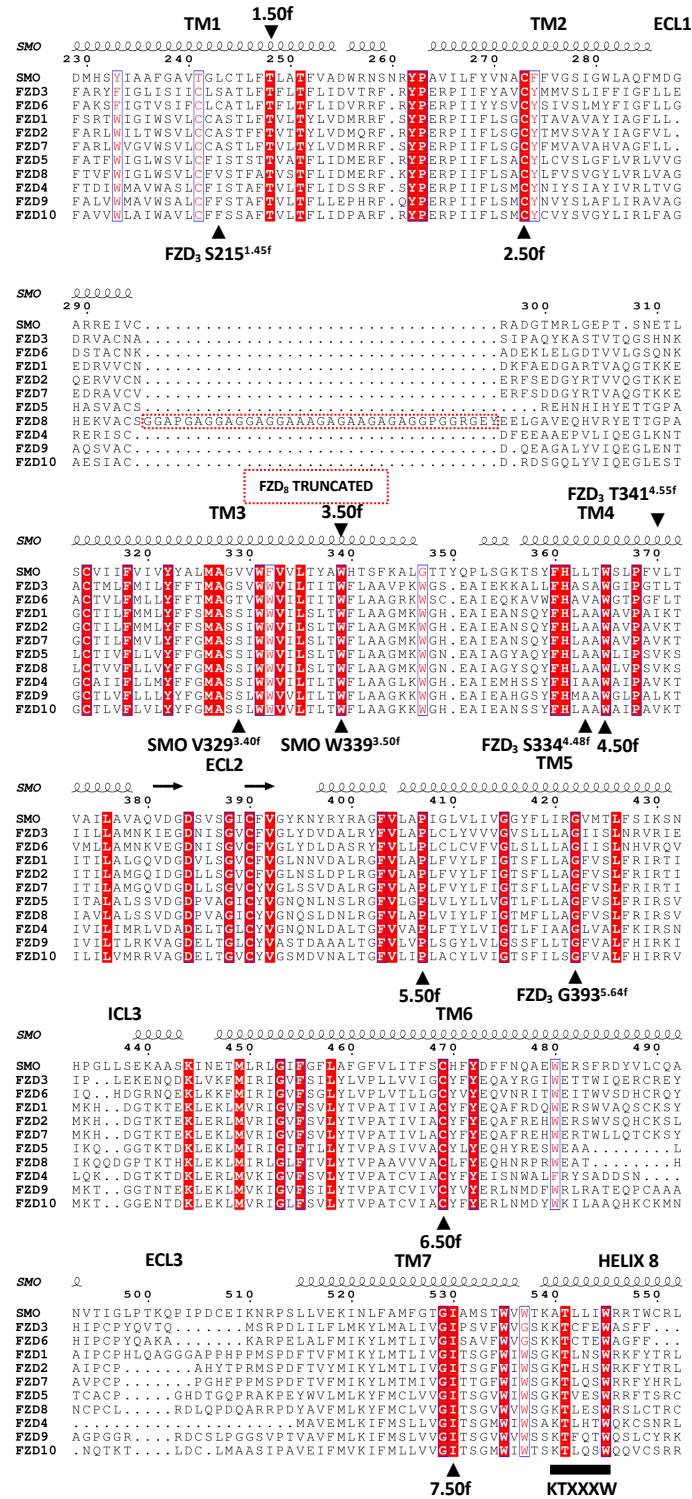
In summary, these results reveal exciting new clues about the function of FZD receptors and provide an important foundation for future structural studies to build on.

Appendices

Appendix 1 Amino acid sequence alignment of human Class F GPCRs



Appendix 1 Amino acid sequence alignment of human Class F GPCRs



Appendix 1 Amino acid sequence alignment of human Class F GPCRs

C-TERMINAL TAIL						
SMO	560	570	580	590	600	610
SMO	TGQSDDEP	KRIKKSKMIAKAF	SKRHLLQNP	QQLSFSMHTV	SHDGPVAGL	AFDLNEPSA
FZD3		HGRKKKIIVNESR	OVLPQEPD	..FAQSLLR		DPNTPIRKSRG
FZD6		KRNRRKRDPISE	SRVLQESCE	FFLKHN	SKVKHKKHYP	SSSHKLKVISK
FZD1		TNSKQGETTV				
FZD2		TNSRHGETTV				
FZD7		SHSSKGETAV				
FZD5		CCRPRRGHKSGG	AMAA.....	GDYFEASA	ALTGRTGPP	GPAAATYHKQ
FZD8		CWASKGAAVCGG	AGATAAGGGG	GGGGGGGGG	GGGGGGG	GLYSDVSTGL
FZD4		VNSCKVKREKRN	GNWVKGKGE	TVV.....		
FZD9		IAAGR.ARAKAC	RAPGSYGRG	THCHYKAPT	VVLHMTKT	DPSPLENPTH
FZD10		LKK.K.SRRKP	PASVITSGGI	YKKAQHPQ	KT...HHG	KYEIPAQSPT

SMO	620	630	640	650	660	670
SMO	DVSSAWAQHV	TKMVARRGAIL	PQDISVTP	VATVPVP	PEEQANLWL	VEAEISP
FZD3		TSQTGSTHAS	TQL...AMV	DDQSKAGS	IHSK.....	VSSYHGS
FZD6		TSTGATANH	GTSAVA...IT	SHDYLGG	ETLTIQ.....	TSPETSM
FZD1						
FZD2						
FZD7						
FZD5						
FZD8		SSVSYPKQ	MPLSQV.....			
FZD4						
FZD9						
FZD10						

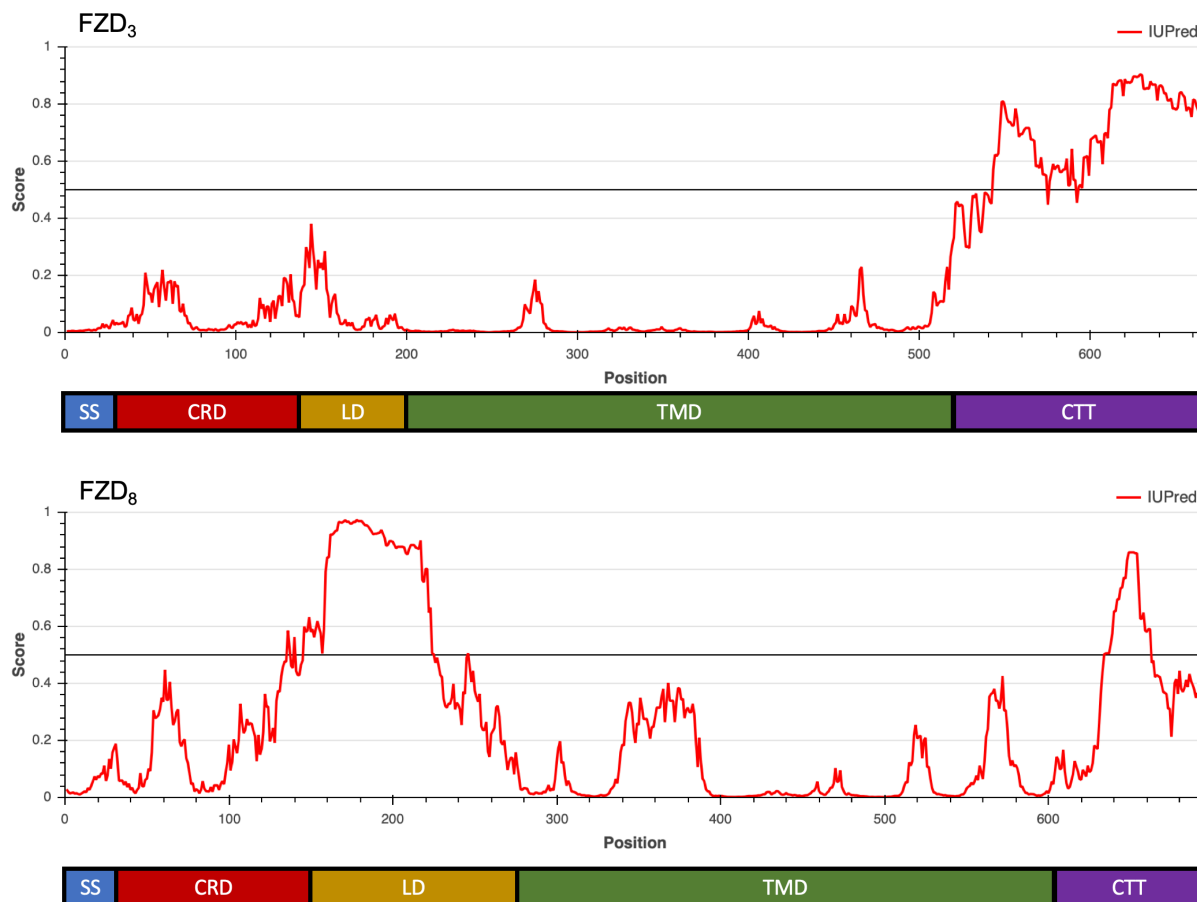
SMO	680	690	700	710	720	730
SMO	KRRRKRKKE	VCPLAPPEL	HPPAPAP	STIPRLP	QLPRQKCL	VAAAGAWG
FZD3		TPCSYRGM	..EERLPHG	SMR.LTDH	SRHSSSH	RLN.EQSR
FZD6		TPRLREQD	CGEPASPA	ASISR.LS	GEQVDGK	QAGSVSE
FZD1						
FZD2						
FZD7						
FZD5						
FZD8						
FZD4						
FZD9						
FZD10						

SMO	740	750	760	770	780
SMO	VSNFFCPEP	SPPPQDP	FLPSAPAP	VAWAHGR	RQGLGPIH
FZD3		SMNRVIEE	DGTSA.....		
FZD6		SNNLQVP	SSSEPSS	LKGSTSL	LVHPVSG
FZD1					
FZD2					
FZD7					
FZD5					
FZD8					
FZD4					
FZD9					
FZD10					

Appendix 1

Amino acid sequence alignment of all human Class F GPCRs. Alignment of human SMO (Uniprot ID Q99835); FZD₁ (Q9UP38); FZD₂ (Q14332); FZD₃ (Q9NPG1); FZD₄ (Q9ULV1); FZD₅ (Q13467); FZD₆ (O60353); FZD₇ (O75084); FZD₈ (Q9H461); FZD₉ (O00144); and FZD₁₀ (Q99835). Secondary structure elements, as well as important amino acid positions, CRD disulphide bonds (1-5, dashed blue lines), and construct modifications discussed in this thesis, are annotated. Sequences were aligned using MultAlin (Corpet, 1988) and formatted using Espritt (Gouet *et al.*, 1999).

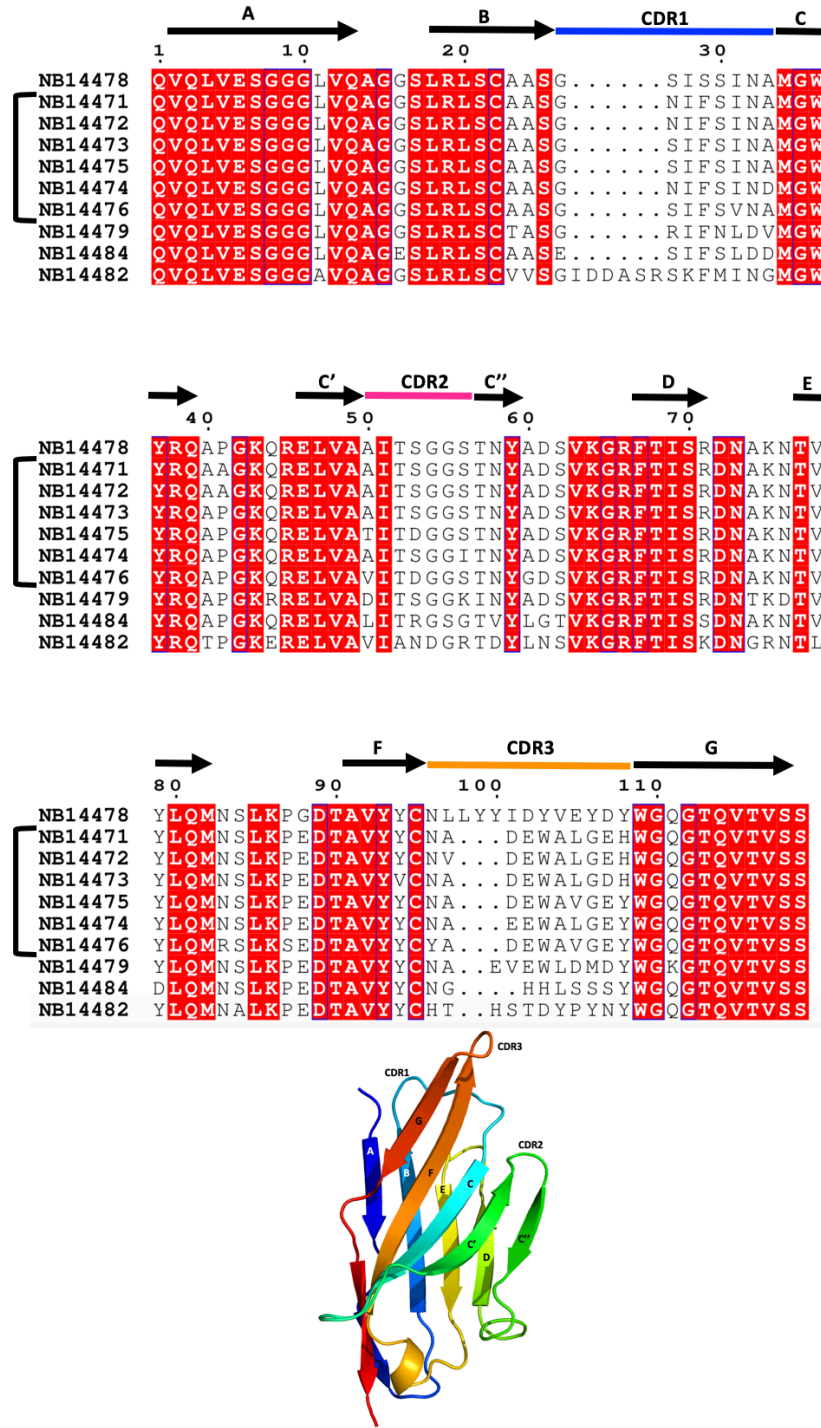
Appendix 2 Disorder analysis of human FZD₃ and FZD₈ amino acid sequences



Appendix 2

Analysis of the amino acid sequences of two representative human FZD receptors (FZD₃ and FZD₈) using the disorder prediction server IUPred2A (Mészáros, Erdős and Dosztányi, 2018). The predicted positions of the domains of the receptors are indicated. The LD and C-terminal tail (CTT) are predicted to be highly disordered in all FZDs, however FZD₈ exhibits two longer regions with greater predicted disorder, one in the LD and one in the ICL1, whereas FZD_{3/6} possess longer disordered C-terminal tails.

Appendix 3 Amino acid sequence alignment of anti-FZD₃ Nbs



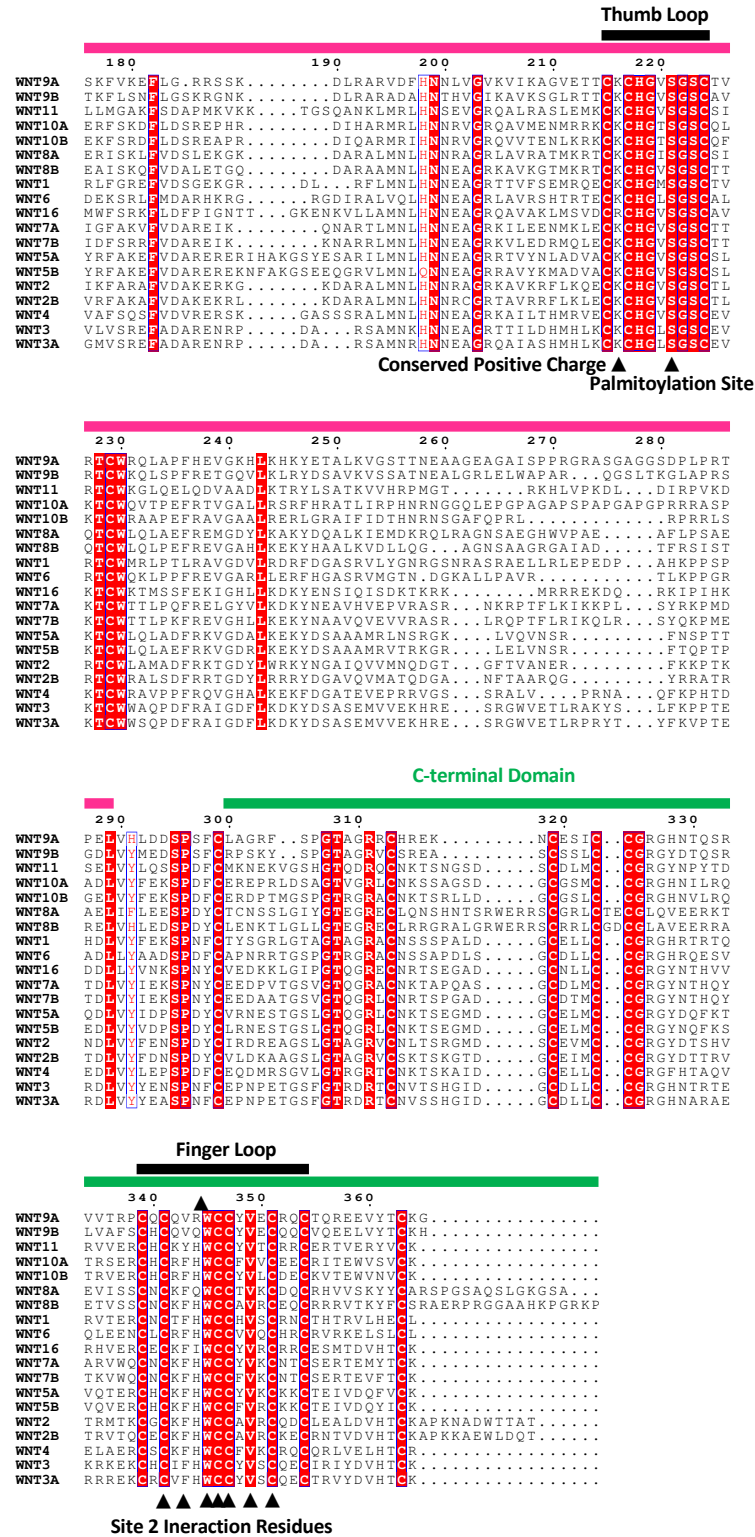
Appendix 3

Amino acid sequence alignment of all anti-FZD₃ Nbs with framework regions (A-G) and CDRs indicated. The large family comprising six highly similar Nbs is indicated by the bracket. Below, crystal structure of Nb₁₄₄₇₈ with positions of framework regions and CDRs indicated.

Appendix 4 Amino acid sequence alignment of human WNT proteins

		Signal Peptide									
		1		10		20		30		40	
WNT9A		M	L	D	G	S	P	L	A	R	W
WNT9B		M	R	P	P	P	A	L	A	L	A
WNT11		M	R	A	R	P	Q	V	C	E	A
WNT10A		M	G	S	A	H	P	R	P	W	L
WNT10B		M	L	E	E	P	R	P	P	P	S
WNT8A		M	L	E	E	P	R	P	P	P	S
WNT8B		M	L	E	E	P	R	P	P	P	S
WNT1		M	G	L	W	A	L	P	G	W	S
WNT6		M	L	P	P	L	P	S	R	L	G
WNT16		M	E	R	H	P	P	M	Q	L	T
WNT7A		M	N	R	K	A	R	R	C	L	G
WNT7B		M	H	R	N	F	R	K	W	I	F
WNT5A		M	K	K	S	I	G	I	L	S	P
WNT5B		M	P	S	L	L	L	L	L	L	L
WNT2		M	N	A	P	L	G	G	I	W	L
WNT2B		M	L	R	P	G	G	A	E	E	A
WNT4		M	S	P	R	S	C	L	R	S	L
WNT3		M	E	P	H	L	L	G	L	L	G
WNT3A		M	A	P		L	G	Y	F	L	
		N-Terminal Domain									
		50		60		70		80		90	
WNT9A		L	E	P	E	A	A	A	Q		
WNT9B		G	T	A	A	P	A	Q	G	G	A
WNT11		S	K	T	P	S	A	L	A	L	N
WNT10A		L	P	P	E	P	V	L	N	A	N
WNT10B		L	P	P	E	P	L	N	A	N	T
WNT8A		W	A	A	L	G	I	C			
WNT8B		K	P	S	V	I	C	L	F	T	C
WNT1		T	N	L	T	D	S	K	S	L	Q
WNT6		V	G	S	P	L	V	M	D	P	T
WNT16		W	L	G	I	A	S	F	G	V	P
WNT7A		G	G	F	S	S	V	A	L	G	A
WNT7B		G	A	L	S	S	V	A	L	G	A
WNT5A		V	O	M	S	E	V	I	I	G	A
WNT5B		V	O	R	P	E	M	F	I	I	G
WNT2		W	Y	M	R	A	T	G	G	S	S
WNT2B		W	Y	I	G	A	L	G			
WNT4		G	S	I							
WNT3		T	S	L	G	S	Q	P	L	L	C
WNT3A		S	S	L	G	S	Q	P	I	L	C
		110		120		130		140		150	
WNT9A		T	L	E	G	R	Y	R	A	S	L
WNT9B		S	L	E	G	R	M				
WNT11		S	S	I	E	L	A	P	N	Y	L
WNT10A		S	S	L	E	T	R	N	K	I	P
WNT10B		S	A	L	E	G	G	R	L	P	H
WNT8A		P	E	N	A	L	Q	S	T	H	N
WNT8B		P	E	R	A	L	Q	S	S	H	G
WNT1		P	T	A							
WNT6		F	G	R	I	L	Q	D	I	R	E
WNT16		M	I	T	A	A	T	T	A	P	M
WNT7A		S	A	L	G						
WNT7B		S	A	L	G						
WNT5A		S	T	V	D						
WNT5B		S	T	A	D						
WNT2		N	I	L	D	R	D	H	S	L	F
WNT2B		T	I	L	D	R	D	H	S	L	F
WNT4		S	T	L							
WNT3		T	I	D	D	S	L	A	I	F	G
WNT3A		T	T	V	H	D	S	L	A	I	F
		160		170							
WNT9A		A									
WNT9B		S									
WNT11		V	P	G	E						
WNT10A		S	R	R	G	E	E	A	F	R	R
WNT8A		S	N	N	G						
WNT8B		S	R	N	G						
WNT1		R	R	R	G						
WNT6		P	R	G							
WNT16		T	L	Q	N						
WNT7A		E	K	Q	G						
WNT7B		E	K	Q	G						
WNT5A		R	A	A							
WNT5B		R	T	A							
WNT2		K	K	M	G						
WNT2B		Y	T	R	G						
WNT4		T	V	H	G						
WNT3		H	H	K	G						
WNT3A		R	H	Q	G						

Appendix 4 Amino acid sequence alignment of human WNT proteins



Appendix 4

Amino acid sequence alignment of human WNT proteins with major domains and important residues annotated.

Appendix 5 List of constructs in the pHLsec vector

FZD ₁				
Name	Start	End	Signal peptide	Tags/Details
FZD ₁ 1-640	1	640	Native	C-terminal 3C, mV, 1D4
FZD ₁ 1-644	1	644	Native	C-terminal 3C, mV, 1D4
FZD ₁ 1-647	1	647	Native	C-terminal 3C, mV, 1D4
FZD ₁ 105-640	105	640	pHLsec	C-terminal 3C, mV, 1D4
FZD ₁ 105-644	105	644	pHLsec	C-terminal 3C, mV, 1D4
FZD ₁ 105-647	105	647	pHLsec	C-terminal 3C, mV, 1D4
FZD ₁ 108-640	108	640	pHLsec	C-terminal 3C, mV, 1D4
FZD ₁ 108-644	108	644	pHLsec	C-terminal 3C, mV, 1D4
FZD ₁ 108-647	108	647	pHLsec	C-terminal 3C, mV, 1D4
FZD ₁ 111-640	111	640	pHLsec	C-terminal 3C, mV, 1D4
FZD ₁ 111-644	111	644	pHLsec	C-terminal 3C, mV, 1D4
FZD ₁ 111-647	111	647	pHLsec	C-terminal 3C, mV, 1D4
FZD ₁ 113-640	113	640	pHLsec	C-terminal 3C, mV, 1D4
FZD ₁ 113-644	113	644	pHLsec	C-terminal 3C, mV, 1D4
FZD ₁ 113-647	113	647	pHLsec	C-terminal 3C, mV, 1D4
Name	Start	End	Signal peptide	Details
BRIL_FZD ₁ 1-640	1	640	Native	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 1-644	1	644	Native	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 1-647	1	647	Native	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 105-640	105	640	pHLsec	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 105-644	105	644	pHLsec	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 105-647	105	647	pHLsec	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 108-640	108	640	pHLsec	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4

Appendix 5 List of constructs in the pHLsec vector

BRIL_FZD ₁ 108-644	108	644	pHLsec	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 108-647	108	647	pHLsec	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 111-640	111	640	pHLsec	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 111-644	111	644	pHLsec	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 111-647	111	647	pHLsec	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 113-640	113	640	pHLsec	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 113-644	113	644	pHLsec	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4
BRIL_FZD ₁ 113-647	113	647	pHLsec	BRIL insert between M520 and T527, C-terminal 3C, mV, 1D4

FZD2				
Name	Start	End	Signal peptide	Details
FZD ₂ 1-558	1	558	Native	C-terminal 3C, mV, 1D4
FZD ₂ 34-558	34	558	pHLsec	C-terminal 3C, mV, 1D4
Name	Start	End	Signal peptide	Details
BRIL_FZD ₂ 1-558	1	558	Native	C-terminal 3C, mV, 1D4
BRIL_FZD ₂ 34-558	34	558	pHLsec	C-terminal 3C, mV, 1D4

FZD3				
Name	Start	End	Signal peptide	Details
FZD ₃ 1-517	1	511	Native	C504S, ending SRLTGQ, C-terminal 3C, mV, 1D4
FZD ₃ 23-517	23	511	pHLsec	C504S, ending SRLTGQ, C-terminal 3C, mV, 1D4

Appendix 5 List of constructs in the pHLsec vector

Name	Start	End	Signal peptide	Details
BRIL_FZD ₃ 1-517	1	511	Native	C504S, ending SRLTGQ, BRIL insert between I404 and Q411, 3C site, mV, 1D4 tag
BRIL_FZD ₃ 23-517	23	511	pHLsec	C504S, ending SRLTGQ, BRIL insert between I404 and Q411, 3C site, mV, 1D4 tag

FZD8				
Name	Start	End	Signal peptide	Details
FZD ₈ 1-623 Δ173-203 Δ342-376	1	623	Native	Δ173-203, Δ342-376, Ending LTGQ, C-terminal 3C, mV, 1D4
FZD ₈ 28-623 Δ173-203 Δ342-376	28	623	pHLsec	Δ173-203, Δ342-376, Ending LTGQ, C-terminal 3C, mV, 1D4
Name	Start	End	Signal peptide	Details
BRIL_FZD ₈ 1-623 Δ173-203 Δ342-376	1	623	Native	Δ173-203, Δ342-376, Ending LTGQ BRIL insert between I448 and T457, C-terminal 3C, mV, 1D4
BRIL_FZD ₈ 28-623 Δ173-203 Δ342-376	28	623	pHLsec	Δ173-203, Δ342-376, Ending LTGQ BRIL insert between I448 and T457, C-terminal 3C, mV, 1D4

FZD10				
Name	Start	End	Signal peptide	Details
FZD ₁₀ 1-541	1	541	Native	C535S, ending LTGQ, C-terminal 3C, mV, 1D4
FZD ₁₀ 21-541	21	541	pHLsec	C535S, ending LTGQ, C-terminal 3C, mV, 1D4
Name	Start	End	Signal peptide	Details
BRIL_FZD ₁₀ 1-541	1	541	Native	C535S, ending LTGQ, BRIL insert between M427 and T434, C-terminal 3C, mV, 1D4

Appendix 5 List of constructs in the pHLsec vector

BRIL_FZD ₁₀ 21-541	21	541	pHLsec	C535S, ending LTGQ, BRIL insert between M427 and T434, C-terminal 3C, mV, 1D4

FZD3				
Name	Start	End	Signal peptide	Details
BRIL_FZD ₃ S215A	23	517	pHLsec	C504S, ending SRLTGQ, BRIL insert between I404 and Q411, C-terminal 3C, mV, 1D4
BRIL_FZD ₃ S301F	23	517	pHLsec	C504S, ending SRLTGQ, BRIL insert between I404 and Q411, C-terminal 3C, mV, 1D4
BRIL_FZD ₃ S334A	23	517	pHLsec	C504S, ending SRLTGQ, BRIL insert between I404 and Q411, C-terminal 3C, mV, 1D4
BRIL_FZD ₃ T341A	23	517	pHLsec	C504S, ending SRLTGQ, BRIL insert between I404 and Q411, C-terminal 3C, mV, 1D4
BRIL_FZD ₃ G393V	23	517	pHLsec	C504S, ending SRLTGQ, BRIL insert between I404 and Q411, C-terminal 3C, mV, 1D4
BRIL_FZD ₃ 26-517	26	517	pHLsec	C504S, ending SRLTGQ, BRIL insert between I404 and Q411, C-terminal 3C, mV, 1D4
BRIL_FZD ₃ 26-517 S215A S334A	26	517	pHLsec	C504S, ending SRLTGQ, BRIL insert between I404 and Q411, C-terminal 3C, mV, 1D4
BRIL_FZD ₃ 26-517 S215A S334A G393V	26	517	pHLsec	C504S, ending SRLTGQ, BRIL insert between I404 and Q411, C-terminal 3C, mV, 1D4

Appendix 5

Selected pHLsec constructs

Appendix 6 List of constructs in the pHR-CMV-TetO₂ vector

FZD ₃				
Name	Start	End	Signal peptide	Details
BRIL_FZD ₃ 26-517 S215A S334A G393V	26	517	pHR-CMV-TetO ₂	C504S, ending SRLTGQ, BRIL insert between I404 and Q411, N-terminal TS, mV, 3C
FZD ₃ 1-666	1	666	Native	Full length, wild type, untagged
FZD ₃ 1-517	1	517	Native	Untagged
BRIL_FZD ₃ 1-666	1	666	Native	Untagged
FZD ₃ 1-666 G393V	1	666	Native	Untagged

FZD ₃ CRD				
Name	Start	End	Signal peptide	Details
FZD ₃ 26-138	26	138	pHR-CMV-TetO ₂	C-terminal 3C, mV, TS

FZD ₁₀ CRD				
Name	Start	End	Signal peptide	Details
FZD ₃₃₁₋₁₅₇	31	157	pHR-CMV-TetO ₂	C-terminal 3C, mV, TS

Nb ₁₄₄₇₈				
Name	Start	End	Signal peptide	Details
Nb ₁₄₄₇₈	1	119	pHR-CMV-TetO ₂	Untagged

ROR2_LRP6 Fusion				
Name	Start	End	Signal peptide	Details
ROR2_LRP6 Fusion	SEE	DETAILS	ROR2	ROR2 1-403 fused to LRP6 1394-1613

Nb ₁₄₄₇₈ _DKK1C Fusion				
Name	Start	End	Signal peptide	Details
Nb ₁₄₄₇₈ _DKK1C Fusion	SEE	DETAILS	pHR-CMV-TetO ₂	Nb ₁₄₄₇₈ 1-119 fused to DKK1 177-266

Appendix 6 List of constructs in the pHR-CMV-TetO₂ vector

Biotin ligase (<i>E. Coli</i>)				
Name	Start	End	Signal peptide	Details
BirA	1	321	Igk	N-terminal HA tag, C-terminal KDEL

All FZD CRDs				
Name	Start	End	Signal peptide	Tags/Details
FZD ₁ 113-251	113	251	pHR-CMV-TetO ₂	C-terminal 3C, AVI, HIS6
FZD ₂ 36-158	36	158	pHR-CMV-TetO ₂	C-terminal 3C, AVI, HIS6
FZD ₃ 26-143	26	143	pHR-CMV-TetO ₂	C-terminal 3C, AVI, HIS6
FZD ₄ 42-179	42	179	pHR-CMV-TetO ₂	C-terminal 3C, AVI, HIS6
FZD ₅ 31-176	31	176	pHR-CMV-TetO ₂	C-terminal 3C, AVI, HIS6
FZD ₆ 22-139	22	139	pHR-CMV-TetO ₂	C-terminal 3C, AVI, HIS6
FZD ₇ 33-173	33	173	pHR-CMV-TetO ₂	C-terminal 3C, AVI, HIS6
FZD ₈ 28-153	28	153	pHR-CMV-TetO ₂	C-terminal 3C, AVI, HIS6
FZD ₉ 36-162	36	162	pHR-CMV-TetO ₂	C-terminal 3C, AVI, HIS6
FZD ₁₀ 31-157	31	157	pHR-CMV-TetO ₂	C-terminal 3C, AVI, HIS6

Appendix 6

Selected pHR-CMV-TetO₂ constructs

Appendix 7 Useful abbreviations and symbols

°C	Degrees Celsius
7TM	GPCR seven-pass transmembrane domain
APC	Adenomatous polyposis coli
BRIL	Thermostabilised apocytochrome b ₅₆₂ containing mutations M7W, H102I and R106L
βTrCP	B-transducin repeat-containing E3 ubiquitin protein ligase
CAMKII	Calcium/calmodulin-dependent kinase II
CELSR	Cadherin EGF LAG seven-pas G-type receptor
CHS	Cholesteryl hemisuccinate
CK1	Casein kinase 1
CRD	Cysteine-rich domain
CTT	C-terminal tail
DDM	n-Dodecyl β-D-maltoside
DKK	Dickopff
DME	Dimethyl ether
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DVL	Dishevelled
<i>e</i>	The electric charge carried by a single proton (1.602176634 x 10 ⁻¹⁹ C)
ECL	Extracellular loop
EDTA	Ethylenediaminetetraacetic acid

EGF	Epidermal growth factor receptor
EGFR	Epidermal growth factor receptor
EndoF1	Endo- β -N-acetylglucosaminidase F1
ER	Endoplasmic reticulum
FBS	Foetal bovine serum
FSEC	Fluorescence size-exclusion chromatography
FZD	Frizzled
<i>g</i>	Acceleration of gravity at the Earth's surface (9.81 m/s ²)
GAG	Glycosaminoglycans
GlcNAc	N-acetyl-D-glucosamine
GnTI-	α -1,3-mannosyl-glycoprotein 2- β -N-acetylglucosaminyltransferase-deficient variant of the HEK293S cell line
GPCR	G protein-coupled receptor
GPR124	Orphan GPCR 124
GSK3	Glycogen synthase kinase 3
H8	GPCR helix 8
HCC	Hepatocellular carcinoma
HEK	Human embryonic kidney (cells)
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid
HRP	Horseradish peroxidase
HPSG	Heparan sulphate proteoglycan
ICL	Intracellular loop
Ig	Immunoglobulin

IPTG	Isopropyl β -D-1-thiogalactopyranoside
K_d	Dissociation constant
kDa	Kilodalton
KRM	Kremen
LB	Lysogeny broth
LCP	Lipidic cubic phase
LD	Linker domain
LEF	Lymphoid enhancer factor
LGR	leucine-rich repeat-containing GPCR
LRP	Low-density lipoprotein receptor-related protein
MAP3K7	Mitogen-activated protein kinase kinase kinase 7
MMTV	Mouse mammary tumour virus
MW	Molecular weight
Nb	Nb
Nb ₁₄₄₇₈ _DKK1C	Nb ₁₄₄₇₈ fused to the DKK1 C-terminal domain
NEAA	Non-essential amino acids
NHERF1	Na ⁺ /H ⁺ exchange regulatory factor 1
NLK	Nemo-like kinase
NTR	Netrin-like domain
PAM	Palmitoleic acid
PBS	Phosphate-buffered saline
PCP	Planar cell polarity
PCR	Polymerase chain reaction
PDB	Protein Data Bank

PEI	Polyethyleneimine
PKC	Protein kinase C
RECK	Reversion-inducing cysteine-rich protein with Kazal motifs
RMSD	Root mean square deviation
RNF43	Ring finger protein 43
ROR	Retinoic acid-related orphan tyrosine kinase receptor
RSPO	R-spondin
RYK	Related to tyrosine kinase receptor
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SFRP	Secreted Frizzled-related protein
SMO	Smoothed
TCF	T-cell factor
TetR	Tet repressor protein
TM1-7	GPCR transmembrane helices 1-7
TMD	Transmembrane domain
TNBC	Triple-negative breast cancer
Tris	Tris(hydroxymethyl)aminomethane
VANGL	Van Gogh-like
VEGF	Vascular endothelial growth factor
WIF	WNT inhibitory factor
WNT	A mnemonic based on 'Wingless-related integration site'
ZNRF3	Zinc and ring finger 3

References

- Adams, P. D. *et al.* (2010) ‘PHENIX: A comprehensive Python-based system for macromolecular structure solution’, *Acta Crystallographica Section D: Biological Crystallography*, 66(2), pp. 213–221. doi: 10.1107/S0907444909052925.
- Agostino, M., Öther-Gee Pohl, S. and Dharmarajan, A. (2019) ‘Wnt Binding Affinity Prediction for Putative Frizzled-Type Cysteine-Rich Domains’, *International Journal of Molecular Sciences*, pp. 1–11.
- Agostino, M., Pohl, S. Ö. G. and Dharmarajan, A. (2017) ‘Structure-based prediction of Wnt binding affinities for frizzled-type cysteine-rich domains’, *Journal of Biological Chemistry*, 292(27), pp. 11218–11229. doi: 10.1074/jbc.M117.786269.
- Alexander, S. P. H. *et al.* (2017) ‘THE CONCISE GUIDE TO PHARMACOLOGY 2017/18: G protein-coupled receptors’, *British Journal of Pharmacology*, 174, pp. S17–S129. doi: 10.1111/bph.13878.
- Angers, S. and Moon, R. T. (2009) ‘Proximal events in Wnt signal transduction.’, *Nature reviews. Molecular cell biology*. Nature Publishing Group, 10(7), pp. 468–77. doi: 10.1038/nrm2717.
- Aricescu, A. R., Lu, W. and Jones, E. Y. (2006) ‘A time- and cost-efficient system for high-level protein production in mammalian cells’, *Acta Crystallographica Section D: Biological Crystallography*, 62(10), pp. 1243–1250. doi: 10.1107/S0907444906029799.
- Ballesteros, J. A. and Weinstein, H. (1995) ‘Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors’, *Methods in Neurosciences*. doi: 10.1016/S1043-9471(05)80049-7.
- Bang, I. *et al.* (2018) ‘Biophysical and functional characterization of Norrin signaling through Frizzled4’, *Proceedings of the National Academy of Sciences of the United States of America*, 115(35), pp. 8787–8792. doi: 10.1073/pnas.1817333115.
- Behrens, J. *et al.* (1996) ‘Functional Interaction of β -Catenin with the Transcription Factor LEF-1’, *Nature*, pp. 638–642. doi: 10.1038/382638a0.
- Bhanot, P. *et al.* (1996) ‘A new member of the frizzled family from Drosophila functions as a Wingless receptor.’, *Nature*, 382(6588), pp. 225–230. doi:

10.1038/382225a0.

- Bowin, C.-F., Inoue, A. and Schulte, G. (2019) 'WNT-3A-induced β -catenin signaling does not require signaling through heterotrimeric G proteins', *Journal of Biological Chemistry*, 7(1), p. jbc.AC119.009412. doi: 10.1074/jbc.AC119.009412.
- Brott, B. K. and Sokol, S. Y. (2002) 'Regulation of Wnt/LRP signaling by distinct domains of Dickkopf proteins.', *Molecular and cellular biology*, 22(17), pp. 6100–10. doi: 10.1128/MCB.22.17.6100.
- Bu, Q. *et al.* (2017) 'The crystal structure of full-length Sizzled from *Xenopus laevis* yields insights into Wnt-antagonistic function of secreted Frizzled-related proteins', *Journal of Biological Chemistry*, 292(39), pp. 16055–16069. doi: 10.1074/jbc.M117.791756.
- Byrne, E. F. X. *et al.* (2016) 'Structural basis of Smoothed regulation by its extracellular domains', *Nature*. Nature Publishing Group, 535(7613), pp. 1–21. doi: 10.1038/nature18934.
- Caldwell, G. M. *et al.* (2004) 'The Wnt Antagonist sFRP1 in Colorectal Tumorigenesis', *Cancer Research*, 64(3), pp. 883–888. doi: 10.1158/0008-5472.CAN-03-1346.
- Chang, T.-H. *et al.* (2015) 'Structure and functional properties of Norrin mimic Wnt for signalling with Frizzled4, Lrp5/6, and proteoglycan', *eLife*, 4, pp. 1–27. doi: 10.7554/eLife.06554.
- Chen, C. -m. *et al.* (2004) 'Evidence that the cysteine-rich domain of *Drosophila* Frizzled family receptors is dispensable for transducing Wingless', *Proceedings of the National Academy of Sciences*, 101(45), pp. 15961–15966. doi: 10.1073/pnas.0407103101.
- Chen, K.-Y. M. *et al.* (2012) 'Naturally evolved G protein-coupled receptors adopt metastable conformations', *Proceedings of the National Academy of Sciences*, 109(33), pp. 13284–13289. doi: 10.1073/pnas.1205512109.
- Chen, P. *et al.* (2018) 'Structural basis for recognition of frizzled proteins by clostridium difficile toxin B', *Science*, 360(6389), pp. 664–669. doi: 10.1126/science.aar1999.
- Cho, C. *et al.* (2019) 'Molecular determinants in Frizzled, Reck, and Wnt7a for ligand-specific signaling in neurovascular development', *eLife*, 8, pp. 1–19. doi:

10.7554/eLife.47300.

- Chun, E. *et al.* (2012) ‘Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors’, *Structure*. Elsevier Ltd, 20(6), pp. 967–976. doi: 10.1016/j.str.2012.04.010.
- Clevers, H. and Nusse, R. (2012) ‘Wnt/ β -catenin signaling and disease’, *Cell*, 149(6), pp. 1192–1205. doi: 10.1016/j.cell.2012.05.012.
- Cong, F., Schweizer, L. and Varmus, H. (2004) ‘Wnt signals across the plasma membrane to activate the β -catenin pathway by forming oligomers containing its receptors, Frizzled and LRP’, *Development*, 131(20), pp. 5103–5115. doi: 10.1242/dev.01318.
- Corpet, F. (1988) ‘Multiple sequence alignment with hierarchical clustering’, *Nucleic Acids Research*, 16(22), pp. 10881–10890.
- Cross, F. R. *et al.* (1997) ‘Constitutive Transcriptional Activation by a β -Catenin–Tcf Complex in APC^{-/-} Colon Carcinoma’, *Science*, 275(March), pp. 3–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9065401>.
- Cruciat, C. M. and Niehrs, C. (2013) ‘Secreted and transmembrane Wnt inhibitors and activators’, *Cold Spring Harbor Perspectives in Biology*, 5(3), pp. 1–26. doi: 10.1101/cshperspect.a015081.
- Dang, L. T. *et al.* (2019) ‘Receptor subtype discrimination using extensive shape complementary designed interfaces’, *Nature Structural and Molecular Biology*. Springer US, 26(6), pp. 407–414. doi: 10.1038/s41594-019-0224-z.
- Daniels, D. L. and Weis, W. I. (2005) ‘ β -catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation’, *Nature Structural and Molecular Biology*, 12(4), pp. 364–371. doi: 10.1038/nsmb912.
- Dann, C. E. *et al.* (2001) ‘Insights into Wnt binding and signalling from the structures of two Frizzled cysteine-rich domains’, *Nature*, 412(July), pp. 86–90. doi: 10.1038/35083601.
- DeAlmeida, V. I. *et al.* (2007) ‘The soluble Wnt receptor Frizzled8CRD-hFc inhibits the growth of teratocarcinomas in vivo’, *Cancer Research*, 67(11), pp. 5371–5379. doi: 10.1158/0008-5472.CAN-07-0266.
- DeBruine, Z. J. *et al.* (2017) ‘Wnt5a promotes Frizzled-4 signalosome assembly by stabilizing cysteine-rich domain dimerization’, *Genes and Development*, 31(9), pp.

- 916–926. doi: 10.1101/gad.298331.117.
- Dijksterhuis, J. P. *et al.* (2015) ‘Systematic mapping of WNT-FZD protein interactions reveals functional selectivity by distinct WNT-FZD pairs’, *Journal of Biological Chemistry*, 290(11), pp. 6789–6798. doi: 10.1074/jbc.M114.612648.
 - Dijksterhuis, J. P., Petersen, J. and Schulte, G. (2014) ‘WNT/Frizzled signalling: Receptor-ligand selectivity with focus on FZD-G protein signalling and its physiological relevance: IUPHAR Review 3’, *British Journal of Pharmacology*, 171(5), pp. 1195–1209. doi: 10.1111/bph.12364.
 - Dong, B. *et al.* (2018) ‘Functional redundancy of frizzled 3 and frizzled 6 in planar cell polarity control of mouse hair follicles’, *Development (Cambridge)*, 145(19). doi: 10.1242/dev.168468.
 - Elegheert, J. *et al.* (2018) ‘Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins’, *Nature Protocols*, 13(12), pp. 2991–3017. doi: 10.1038/s41596-018-0075-9.
 - Emsley, P. and Cowtan, K. (2004) ‘Coot: Model-building tools for molecular graphics’, *Acta Crystallographica Section D: Biological Crystallography*. International Union of Crystallography, 60(12 I), pp. 2126–2132. doi: 10.1107/S0907444904019158.
 - Eubelen, M. *et al.* (2018) ‘A molecular mechanism for Wnt ligand-specific signaling’, *Science*, 361(6403). doi: 10.1126/science.aat1178.
 - Evans, P. R. and Murshudov, G. N. (2013) ‘How good are my data and what is the resolution?’, *Acta Crystallographica Section D: Biological Crystallography*. International Union of Crystallography, 69(7), pp. 1204–1214. doi: 10.1107/S0907444913000061.
 - Foord, S. M. *et al.* (2005) ‘G Protein-Coupled Receptor List’, *International Union of Pharmacology XLVI*, 57(2), pp. 279–288. doi: 10.1124/pr.57.2.5.2.
 - Gammons, M. V. *et al.* (2016) ‘Wnt Signalingosome Assembly by DEP Domain Swapping of Dishevelled’, *Molecular Cell*. MRC Laboratory of Molecular Biology, 64(1), pp. 92–104. doi: 10.1016/j.molcel.2016.08.026.
 - Generoso, S. F. *et al.* (2015) ‘Pharmacological folding chaperones act as allosteric ligands of Frizzled4’, *Nature Chemical Biology*. Nature Publishing Group, 11(4), pp. 280–286. doi: 10.1038/nchembio.1770.
 - De Genst, E. *et al.* (2006) ‘Molecular basis for the preferential cleft recognition by

- dromedary heavy-chain antibodies', *Proceedings of the National Academy of Sciences*, 103(12), pp. 4586–4591. doi: 10.1073/pnas.0505379103.
- Golan, T. *et al.* (2004) 'The Human Frizzled 6 (HFz6) Acts as a Negative Regulator of the Canonical Wnt- β -Catenin Signaling Cascade', *Journal of Biological Chemistry*, 279(15), pp. 14879–14888. doi: 10.1074/jbc.M306421200.
 - Gouet, P. *et al.* (1999) 'ESPrict: Analysis of multiple sequence alignments in PostScript', *Bioinformatics*, 15(4), pp. 305–308. doi: 10.1093/bioinformatics/15.4.305.
 - Grainger, S. *et al.* (2019) 'EGFR is required for Wnt9a-Fzd9b signalling specificity in haematopoietic stem cells', *Nature Cell Biology*. Springer US, 21(6), pp. 721–730. doi: 10.1038/s41556-019-0330-5.
 - Grumolato, L. *et al.* (2010) 'Canonical and noncanonical Wnts use a common mechanism to activate completely unrelated coreceptors', *Genes and Development*, 24(22), pp. 2517–2530. doi: 10.1101/gad.1957710.
 - Gujral, T. S. *et al.* (2014) 'A noncanonical Frizzled2 pathway regulates epithelial-mesenchymal transition and metastasis', *Cell*. Elsevier Inc., 159(4), pp. 844–856. doi: 10.1016/j.cell.2014.10.032.
 - Gurney, A. *et al.* (2012) 'Wnt pathway inhibition via the targeting of Frizzled receptors results in decreased growth and tumorigenicity of human tumors', *Proceedings of the National Academy of Sciences*, 109(29), pp. 11717–11722. doi: 10.1073/pnas.1120068109.
 - Hansen, C. *et al.* (2009) 'Wnt-5a-induced phosphorylation of DARPP-32 inhibits breast cancer cell migration in a CREB-dependent manner', *Journal of Biological Chemistry*, 284(40), pp. 27533–27543. doi: 10.1074/jbc.M109.048884.
 - Hao, H.-X. *et al.* (2012) 'ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner', *Nature*. Nature Publishing Group, 485(7397), pp. 195–200. doi: 10.1038/nature11019.
 - He, A. T. *et al.* (2008) 'Identification of c-MYC as a Target of the APC Pathway Identification Target of as the Pathway', *Advancement Of Science*, 281(5382), pp. 1509–1512.
 - Hirai, H. *et al.* (2019) 'Crystal structure of a mammalian Wnt-frizzled complex', *Nature Structural and Molecular Biology*. Springer US, 26(May). doi: 10.1038/s41594-019-0216-z.

- Ho, S. N. *et al.* (1989) 'Site-directed mutagenesis by overlap extension using the polymerase chain reaction', *Gene*. doi: 10.1016/0378-1119(89)90358-2.
- Hodgess, R. S. *et al.* (1988) 'Antigen-Antibody Interaction', *Journal of Biological Chemistry*, 263(24), pp. 11768–11775.
- Hsieh, J.-C. *et al.* (1999) 'A new secreted protein that binds to Wnt proteins and inhibits their activities', *Nature*, 398(6726), pp. 431–436. doi: 10.1038/18899.
- Hua, Z. L. *et al.* (2014) 'Partial interchangeability of Fz3 and Fz6 in tissue polarity signaling for epithelial orientation and axon growth and guidance', *Development*, 141(20), pp. 3944–3954. doi: 10.1242/dev.110189.
- Huang, P. *et al.* (2018) 'Structural Basis of Smoothed Activation in Hedgehog Signaling', *Cell*. Elsevier, 174(2), pp. 312–324.e16. doi: 10.1016/j.cell.2018.04.029.
- Isberg, V. *et al.* (2015) 'Generic GPCR residue numbers - Aligning topology maps while minding the gaps', *Trends in Pharmacological Sciences*. Elsevier Ltd, 36(1), pp. 22–31. doi: 10.1016/j.tips.2014.11.001.
- Ishitani, T. *et al.* (1999) 'The TAK1-NLK-MAPK-related pathway antagonizes signalling between β -catenin and transcription factor TCF', *Nature*, 399(6738), pp. 798–802. doi: 10.1038/21674.
- Janda, C. Y. *et al.* (2012) 'Structural Basis of Wnt Recognition by Frizzled', *Science (New York, N.Y.)*, 337(6090), pp. 59–64. doi: 10.1126/science.1222879.
- Janda, C. Y. *et al.* (2017) 'Surrogate Wnt agonists that phenocopy canonical Wnt and β -catenin signalling', *Nature*. Nature Publishing Group, 545(7653), pp. 234–237. doi: 10.1038/nature22306.
- Jenny, A. *et al.* (2005) 'Diego and Prickle regulate frizzled planar cell polarity signalling by competing for Dishevelled binding', *Nature Cell Biology*, 7(7), pp. 691–697. doi: 10.1038/ncb1271.
- Kabsch, W. (2010) 'XDS', *Acta Crystallographica Section D Biological Crystallography*, 66(2), pp. 125–132. doi: 10.1107/S0907444909047337.
- Kakugawa, S. *et al.* (2015) 'Notum deacylates Wnt proteins to suppress signalling activity', *Nature*. Nature Publishing Group, 519(7542), pp. 187–192. doi: 10.1038/nature14259.
- Katanaev, V. L. *et al.* (2005) 'Trimeric G protein-dependent frizzled signaling in *Drosophila*', *Cell*, 120(1), pp. 111–122. doi: 10.1016/j.cell.2004.11.014.

- Kilander, M. B. C. *et al.* (2014) 'Disheveled regulates precoupling of heterotrimeric G proteins to Frizzled 6', *FASEB Journal*, 28(5), pp. 2293–2305. doi: 10.1096/fj.13-246363.
- King, T. D. *et al.* (2012) 'Frizzled7 as an emerging target for cancer therapy', *Cellular Signalling*. Elsevier Inc., 24(4), pp. 846–851. doi: 10.1016/j.cellsig.2011.12.009.
- Komiya, Y. and Habas, R. (2008) 'Wnt signal transduction pathways', *Organogenesis*, 4(2), pp. 68–75. doi: 10.4161/org.4.2.5851.
- Kong, J. H., Siebold, C. and Rohatgi, R. (2019) 'Biochemical mechanisms of vertebrate hedgehog signaling', *Development*, 146(10), p. dev166892. doi: 10.1242/dev.166892.
- Koo, B. K. *et al.* (2012) 'Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors', *Nature*. Nature Publishing Group, 488(7413), pp. 665–669. doi: 10.1038/nature11308.
- Koval, A. and Katanaev, V. L. (2011) 'Wnt3a stimulation elicits G-protein-coupled receptor properties of mammalian Frizzled proteins.', *The Biochemical journal*, 433(3), pp. 435–440. doi: 10.1042/BJ20101878.
- De Lau, W. *et al.* (2011) 'Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling', *Nature*. Nature Publishing Group, 476(7360), pp. 293–297. doi: 10.1038/nature10337.
- Li, C. *et al.* (2019) 'Down-regulation of FZD3 receptor suppresses growth and metastasis of human melanoma independently of canonical WNT signaling', *Proceedings of the National Academy of Sciences*, 116(10), pp. 4548–4557. doi: 10.1073/pnas.1813802116.
- Li, V. S. W. *et al.* (2012) 'Wnt Signaling through Inhibition of β -Catenin Degradation in an Intact Axin1 Complex', *Cell*. Elsevier Inc., 149(6), pp. 1245–1256. doi: 10.1016/j.cell.2012.05.002.
- Lin, K. *et al.* (1997) 'The cysteine-rich frizzled domain of Frzb-1 is required and sufficient for modulation of Wnt signaling', *Proceedings of the National Academy of Sciences*, 94(October), pp. 1–5. Available at: papers2://publication/uuid/1B9DCBC9-995E-4413-B827-0555403020EA.
- Lin, X. (2004) 'Functions of heparan sulfate proteoglycans in cell signaling during development', *Development*, 131(24), pp. 6009–6021. doi: 10.1242/dev.01522.
- Liu, T. *et al.* (2001) 'G protein signaling from activated rat Frizzled-1 to the β -catenin-

- lef-tcf pathway', *Science*, 292(5522), pp. 1718–1722. doi: 10.1126/science.1060100.
- MacDonald, B. T. and He, X. (2012) 'Frizzled and LRP5/6 receptors for wnt/ β -catenin signaling', *Cold Spring Harbor Perspectives in Biology*, 4(12). doi: 10.1101/cshperspect.a007880.
 - Macheda, M. L. *et al.* (2012) 'The Wnt receptor Ryk plays a role in mammalian planar cell polarity signaling', *Journal of Biological Chemistry*, 287(35), pp. 29312–29323. doi: 10.1074/jbc.M112.362681.
 - Malinauskas, T. *et al.* (2011) 'Modular mechanism of Wnt signaling inhibition by Wnt inhibitory factor 1.', *Nature structural & molecular biology*. Nature Publishing Group, 18(8), pp. 886–93. doi: 10.1038/nsmb.2081.
 - Mao, B. *et al.* (2001) 'LDL-receptor-related protein 6 is a receptor for Dickkopf proteins', *Nature*, 411(6835), pp. 321–325. doi: 10.1038/35077108.
 - Mao, B. *et al.* (2002) 'Kremen proteins are Dickkopf receptors that regulate Wnt/ β -catenin signalling', *Nature*, 417(June), pp. 6–9.
 - Mao, B. and Niehrs, C. (2003) 'Kremen2 modulates Dickkopf2 activity during Wnt/LRP6 signaling', *Gene*, 302(1–2), pp. 179–183. doi: 10.1016/S0378-1119(02)01106-X.
 - McCoy, A. J. *et al.* (2007) 'Phaser crystallographic software', *Journal of Applied Crystallography*. International Union of Crystallography, 40(4), pp. 658–674. doi: 10.1107/S0021889807021206.
 - Mészáros, B., Erdős, G. and Dosztányi, Z. (2018) 'IUPred2A: Context-dependent prediction of protein disorder as a function of redox state and protein binding', *Nucleic Acids Research*, 46(W1), pp. W329–W337. doi: 10.1093/nar/gky384.
 - Mii, Y. and Taira, M. (2009) 'Secreted Frizzled-related proteins enhance the diffusion of Wnt ligands and expand their signalling range', *Development*, 136(24), pp. 4083–4088. doi: 10.1242/dev.032524.
 - Molenaar, M. *et al.* (1996) 'XTcf-3 transcription factor mediates β -catenin-induced axis formation in xenopus embryos', *Cell*, 86(3), pp. 391–399. doi: 10.1016/S0092-8674(00)80112-9.
 - Morin, P. J. *et al.* (1997) 'Activation of β -catenin-Tcf signaling in colon cancer by mutations in β -catenin or APC', *Science*, 275(5307), pp. 1787–1790. doi: 10.1126/science.275.5307.1787.

- Mullis, K. *et al.* (1986) 'Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. 1986.', *Biotechnology (Reading, Mass.)*, 24(Table 1), pp. 17–27. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1422010>.
- Muyldermans, S. (2013) 'Nanobodies: Natural Single-Domain Antibodies', *Annual Review of Biochemistry*, 82(1), pp. 775–797. doi: 10.1146/annurev-biochem-063011-092449.
- Nagasaki, K. *et al.* (2018) 'Nonsense mutations in FZD2 cause autosomal-dominant omodysplasia: Robinow syndrome-like phenotypes', *American Journal of Medical Genetics, Part A*, 176(3), pp. 739–742. doi: 10.1002/ajmg.a.38623.
- Niehrs, C. (2012) 'The complex world of WNT receptor signalling.', *Nature reviews. Molecular cell biology*. Nature Publishing Group, 13(12), pp. 767–779. doi: 10.1038/nrm3470.
- Nile, A. H. *et al.* (2017) 'Unsaturated fatty acyl recognition by Frizzled receptors mediates dimerization upon Wnt ligand binding', *Proceedings of the National Academy of Sciences*, 114(16), pp. 4147–4152. doi: 10.1073/pnas.1618293114.
- Nile, A. H. *et al.* (2018) 'A selective peptide inhibitor of Frizzled 7 receptors disrupts intestinal stem cells article', *Nature Chemical Biology*. Springer US, 14(6), pp. 582–590. doi: 10.1038/s41589-018-0035-2.
- Nile, A. H. and Hannoush, R. N. (2019) 'Fatty acid recognition in the Frizzled receptor family', *Journal of Biological Chemistry*, 294(2), pp. 726–736. doi: 10.1074/jbc.REV118.005205.
- Nusse, R. *et al.* (1991) 'A new nomenclature for int-1 and related genes: The Wnt gene family', *Cell*, 64(2), p. 231. doi: 10.1016/0092-8674(91)90633-a.
- Nusse, R. and Varmus, H. E. (1982) 'Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome', *Cell*, 31(1), pp. 99–109. doi: 10.1016/0092-8674(82)90409-3.
- Pace, C. N. and Scholtz, J. M. (1998) 'A helix propensity scale based on experimental studies of peptides and proteins.', *Biophysical journal*, 75(1), pp. 422–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9649402>0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1299714.
- Pardon, E. *et al.* (2014) 'A general protocol for the generation of Nanobodies for structural biology', *Nature Protocols*, 9(3), pp. 674–693. doi: 10.1038/nprot.2014.039.

- Park, H. W. *et al.* (2015) ‘Alternative Wnt Signaling Activates YAP/TAZ’, *Cell*, 162(4), pp. 780–794. doi: 10.1016/j.cell.2015.07.013.
- Pavlovic, Z. *et al.* (2018) ‘A synthetic anti-Frizzled antibody engineered for broadened specificity exhibits enhanced anti-tumor properties’, *mAbs*. Taylor & Francis, 10(8), pp. 1157–1167. doi: 10.1080/19420862.2018.1515565.
- Petersen, J. *et al.* (2017) ‘Agonist-induced dimer dissociation as a macromolecular step in G protein-coupled receptor signaling’, *Nature Communications*. Springer US, 8(1). doi: 10.1038/s41467-017-00253-9.
- Pietrobono, S. and Stecca, B. (2018) ‘Targeting the Oncoprotein Smoothed by Small Molecules: Focus on Novel Acylguanidine Derivatives as Potent Smoothed Inhibitors’, *Cells*, 7(12), p. 272. doi: 10.3390/cells7120272.
- Povelones, M. and Nusse, R. (2005) ‘The role of the cysteine-rich domain of Frizzled in Wingless-Armadillo signaling’, *EMBO Journal*, 24(19), pp. 3493–3503. doi: 10.1038/sj.emboj.7600817.
- Predonzani, A. *et al.* (2008) ‘In vivo site-specific biotinylation of proteins within the secretory pathway using a single vector system’, *BMC Biotechnology*, 8, pp. 1–11. doi: 10.1186/1472-6750-8-41.
- Qi, X. *et al.* (2019) ‘Cryo-EM structure of oxysterol-bound human Smoothed coupled to a heterotrimeric Gi’, *Nature*. Springer US. doi: 10.1038/s41586-019-1286-0.
- Ralser, M. *et al.* (2006) ‘An efficient and economic enhancer mix for PCR’, *Biochemical and Biophysical Research Communications*, 347(3), pp. 747–751. doi: 10.1016/j.bbrc.2006.06.151.
- Raman, S. *et al.* (2019) ‘Structure-guided design fine-tunes pharmacokinetics, tolerability, and antitumor profile of multispecific frizzled antibodies’, *Proceedings of the National Academy of Sciences*, 116(14), pp. 6812–6817. doi: 10.1073/pnas.1817246116.
- Read, R. J. (1986) ‘Improved Fourier coefficients for maps using phases from partial structures with errors’, *Acta Crystallographica Section A*, 42(3), pp. 140–149. doi: 10.1107/S1600536804004209.
- Reeves, P. J. *et al.* (2002) ‘Structure and function in rhodopsin: High-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-

- inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line', *Proceedings of the National Academy of Sciences of the United States of America*, 99(21), pp. 13419–13424. doi: 10.1073/pnas.212519299.
- Reichsman, F., Smith, L. and Cumberledge, S. (1996) 'Glycosaminoglycans Can Modulate Extracellular Localization of the', *J Cell Biol*, 135(3), pp. 819–827. doi: 10.1083/jcb.135.3.819.
 - Riccio, G. *et al.* (2018) 'A Negative Allosteric Modulator of WNT Receptor Frizzled 4 Switches into an Allosteric Agonist', *Biochemistry*, 57(5), pp. 839–851. doi: 10.1021/acs.biochem.7b01087.
 - Rijsewijk, F. *et al.* (1987) 'The Drosophila homology of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless', *Cell*, 50(4), pp. 649–657. doi: 10.1016/0092-8674(87)90038-9.
 - Rimkus, T. K. *et al.* (2016) 'Targeting the sonic hedgehog signaling pathway: Review of smoothened and GLI inhibitors', *Cancers*, 8(2), pp. 1–23. doi: 10.3390/cancers8020022.
 - Ring, L. *et al.* (2014) 'β-Catenin-dependent pathway activation by both promiscuous "canonical" WNT3a-, and specific "noncanonical" WNT4- and WNT5a-FZD receptor combinations with strong differences in LRP5 and LRP6 dependency', *Cellular Signalling*. Elsevier Inc., 26(2), pp. 260–267. doi: 10.1016/j.cellsig.2013.11.021.
 - Rosenbaum, D. M., Rasmussen, S. G. F. and Kobilka, B. K. (2009) 'The structure and function of G-protein-coupled receptors', *Nature*, 459(7245), pp. 356–363. doi: 10.1038/nature08144.
 - Rossman, M. G. and Blow, D. M. (1962) 'The Detection of Sub-Units Within the Crystallographic Asymmetric Unit', *Acta Crystallographica Section D Biological Crystallography*, pp. 24–31. doi: 10.1107/S0365110X62000067.
 - Saal, H. M. *et al.* (2015) 'A mutation in FRIZZLED2 impairs Wnt signaling and causes autosomal dominant omodysplasia', *Human Molecular Genetics*, 24(12), pp. 3399–3409. doi: 10.1093/hmg/ddv088.
 - Saini, S. *et al.* (2009) 'Functional significance of secreted frizzled-related protein 1 in metastatic renal cell carcinomas', *Cancer Research*, 69(17), pp. 6815–6822. doi: 10.1158/0008-5472.CAN-09-1254.
 - Sato, A. *et al.* (2010) 'Wnt5a regulates distinct signalling pathways by binding to

- Frizzled2', *EMBO Journal*. Nature Publishing Group, 29(1), pp. 41–54. doi: 10.1038/emboj.2009.322.
- Schmidt, T. G. M. *et al.* (2013) 'Development of the Twin-Strep-tag® and its application for purification of recombinant proteins from cell culture supernatants', *Protein Expression and Purification*. Elsevier Inc., 92(1), pp. 54–61. doi: 10.1016/j.pep.2013.08.021.
 - Seifert, J. R. K. and Mlodzik, M. (2007) 'Frizzled/PCP signalling: A conserved mechanism regulating cell polarity and directed motility', *Nature Reviews Genetics*, 8(2), pp. 126–138. doi: 10.1038/nrg2042.
 - Seměnov, M. V. *et al.* (2001) 'Head inducer dickkopf-1 is a ligand for Wnt coreceptor LRP6', *Current Biology*, 11(12), pp. 951–961. doi: 10.1016/S0960-9822(01)00290-1.
 - Serrano-vega, M. J. *et al.* (2008) 'Conformational thermostabilization of the β 1-adrenergic receptor in a detergent-resistant form', *Proceedings of the National Academy of Sciences*, 105(3), pp. 877–882.
 - Shafer, B. *et al.* (2011) 'Vangl2 Promotes Wnt/Planar Cell Polarity-like Signaling by Antagonizing Dvl1-Mediated Feedback Inhibition in Growth Cone Guidance', *Developmental Cell*. Elsevier Inc., 20(2), pp. 177–191. doi: 10.1016/j.devcel.2011.01.002.
 - Sharma, R. P. and Chopra, V. L. (1976) 'Effect of the wingless (wg1) mutation on wing and haltere development in *Drosophila melanogaster*', *Developmental Biology*, 48(2), pp. 461–465. doi: 10.1016/0012-1606(76)90108-1.
 - Shen, G. *et al.* (2015) 'Structural basis of the Norrin-Frizzled 4 interaction', *Cell Research*, 25(9), pp. 1078–1081. doi: 10.1038/cr.2015.92.
 - Slusarski, D. C., Corces, V. G. and Moon, R. T. (1997) 'Interaction of Wnt and a Frizzled homologue triggers G-protein-linked phosphatidylinositol signalling', *Nature*, 390(6658), pp. 410–413. doi: 10.1038/37138.
 - Smallwood, P. M. *et al.* (2007) 'Mutational analysis of Norrin-Frizzled4 recognition', *Journal of Biological Chemistry*, 282(6), pp. 4057–4068. doi: 10.1074/jbc.M609618200.
 - Steinhart, Z. *et al.* (2017) 'Genome-wide CRISPR screens reveal a Wnt-FZD5 signaling circuit as a druggable vulnerability of RNF43-mutant pancreatic tumors', *Nature Medicine*, 23(1), pp. 60–68. doi: 10.1038/nm.4219.

- Strakova, K. *et al.* (2017) ‘The tyrosine Y2502.39 in Frizzled 4 defines a conserved motif important for structural integrity of the receptor and recruitment of Disheveled’, *Cellular Signalling*. Elsevier, 38(December 2016), pp. 85–96. doi: 10.1016/j.cellsig.2017.06.018.
- Strutt, D. I., Weber, U. and Mlodzik, M. (1997) ‘The role of RhoA in tissue polarity and frizzled signalling’, *Nature*, 387(6630), pp. 292–295. doi: 10.1038/387292a0.
- Subramanian, S. and Srienc, F. (1996) ‘Quantitative analysis of transient gene expression in mammalian cells using the green fluorescent protein’, *Journal of Biotechnology*, 49(1–3), pp. 137–151. doi: 10.1016/0168-1656(96)01536-2.
- Tam, R. *et al.* (2017) ‘WNT antagonists exhibit unique combinatorial antitumor activity with taxanes by potentiating mitotic cell death’, *Science Advances*, 3(6), p. e1700090. doi: 10.1126/sciadv.1700090.
- Tao, Y. *et al.* (2019) ‘Tailored tetravalent antibodies potently and specifically activate Wnt / Frizzled pathways in cells , organoids and mice’, *eLife*, pp. 1–16.
- Tapia-Rojas, C. and Inestrosa, N. C. (2018) ‘Loss of canonical Wnt signaling is involved in the pathogenesis of Alzheimer’s disease’, *Neural Regeneration Research*, 13(10), pp. 1705–1710. doi: 10.4103/1673-5374.238606.
- Türkmen, S. *et al.* (2017) ‘A Novel de novo FZD2 Mutation in a Patient with Autosomal Dominant Omodysplasia’, *Molecular Syndromology*, 8(6), pp. 318–324. doi: 10.1159/000479721.
- Umbhauer, M. *et al.* (2000) ‘The C-terminal cytoplasmic Lys-thr-X-X-X-Trp motif in frizzled receptors mediates Wnt/beta-catenin signalling.’, *The EMBO journal*, 19(18), pp. 4944–54. doi: 10.1093/emboj/19.18.4944.
- Üren, A. *et al.* (2000) ‘Secreted frizzled-related protein-1 binds directly to wingless and is a biphasic modulator of Wnt signaling’, *Journal of Biological Chemistry*, 275(6), pp. 4374–4382. doi: 10.1074/jbc.275.6.4374.
- Usui, T. *et al.* (1999) ‘Flamingo, a seven-pass transmembrane cadherin, regulates planar cell polarity under the control of Frizzled’, *Cell*, 98(5), pp. 585–595. doi: 10.1016/S0092-8674(00)80046-X.
- Valnohova, J. *et al.* (2018) ‘Functional dissection of the N-terminal extracellular domains of Frizzled 6 reveals their roles for receptor localization and Dishevelled recruitment’, *Journal of Biological Chemistry*, 293(46), pp. 17875–17887. doi:

- 10.1074/jbc.RA118.004763.
- Voloshanenko, O. *et al.* (2017) ‘Mapping of Wnt-Frizzled interactions by multiplex CRISPR targeting of receptor gene families’, *FASEB Journal*, 31(11), pp. 4832–4844. doi: 10.1096/fj.201700144R.
 - Vonrhein, C. *et al.* (2011) ‘Data processing and analysis with the autoPROC toolbox’, *Acta Crystallographica Section D: Biological Crystallography*. International Union of Crystallography, 67(4), pp. 293–302. doi: 10.1107/S0907444911007773.
 - Wacker, D., Stevens, R. C. and Roth, B. L. (2017) ‘How Ligands Illuminate GPCR Molecular Pharmacology’, *Cell*. Elsevier Inc., 170(3), pp. 414–427. doi: 10.1016/j.cell.2017.07.009.
 - Wang, C. *et al.* (2013) ‘Structure of the human smoothened receptor bound to an antitumour agent’, *Nature*. Nature Publishing Group, 497(7449), pp. 338–343. doi: 10.1038/nature12167.
 - Wang, C. *et al.* (2014) ‘Structural basis for Smoothened receptor modulation and chemoresistance to anticancer drugs’, *Nature Communications*. Nature Publishing Group, 5, pp. 1–11. doi: 10.1038/ncomms5355.
 - Wang, Y. *et al.* (2016) ‘Frizzled Receptors in Development and Disease.’, *Current topics in developmental biology*, 117, pp. 113–139. doi: 10.1016/bs.ctdb.2015.11.028.
 - Wheeler, D. S. *et al.* (2011) ‘Direct interaction between NHERF1 and Frizzled regulates B-catenin signaling’, *Oncogene*. Nature Publishing Group, 30(1), pp. 32–42. doi: 10.1038/onc.2010.389.
 - White, J. J. *et al.* (2018) ‘WNT Signaling Perturbations Underlie the Genetic Heterogeneity of Robinow Syndrome’, *American Journal of Human Genetics*. Elsevier Company., 102(1), pp. 27–43. doi: 10.1016/j.ajhg.2017.10.002.
 - Willert, K. *et al.* (2003) ‘Wnt proteins are lipid-modified and can act as stem cell growth factors.’, *Nature*, 423(May). doi: 10.1038/nature01611.
 - Winter, G. (2010) ‘Xia2: An expert system for macromolecular crystallography data reduction’, *Journal of Applied Crystallography*, 43(1), pp. 186–190. doi: 10.1107/S0021889809045701.
 - Winter, G. *et al.* (2018) ‘DIALS: Implementation and evaluation of a new integration package’, *Acta Crystallographica Section D: Structural Biology*. International Union of Crystallography, 74, pp. 85–97. doi: 10.1107/S2059798317017235.

- Wong, H.-C. *et al.* (2003) ‘Direct binding of the PDZ domain of Dishevelled to a conserved internal sequence in the C-terminal region of Frizzled.’, *Molecular Cell*, 12(5), pp. 1251–60. doi: 10.1016/s1097-2765(03)00427-1.
- Wong, S. C. C. *et al.* (2013) ‘Clinical significance of frizzled homolog 3 protein in colorectal cancer patients’, *PLoS ONE*, 8(11), pp. 1–12. doi: 10.1371/journal.pone.0079481.
- Wright, S. C. *et al.* (2019) ‘A conserved molecular switch in Class F receptors regulates receptor activation and pathway selection’, *Nature Communications*, 10(1). doi: 10.1038/s41467-019-08630-2.
- Wu, J. and Mlodzik, M. (2008) ‘The Frizzled Extracellular Domain Is a Ligand for Van Gogh/Stbm during Nonautonomous Planar Cell Polarity Signaling’, *Developmental Cell*, 15(3), pp. 462–469. doi: 10.1016/j.devcel.2008.08.004.
- Xu, Q. *et al.* (2004) ‘Vascular Development in the Retina and Inner Ear’, *Cell*, 116(6), pp. 883–895. doi: 10.1016/s0092-8674(04)00216-8.
- Yang, S. *et al.* (2018) ‘Crystal structure of the Frizzled 4 receptor in a ligand-free state’, *Nature*, 560(7720), pp. 666–670. doi: 10.1038/s41586-018-0447-x.
- Yang, Y. and Mlodzik, M. (2015) ‘Wnt-Frizzled/Planar Cell Polarity Signaling: Cellular Orientation by Facing the Wind (Wnt)’, *Annual Review of Cell and Developmental Biology*, 31(1), pp. 623–646. doi: 10.1146/annurev-cellbio-100814-125315.
- Zebisch, M. *et al.* (2016) ‘Structure of the Dual-Mode Wnt Regulator Kremen1 and Insight into Ternary Complex Formation with LRP6 and Dickkopf’, *Structure*. The Authors, 24(9), pp. 1599–1605. doi: 10.1016/j.str.2016.06.020.
- Zeng, C. M., Chen, Z. and Fu, L. (2018) ‘Frizzled receptors as potential therapeutic targets in human cancers’, *International Journal of Molecular Sciences*, 19(5). doi: 10.3390/ijms19051543.
- Zeng, X. *et al.* (2005) ‘A dual-kinase mechanism for Wnt co-receptor phosphorylation and activation’, *Nature*, 438(7069), pp. 873–877. doi: 10.1038/nature04185.
- Zhang, X., Dong, S. and Xu, F. (2018) ‘Structural and Druggability Landscape of Frizzled G Protein-Coupled Receptors’, *Trends in Biochemical Sciences*. Elsevier Ltd, 43(12), pp. 1033–1046. doi: 10.1016/j.tibs.2018.09.002.
- Zhang, X., Gaspard, J. P. and Chung, D. C. (2001) ‘Regulation of vascular endothelial

growth factor by the Wnt and K-ras pathways in colonic neoplasia', *Cancer Research*, 61(16), pp. 6050–6054.