Structural and functional studies of Wnt signalling pathway antagonists

Tomas Malinauskas

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



Jesus College
Nuffield Department of Clinical Medicine
Division of Structural Biology
University of Oxford

Trinity Term 2011

Abstract

Wnt morphogens control embryonic development and adult tissue homeostasis. Deregulation of the Wnt signalling pathway leads to human diseases. This thesis covers work on two antagonists of the Wnt signalling pathway, Wnt inhibitory factor 1 (WIF-1) (**Chapters 1** to **5**) and secreted Frizzled-related protein 1 (sFRP-1) (**Chapters 1** and **6**). In vertebrates the N-terminal WIF domain of the six-domain WIF-1 and the cysteine-rich domain of the two-domain sFRP-1 bind to Wnt proteins and inhibit signal transduction.

My human WIF domain crystal structure reveals a novel binding site for phospholipid. Two acyl chains of the phospholipid extend deep into the domain while the lipid head group is surface exposed. Biophysical and cellular assays, combined with structure-guided mutagenesis, indicate a WIF domain Wnt-binding surface proximal to the lipid head group, but also implicate the five epidermal growth factor (EGF)-like domains (EGFs I–V) in Wnt binding. The crystal structure of the six-domain WIF-1 reveals EGFs I–V wrapped-back to interface with the WIF domain at EGF III. Binding studies locate a heparan sulfate proteoglycan-binding site in EGFs II–V, consistent with highly conserved, positively charged residues on EGF IV.

My human sFRP-1 cysteine-rich domain crystal structures reveal a conserved hydrophobic homodimer interface, conformationally variant N- and C-termini, and a conformationally invariant, putative Wnt-binding site. The site is located proximal to a positively charged surface region, which is consistent with the observed sFRP-1 cysteine-rich domain binding to heparin and heparan sulfate.

The combination of heparan sulfate proteoglycan- and Wnt-binding properties suggests a model for WIF-1 and sFRP-1 localisation, and signalling inhibition, within Wnt morphogen gradients. The findings described in this thesis provide insights into the medically-important mechanisms of Wnt signalling pathway.

Acknowledgements

Firstly, I would like to thank Prof. Yvonne Jones for presenting me with the opportunity to work in the Division of Structural Biology, her guidance and encouragement during the past four years.

I would also like to thank people who have been directly involved with my doctoral studies: Drs. Christian Siebold and Radu Aricescu for advice and guidance; Weixian Lu and Dr. Yuguang Zhao for help with tissue culture; Dr. Karl Harlos for advice on crystal handling; Prof. Chris O'Callaghan for access to surface plasmon resonance equipment; Drs. Benedikt Kessler and Katalin di Gleria for help with mass spectrometry analysis; Dr. Bert Janssen for advice on solvent flattening; Margaret Jones for help and advice in the laboratory; Christian Bell for advice and discussion.

I thank all members of the Division of Structural Biology for making this place an excellent environment to work in; Cancer Research UK for funding; Diamond Light Source and the European Synchrotron Radiation Facility for giving me the opportunity to collect X-ray diffraction data; the University of Oxford and Jesus College for their support.

Finally, I thank my family and friends for all their love, support and encouragement.

Disclaimer

My findings on Wnt inhibitory factor 1 described in this thesis are published in a journal "Nature Structural & Molecular Biology" (volume 18, pages 886–893; Malinauskas et al., 2011). Thus identical figures and parts of the text are present in both publications.

Table of Contents

Abstract **2**
Acknowledgements **3**
Disclaimer **4**
Chapter 1 – Introduction **8**
1.1 Background 8
1.2 An overview of the canonical Wnt/β-catenin signalling pathway 10
1.3 Molecular architecture of Wnt proteins 15
1.4 Extracellular distribution of Wnt proteins 16
1.5 Extracellular antagonists of Wnt signalling 18
1.5.1 Class I: Wnt-binding antagonists 19
1.5.2 Class II: LRP-binding antagonists 23
1.5.3 Class III: Frizzled- and Wnt-binding antagonists 25
1.5.4 Class IV: LRP- and Wnt-binding antagonist 27
1.5.5 Class V: LRP- and Frizzled-binding antagonist 27
1.5.6 *Drosophila* heparan sulfate proteoglycan-modifying protein Notum/Wingful 28
1.6 Aims 28
Chapter 2 – Materials and methods **30**
2.1 Molecular biology 30
2.1.1 Complementary DNA and expression vectors 30
2.1.2 Construct design and polymerase chain reaction 32
2.1.3 DNA restriction digest, ligation and transformation 33
2.2 Protein expression and purification 34
2.2.1 Small scale protein expression and Western blot analysis 34
2.2.2 Large scale expression and purification of WIF-1, sFRP-1 CRD and GPC3 36
2.2.3 Large scale expression and purification of Wnt-3a 37
2.3 Protein crystallisation 38
2.3.1 Crystal optimisation 40
2.3.2 Reductive lysine methylation 41
2.4 Crystal cryo-protection and X-ray diffraction data collection 42
2.5 X-ray diffraction data processing 43
2.6 Structure determination, refinement and validation 43
2.6.1 Structure determination and refinement 43
2.6.2 Structure validation 45
2.7 Surface plasmon resonance (SPR) binding assays 45
2.7.1 SPR binding studies of WIF-1 and Wnt-3a 45

2.7.2 SPR binding studies of sFRP-1 CRD and Wnt-3a 47
2.7.3 SPR binding studies of WIF-1 or sFRP-1 CRD and glycosaminoglycans 47
2.7.4 Preparation of GPC3 constructs 47
2.7.5 SPR binding studies of sFRP-1 CRD and GPC3 48
2.8 Mass spectrometry analysis of WIF-1-bound lipids 48
2.9 WIF-$1_{\Delta C}$–lipid interactions 49
2.10 Analytical heparin affinity chromatography 50
2.11 Cellular assays for Wnt signalling 50
2.12 Binding studies between HEK-293T or CHO cells and WIF-1 constructs 53
2.13 Multi-angle laser light scattering (MALLS) 53
2.14 Analytical ultracentrifugation 54
Chapter 3 – Crystal structures of WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ 55
3.1 Summary 55
3.2 Introduction 55
3.3 Aims 56
3.4 Expression and purification of WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ 57
3.5 Crystallisation and structure determination of WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ 59
3.5.1 Crystallisation, and structure determination of WIF-1_{WD} and mass spectrometry analysis of WIF-1-bound ligands 59
3.5.2 Crystallisation and structure determination of WIF-$1_{WD\text{-}EGF\text{-}I}$ 66
3.6 Key features of the $WIF_{WD\text{-}EGF\text{-}I}$ crystal structure 70
3.61 WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal packing 70
3.62 The WIF domain fold and surface properties 72
3.63 The WIF domain lipid-binding pocket 78
3.64 Exchange of the WIF domain-bound lipids 79
3.65 Comparison of the WIF domain structures determined by NMR spectroscopy and X-ray crystallography 82
3.66 Proteins structurally similar to the WIF domain crystal structure 87
3.67 Structure of the WIF-1 EGF-like domain I 91
3.7 Homology models of WIF domains from *Drosophila* WIF-1 (Shifted) and human RYK 93
3.8 Future perspectives 97
Chapter 4 – Cellular and binding studies of WIF-1–Wnt-3a interactions 98
4.1 Summary 98
4.2 Introduction 98
4.3 Aims 100
4.4 Expression and purification of mouse Wnt-3a 101

4.5 Expression and purification of WIF-1 constructs 103
4.5.1 Expression and purification of wild-type WIF-1 constructs 103
4.5.2 Expression and purification of mutant WIF-1 constructs 106
4.6 Both WIF and EGF-like domains contribute to Wnt-3a signalling inhibition 107
4.7 Both WIF and EGF-like domains bind to Wnt-3a 109
4.8 The Wnt-3a-binding site on the WIF domain is in the proximity of the lipid head group 113
4.9 WIF-$1_{WD\text{-}EGF\text{-}I}$ Met77Trp crystal structure highlights a potential Wnt-3a-binding site 121
4.10 Why does *Drosophila* WIF-1 (Shifted) not bind to Wnt proteins? 125
4.11 Future perspectives 131
Chapter 5 – Contribution of the EGF-like domains to WIF-1 function 132
5.1 Summary 132
5.2 Introduction 132
5.3 Aims 134
5.4 Crystallisation and structure determination of WIF-$1_{\Delta C}$ 134
5.5 The WIF-$1_{\Delta C}$ structure reveals the wrapped-back orientation of the EGF-like domains 140
5.6 WIF-1 binds to glycosaminoglycans via the EGF-like domains II–V 143
5.7 Endocytosis of WIF-1 constructs 150
5.8 Conclusions and future perspectives 153
Chapter 6 – The cysteine-rich domain of the human secreted Frizzled-related protein 1 155
6.1 Summary 155
6.2 Introduction 155
6.3 Aims 161
6.4 Crystallisation and structure determination of the sFRP-1 CRD 161
6.5 The sFRP-1 CRD forms homodimers in the crystals 167
6.6 Structural characteristics of the sFRP-1 CRD 174
6.7 The sFRP-1 CRD interactions with Wnt-3a and glycosaminoglycans 180
6.8 Conclusions and future perspectives 188
Chapter 7 – Further discussions 189
7.1 Similarities between sFRP-1 and WIF-1 189
7.2 Differences between sFRP-1 and WIF-1 190
7.3 Future directions 191
Appendices 195
References 213

Chapter 1 – Introduction

1.1 Background

The human body consists of several hundred distinct cell types and there are about 100 trillion (10^{14}) cells in the human body. All of these cells originate from one fertilized egg. The process by which a multicellular organism grows and develops is a masterpiece of temporal and spatial control of gene expression. At the cellular and molecular levels, all developmental processes are controlled by the cooperative action of different signal transduction pathways between cells. Several families of the key extracellular signalling molecules such as the bone morphogenetic proteins (BMPs), Hedgehogs, the fibroblast growth factors (FGFs), and the Wnts have been identified. The Wnt signalling pathway is essential for embryonic development and adult homeostasis (**Figure 1.1**). Wnt signalling orchestrates a variety of cellular processes, including cell proliferation, differentiation and movement. Deregulation of Wnt signalling leads to developmental disorders and diseases, including cancer.

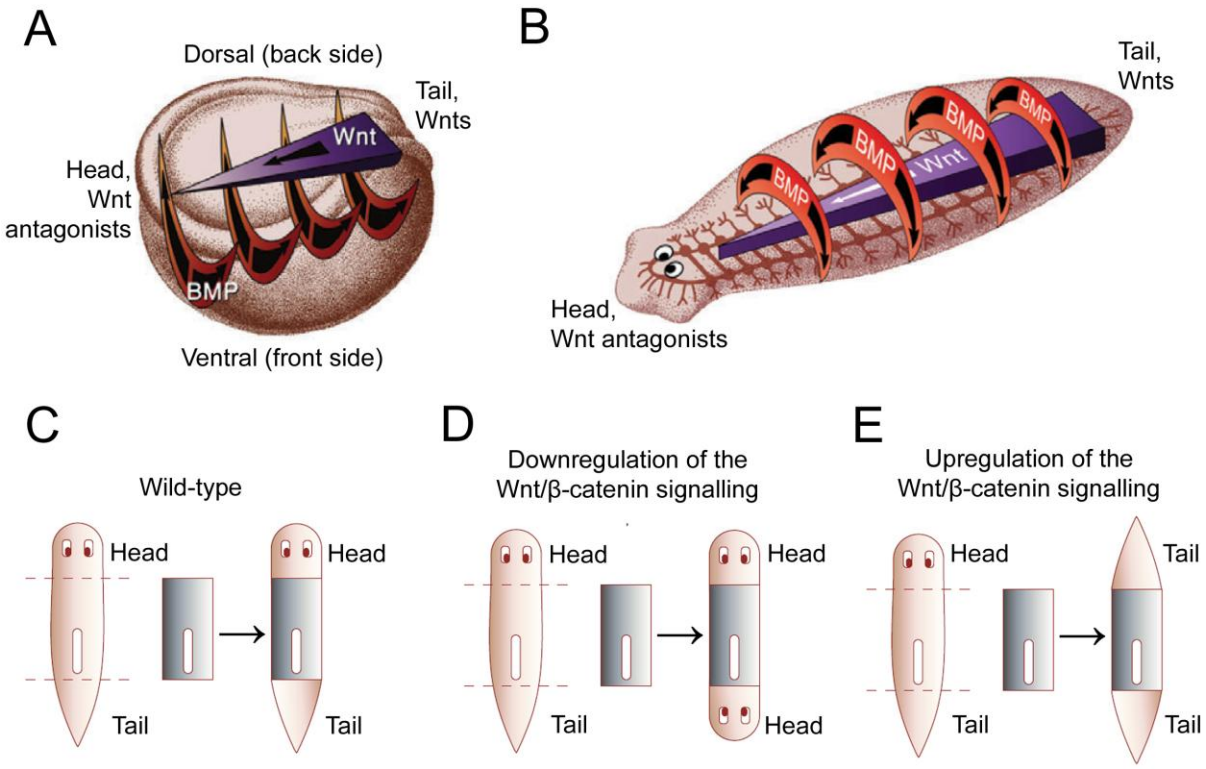


Figure 1.1. **A**, double morphogen gradient model of embryonic axis formation in vertebrates. Morphogens induce distinct cellular responses in a concentration-dependent manner (reviewed by Rogers and Schier, 2011). The model shows how perpendicular activity of Wnts and BMPs regulate anterior-posterior (head-to-tail) and dorsal-ventral (back-to-front) patterning in an amphibian embryo, respectively (reviewed by Niehrs, 2010). **B**, the double morphogen gradient model shows how perpendicular signalling of Wnts and BMPs regulate head-to-tail and back-to-front patterning in a planarian (a flatworm). **C**–**E**, Wnt/β-catenin signalling regulates head and tail formation during planarian regeneration. **C**, an entire animal can be regenerated (indicated with an arrow) from a body fragment approximately 1/300th the size of the original animal (reviewed by Reddien and Sánchez Alvarado, 2004). **D**, knockdown of the Wnt/β-catenin signalling pathway by RNA interference results in formation of an ectopic head from the posterior wound (Gurley et al., 2008; Petersen and Reddien, 2008). Similarly, during mouse embryogenesis, Wnt/β-catenin signalling also needs to be inhibited by Dkk1 (**Section 1.5.2**) for the head to form (Mukhopadhyay et al., 2001). **E**, upregulation of Wnt/β-catenin signalling after knockdown of the negative regulator APC (**Figure 1.2**) causes regeneration of two tails (Gurley et al., 2008). **Figures 1.1**, **A** and **B** are adapted from Figure 3 by Niehrs, 2010. **Figures 1.1**, **C**–**E** are adapted from Figure 1 by Tanaka and Weidinger, 2008.

1.2 An overview of the canonical Wnt/β-catenin signalling pathway

Wnt proteins activate at least three closely related signalling pathways: (1) the canonical Wnt/β-catenin pathway, (2) the Wnt/planar cell polarity (PCP) pathway, and (3) Wnt/Ca^{2+} pathway. This thesis focuses on the canonical Wnt/β-catenin signalling pathway. The key discoveries in the Wnt signalling pathway studies over the past 30 years are reflected in multiple reviews (Nusse and Varmus, 1992; Cadigan and Nusse, 1997; Logan and Nusse, 2004; Clevers, 2006; Polakis, 2007; Amerongen and Nusse, 2009; MacDonald et al., 2009; Chien et al., 2009). The β-catenin independent Wnt signalling pathways (2 and 3) are reviewed by (Seifert and Mlodzik, 2007; Wang and Nathans, 2007) and (Kühl et al., 2000), respectively. The nematode *Caenorhabditis elegans* exhibits an evolutionarily divergent Wnt/β-catenin asymmetry pathway (Phillips and Kimble, 2009).

The current model of signal transduction via the canonical Wnt/β-catenin pathway is summarised in **Figures 1.2** and **1.3**. In the absence of Wnt proteins, the cytoplasmic β-catenin, which is essential for the transcription of Wnt target genes, is kept at low concentration through proteasome-mediated degradation (**Figure 1.2**). The sequence of events leading to β-catenin degradation starts at a multiprotein complex composed of (1) scaffolding protein Axin (Zeng et al., 1997), (2) Adenomatous polyposis coli protein (APC) (Munemitsu et al., 1995), (3) casein kinase 1 α (CK1α) (Liu et al., 2002), and (4) glycogen synthase kinase 3 β (GSK3β) (Ikeda et al., 1998; Sakanaka et al., 1998) and other proteins (reviewed by Kimelman and Xu, 2006). Axin is a 96 kDa molecular weight protein and binds to CK1α, APC, GSK3β and β-catenin via separate regions. β-catenin is sequentially phosphorylated at serine 45 by CK1α (Liu et al., 2002; Amit et al., 2002), and then at threonine 41, serines 37 and 33 by GSK3β (Liu et al., 2002), creating a binding site for β-TrCP1 (β-transducin repeat-containing protein 1). β-TrCP1 is subsequently recognised by the SCF ubiquitin ligase complex, which ubiquitinates β-catenin (Aberle et al., 1997). The

ubiquitinated β-catenin leaves the phosphorylation and ubiquitination complex (also called destruction complex), and is degraded in the proteasome. The continuous destruction of the cytoplasmic β-catenin prevents its accumulation and translocation to the nucleus, keeping transcription of Wnt target genes suppressed by the transcription factors such as lymphoid enhancer-binding factor 1 (LEF) (Riese et al., 1997) and T cell specific transcription factor (TCF) (Brannon et al., 1997).

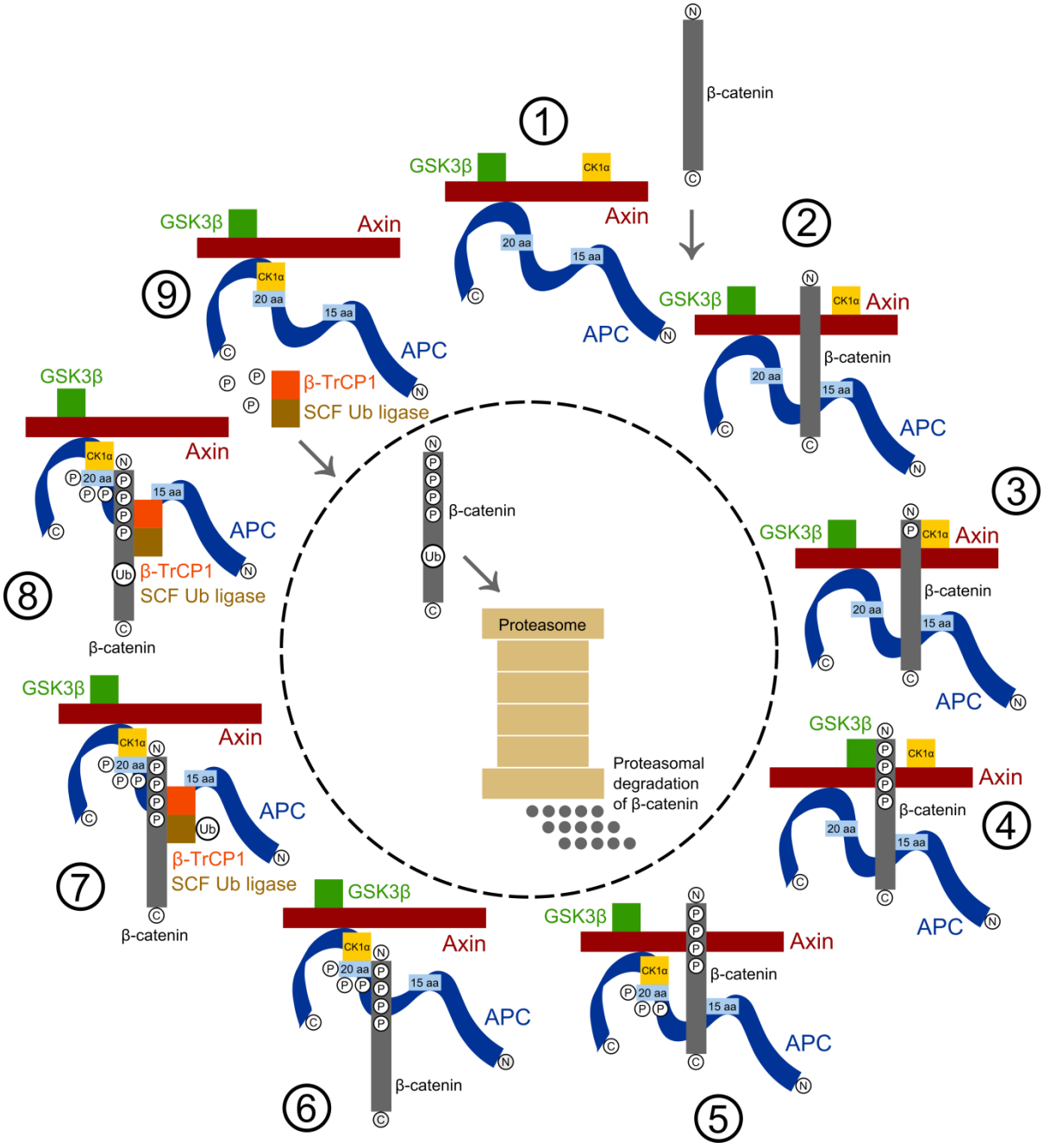

β-catenin
1
GSK3β
CK1α
Axin
20 aa
15 aa
APC
2
3
4
5
6
7
8
9
β-TrCP1
SCF Ub ligase
Ub
P
N
C
Proteasome
Proteasomal
degradation
of β-catenin

Figure 1.2. The β-catenin destruction (also called phosphorylation and ubiquitination) complex. (1) Initially, the destruction complex contains Axin (dark red), GSK3β (green), CK1α (yellow) and Adenomatous polyposis coli protein (APC, dark blue; N-, C-terminal regions, 15 and 20 amino acid (aa) residue repeat regions are indicated). (2) Subsequently, β-catenin (grey, N- and C-terminal regions are indicated) binds to Axin and 15 aa repeat region of APC. (3) β-catenin is successively phosphorylated at Ser45 by CK1α and (4) Thr41, Ser37, Ser33 by GSK3β. (5) The 20 aa repeat region of APC is phosphorylated by CK1α. (6) β-catenin dissociates from Axin and binds to the phosphorylated 20 aa repeat region of APC. (7) β-transducin repeat-containing protein 1 (β-TrCP1, orange) binds to the phosphorylated N-terminal region of β-catenin. (8) SCF ubiquitin ligase (light brown) binds to β-TrCP1 and ubiquitinates (Ub) β-catenin. (9) Molecular mechanisms of the β-catenin destruction complex disassembly remain poorly understood. (10) The phosphorylated and ubiquitinated β-catenin is degraded in the proteasome and (1) the β-catenin destruction complex assembles again. The β-catenin destruction complex contains other components, such as Axin-binding protein phosphatase 2A (Hsu et al., 1999), that were omitted for clarity. Sequence of the molecular events was adapted from Kimelman and Xu, Figure 1, 2006.

The Wnt/β-catenin pathway is activated when Wnt forms a tertiary complex with the seven-pass transmembrane receptor Frizzled (Bhanot et al., 1996) and a single pass low-density lipoprotein receptor-related protein 6 (LRP6) or LRP5 co-receptor (Tamai et al., 2000) (**Figure 1.3**). The Frizzled–Wnt–LRP5/6 complex binds to a PDZ domain of the cytoplasmic protein Dishevelled (Dvl) via the conserved C-terminal region of Frizzled (Wong et al., 2003). The Frizzled–Wnt–LRP5/6–Dvl complex further binds to the Axin–GSK3β–CK1α assembly and promotes phosphorylation of the LRP6 cytoplasmic domain (PPPSPxS motifs) by GSK3β, CK1α and the cell membrane-anchored CK1γ (Tamai et al., 2004; Davidson et al., 2005). The phosphorylated PPPSPxS motif inhibits β-catenin phosphorylation by GSK3β and provides a binding site for Axin (Piao et al., 2008; Wu et al., 2009). The Frizzled–Wnt–LRP5/6–Dvl–Axin complex may lead to further inhibition of β-catenin phosphorylation and degradation via the following model.

Both Dvl and Axin contain DIX domains. The DIX domain of Dvl2 mediates polymerization, which is essential for the Wnt signalling (Schwarz-Romond et al., 2007). DIX domains of the two proteins bind to each other through their polymerization interfaces (Fiedler et al., 2011). The Dvl–Axin–GSK3β complexes in the proximity of the cell membrane may have a

different conformation compared to the Dvl–Axin–GSK3β complexes present in the β-catenin phosphorylation assemblies. As a result, cell membrane-proximal, polymeric Dvl–Axin–GSK3β complexes might be unable to phosphorylate β-catenin. The role of GSK3β and CK1α in both Wnt signalling activation (via LRP6 phosphorylation) and inhibition (via β-catenin phosphorylation) suggests that Wnt/β-catenin signalling is a tightly regulated dynamic process.

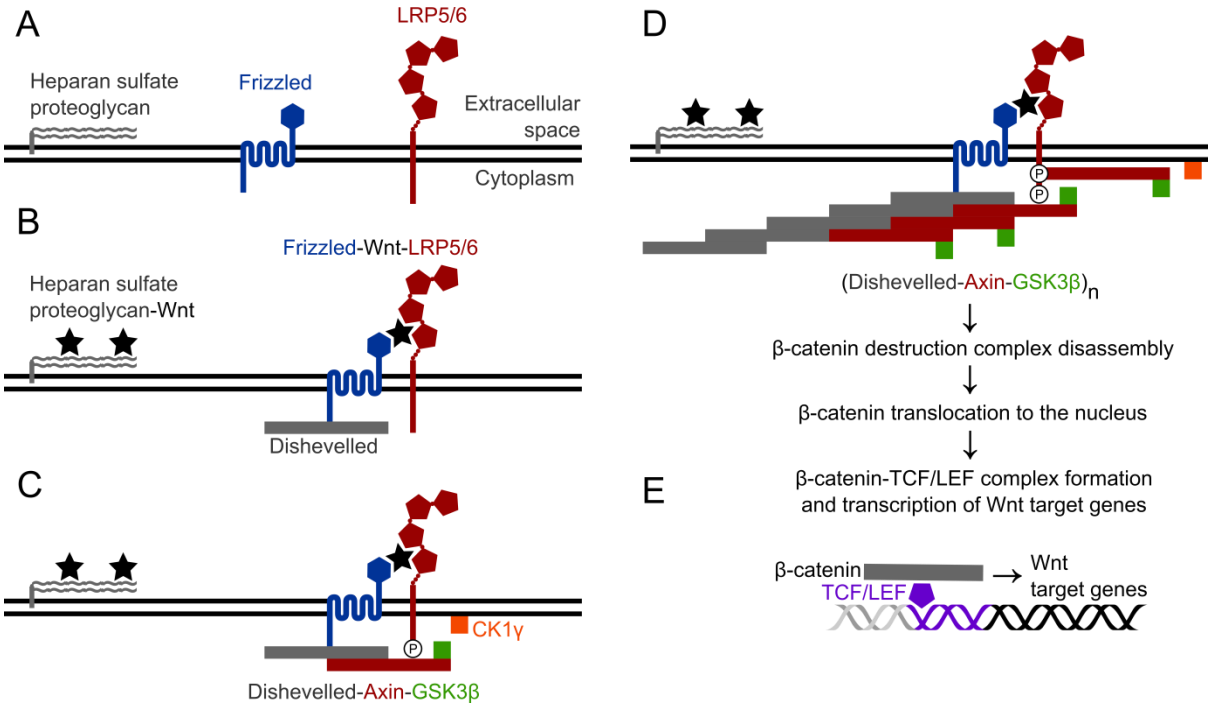


Figure 1.3. Activation of the canonical Wnt/β-catenin signalling pathway. (1) In the absence of Wnt proteins, their cell surface receptors, Frizzled (blue) and LRP5/6 (dark red), are positioned distant from each other. (2) Wnt protein (depicted as a black star) binds to the GPI-anchored, low affinity receptors heparan sulfate proteoglycans (grey). Wnt signalling is activated when Wnt is bound to Frizzled and LRP receptors simultaneously. As a result, cytoplasmic protein Dishevelled (grey) binds to the C-terminal region of Frizzled. (3) Dishevelled binds to the Axin–GSK3β complex. The cytoplasmic region of LRP is phosphorylated by the cell membrane anchored CK1γ (orange) and Axin-bound GSK3β. (4) The phosphorylated LRP region provides a high-affinity binding site for Axin, which recruits GSK3β. GSK3β phosphorylates LRP5/6 further. β-catenin destruction complex disassembles as a result of Axin–GSK3β complex recruitment to the membrane and GSK3β inhibition by the phosphorylated LRP5/6. (5) Unphosphorylated β-catenin accumulates in the cytoplasm, translocates to the nucleus, forms a complex with the TCF/LEF transcription factors (violet) and activates transcription of Wnt target genes (depicted as a black part of the DNA double helix).

In addition, the cell membrane-proximal Dvl–Axin–GSK3β complexes may compete with the β-catenin phosphorylation complexes for the Axin. Its concentration is approximately 5000-fold lower compared to other proteins present in β-catenin phosphorylation complex. Thus Axin may be a limiting molecule in the cytoplasmic Wnt signal transduction cascade (Lee et al., 2003). Low Axin concentration may also allow the Wnt signalling pathway components to participate in cellular processes only distantly related to Wnt signalling. For example, GSK3β phosphorylates NFATs, transcription factors involved in the immune response, and inhibits their translocation to the nucleus (Jope et al., 2007).

Inhibition of β-catenin degradation results in its accumulation in the cytoplasm and nucleus. Nuclear β-catenin binds to TCF/LEF family of transcription factors bound to the gene expression repressor Groucho (called TLE1 in human). TCF proteins specifically bind to a DNA sequence called the Wnt responsive element, CCTTTGWW where W is A or T. β-catenin displaces Groucho from TCF/LEF and recruits other activators of target gene transcription. Wnt target genes are diverse and cell specific. They also encode components of the Wnt signalling pathway itself that suppress or enhance Wnt/β-catenin signalling, providing mechanisms for self-regulation.

1.3 Molecular architecture of Wnt proteins

Wnt proteins are present in all metazoan animals. They contain approximately 350–400 amino acid residues (molecular weight ~40 kDa), an N-terminal secretion signal peptide (in the case of mouse Wnt-3a, 18 amino acid residues), and a conserved pattern of 22–24 cysteine residues (**Appendix 1**). The amino acid sequence of Wnts exhibits no or very limited similarity to other protein families. The domain architecture of Wnts remains unknown. Biochemical studies of Wnts are hindered by their hydrophobic character and inefficient secretion from cells which results in small amounts of purified proteins.

Mouse Wnt-3a was the first purified and biochemically characterized Wnt protein (Willert et

al, 2003). Wnt-3a is lipid modified by palmitoylation at Cys77 (Willert et al, 2003) and palmitoleylation at Ser209 (Takada et al., 2006) which explains its hydrophobic properties. Wnt-3a harbouring a Cys77Ala mutation is secreted by mouse fibroblasts similarly to the wild-type Wnt-3a but its ability to activate Wnt/β-catenin signalling is reduced (Willert et al, 2003). Wnt-3a harbouring a Ser209Ala mutation is retained in the endoplasmic reticulum (Takada et al., 2006). A putative multipass transmembrane O-acyltransferase, Porcupine, is required for the lipid modification and secretion of Wnt-3a (Takada et al., 2006) and *Drosophila melanogaster* Wnt-1 (Wingless) (Zhai et al., 2004). In contrast, *Drosophila* WntD (also called Wnt inhibitor of Dorsal protein or Wnt-8) is not palmitoylated and does not require Porcupine for its secretion (Ching et al., 2008). Recently, it has been reported that endoplasmic reticulum protein Oto attaches hydrophobic glycophosphatidylinositol (GPI)-like anchors to Wnt-1 and Wnt-3a (Zoltewicz et al., 2009), further explaining the hydrophobic properties of Wnts.

Wnt proteins often contain potential Asn-linked glycosylation sites. In the case of Wnt-3a, glycosylation precedes palmitoylation and both modifications are necessary for its secretion (Komekado et al., 2007).

1.4 Extracellular distribution of Wnt proteins

Drosophila Wingless activates target gene expression up to 10 cells away from its source of secretion (Neumann and Cohen, 1997). How do hydrophobic Wnt proteins diffuse in the aqueous extracellular space? Wingless and mouse Wnt-3a associate with lipoprotein particles outside the cell (Panáková et al., 2005). Lipoprotein particles are complexes composed of a central lipid core associated with apolipoproteins that are surrounded by a monolayer of phospholipids. In *Drosophila* larvae with reduced lipoprotein levels, Wingless fails to activate signalling over its normal range. Similarly, Wnt-3a associates with high-density lipoproteins (HDL) and low-density lipoproteins but only HDL allows Wnt-3a secretion from the mouse

fibroblasts (Neumann et al., 2009). In contrast to the wild-type Wnt-3a, Wnt-3a Cys77Ser lacking palmitate modification does not associate with lipoprotein particles and is present as a soluble protein in the aqueous media, suggesting that the second lipid modification at Ser209 is not sufficient to anchor Wnt-3a to lipoprotein particles (Neumann et al., 2009). Recently, the third lipid modification of Wnts has been reported; the endoplasmic reticulum (ER) protein Oto promotes addition of glycophosphatidylinositol (GPI)-like anchors to Wnt-3a and Wnt-1, and mediates Wnt retention in the ER (Zoltewicz et al., 2009). How GPI-like anchor affects Wnt-3a distribution, remains unknown.

The extracellular distribution of Wnts is further regulated by proteoglycans that are located in the extracellular matrix or GPI-anchored to the cell membrane (reviewed by Häcker et al., 2005; Yan and Lin, 2009). Proteoglycans consist of a protein core to which glycosaminoglycan (GAG) chains are covalently linked. GAGs are single chain polysaccharides that consist of repeating disaccharides containing a uronic acid and an amino sugar (reviewed by Rabenstein, 2002). Disaccharides of glucuronic acid and N-acetylglucosamine form heparan sulfate (HS). GPI-anchored heparan sulfate proteoglycans (HSPGs) contain 2–3 HS chains covalently linked near the C-terminus of the protein core, close to the extracellular side of the cell membrane.

Several lines of evidence suggest that HSPGs are essential for the extracellular Wnt distribution. First, mutations of putative glycosyltransferases, affecting the biosynthesis of HS, reduce *Drosophila* Wingless signalling activity and result in the accumulation of Wingless in the front of mutant cells (Takei et al., 2004). Second, a secreted α/β hydrolase Notum (also called Wingful) influences Wingless distribution: overexpression of Notum blocks Wnt signalling while loss of Notum enhances Wnt signalling (Giráldez et al., 2002). Notum shapes Wnt distribution by modifying HSPG Dally-like protein (Dlp), likely by deacetylating N-acetylglucosamine units. Third, *Drosophila* with hypomorphic *dally*

(encoding HSPG Dally) alleles exhibit a phenotype that is similar to those resulting from a loss of Wingless signalling (Lin and Perrimon, 1999). Consistent with HSPG-related defects in *Drosophila*, humans with a mutant gene encoding HSPG GPC3 suffer from Simpson-Golabi-Behmel syndrome (Pilia et al., 1996). Frameshift mutations in the human gene encoding EXT1, an enzyme involved in HS biosynthesis, lead to hereditary multiple exostoses, disorder that is characterized by short stature and benign bone tumours (Ahn et al., 1995). In addition, HSPGs bind to lipoprotein particles (Camejo et al., 1998) that have been shown to transport Wnt-3a between mammalian cells (Neumann et al., 2009), highlighting that extracellular Wnt transport involves a complex interplay between lipid-modified Wnts, lipoprotein particles and HSPGs.

1.5 Extracellular antagonists of Wnt signalling

The Wnt signalling pathway is regulated at multiple steps within and outside the cell. An extensive network of secreted and cell membrane-anchored modulators (agonists and antagonists) shape Wnt signalling in the extracellular space (**Figure 1.4**). Several families of extracellular proteins antagonise the Wnt/β-catenin signalling pathway. Wnt antagonists can be divided into five broad classes based on the mechanism of action.

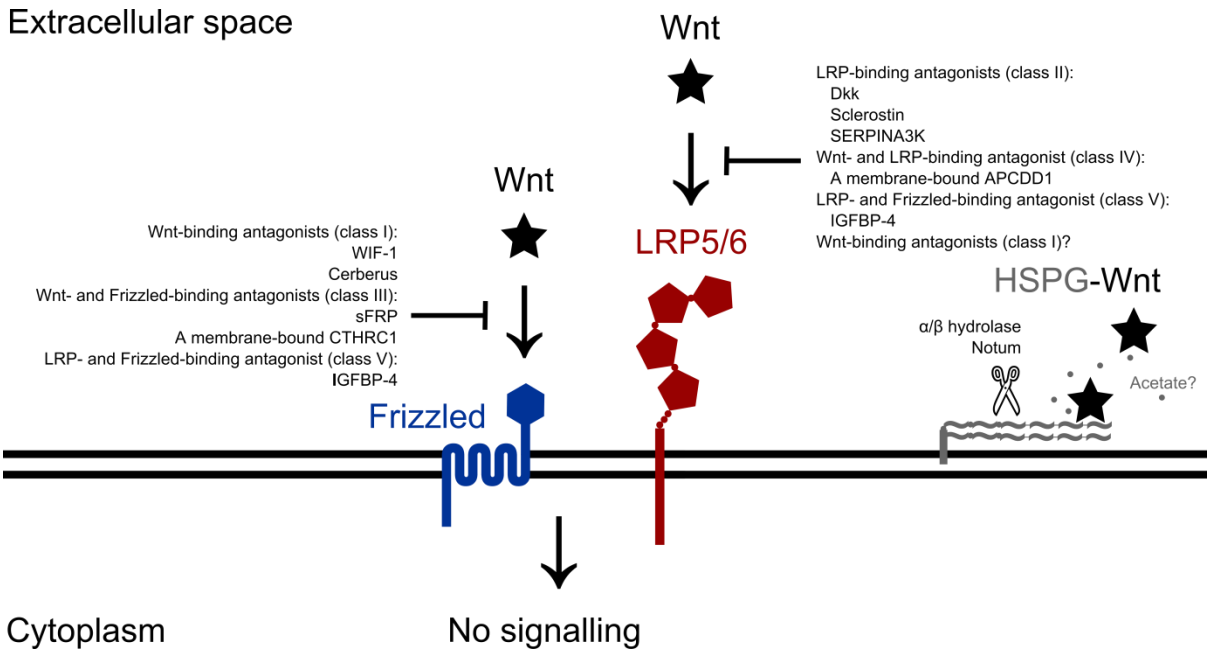


Figure 1.4. Extracellular antagonists of Wnt signalling. Extracellular antagonists are cell membrane-anchored (such as APCDD1 and CTHRC1) or secreted (all the rest) proteins. They can be classified into five broad classes based on the mechanism of action. All of them, except Notum, prevent assembly of the key signal transduction complex composed of Frizzled, Wnt and LRP5/6. No signal is transduced into the cell as a result. Notum downregulates Wnt signal transduction by modulating Wnt protein distribution on the HSPG-covered cell surface.

1.5.1 Class I: Wnt-binding antagonists

Members of the first class of Wnt antagonists include Wnt inhibitory factor 1 (WIF-1) and Cerberus. Both WIF-1 and Cerberus bind to Wnts and prevent them from triggering signalling via assembly of the Frizzled–Wnt–LRP5/6 complex (Piccolo et al., 1999; Hsieh et al., 1999).

1.5.1.1 Cerberus

Cerberus is a multivalent antagonist that binds to two members of the transforming growth factor-β family members (Nodal and BMP) and Wnt proteins via independent sites, and inhibits these three signalling pathways simultaneously (Piccolo et al., 1999). During the embryonic development of *Xenopus*, Cerberus is highly expressed in the Spemann's organizer, a group of cells that induces multiple tissue types and leads to anteroposterior (head-to-tail) polarity (Bouwmeester et al, 1996). Injection of mRNA encoding Cerberus into *Xenopus* embryos induces ectopic heads, and duplicated hearts and livers, suggesting a role for Cerberus in the induction of head structures in the vertebrate embryo. Cerberus is a member of the DAN/Cerberus protein family. Members of the DAN/Cerberus family share a conserved C-terminal cysteine knot (CT) domain (Pearce et al., 1999). The CT domain of Cerberus is composed of approximately 90 amino acid residues and is similar to the C-terminal domain of TGF-β proteins (Meitinger et al., 1993). The CT domains have nine cysteine residues that are predicted to form four intramolecular and one intermolecular disulfide bridges. Secreted mouse Cerberus 1 and *Xenopus* Cerberus form dimers, most likely covalently linked via the intermolecular disulfide bond (Biben et al., 1998; Pearce et al.,

1999). The CT domain contains overlapping Nodal and BMP binding sites. Among the DAN/Cerberus protein family, *Xenopus* Cerberus contains the longest (approximately 90 amino acid residues) N-terminal polypeptide chain which is responsible for Wnt binding (Piccolo et al., 1999). Thus it remains unknown whether Wnt binding and signalling inhibition is conserved among other DAN/Cerberus protein family members or is specific to *Xenopus* Cerberus.

1.5.1.2 Wnt inhibitory factor 1 and other WIF domain containing proteins

1.5.1.2.1 Wnt inhibitory factor 1

Wnt inhibitory factor 1 is evolutionarily conserved across vertebrate families (**Appendix 2**). During embryogenesis in *Xenopus* and zebrafish, expression of WIF-1 is first detectable at the start of somitogenesis in the paraxial mesoderm, which develops into the axial skeleton, skeletal muscle, and part of the skin (Hsieh et al., 1999; Gilbert, 2000). WIF-1 expression continues in adults, and for the mouse is highest in the heart and lungs (Hsieh et al., 1999). Moreover, it is highly expressed at cartilage-mesenchyme interfaces and prevents Wnt-3a-mediated inhibition of chondrogenesis in chicken (Surmann-Schmitt et al., 2009). Targeted disruption of the WIF-1 gene accelerates development of radiation-induced osteosarcomas in mouse (Kansara et al., 2009). Downregulation of WIF-1 expression by gene promoter methylation has been reported in different human cancers (**Table 1.1**), while overexpression of WIF-1 inhibits lung and bladder cancer cell growth (Kim et al., 2007; Tang et al., 2009).

Table 1.1. WIF-1 gene promoter methylation and its expression downregulation in cancer.

Number	Cancer	Other observations	Reference
1.	Acute promyelocytic leukemia (APL)	WIF-1 was preferentially methylated in APL and not in other subtypes of acute myeloid leukemia.	Chim et al., 2006
2.	Astrocytoma	Expression of WIF-1 gene was decreased with the increase of pathological grade. No correlation was found between WIF-1 expression and age or gender.	Yang et al., 2010
3.	Bladder	Transfection of WIF-1 small interfering RNA into bladder tumor cells accelerated their growth.	Urakami et al., 2006
4.	Breast	Methylation of the WIF-1 promoter is attributed to the DNA methyltransferases DNMT1 and DNMT3B.	Ai et al., 2006
5.	Gastrointestinal	Transfection of the WIF-1 gene construct into esophageal and colon cancer cell lines inhibited cell proliferation.	Taniguchi et al., 2005
6.	Kidney	No methylation of WIF-1 gene promoter in normal kidney. WIF-1 induced apoptosis and suppressed tumor growth *in vivo*.	Kawakami et al., 2009
7.	Lung	WIF-1 expression was restored after DNA demethylation in cell lines lacking WIF-1 expression.	Mazieres et al., 2004
8.	Nasopharyngeal (NPC) and esophageal squamous cell (ESCC) carcinomas	WIF-1 expression in NPC and ESCC tumor cells inhibits tumor cell colony formation.	Chan et al., 2007
9.	Osteosarcoma	WIF-1 overexpression decreased tumor growth rate and reduced the number of lung metastasis in mouse.	Rubin et al., 2010
10.	Pleural mesothelioma	Methylation in 4 of 4 mesothelioma cell lines, but not in 2 normal mesothelial cell lines.	Batra et al., 2006
11.	Prostate	WIF-1 expression reduced tumor growth in mouse.	Yee et al., 2010

Human WIF-1 is a secreted glycoprotein of 379 amino acid residues, composed of a secretion signal sequence (residues 1–28), WIF domain (~38–177), five epidermal growth factor (EGF)-like domains and a hydrophilic C-terminus (~339–379). WIF-1 inhibits signal transduction by sequestering Wnt proteins from their receptors. It antagonises signal transduction by *Xenopus* Wnt-8 more than *Drosophila* Wingless, while *Xenopus* Wnt-3a is only weakly inhibited (Hsieh et al., 1999). The WIF domain is as effective an inhibitor of the Wnt signalling pathway as the full length WIF-1 (Hsieh et al., 1999). However, EGF-like domains are necessary for full activity of WIF-1 as was shown using Wnt-dependent axis induction in *Xenopus* embryos as an assay: co-injection of Chordin (an inhibitor of the BMP signalling pathway) and WIF-1 RNA resulted in embryos with a complete secondary axis with two separated eyes (Hsieh et al., 1999). In contrast, co-injection of Chordin and WIF domain RNA did not promote two-eyed ectopic heads. Thus full length WIF-1 (but not WIF domain alone) synergizes with Chordin to generate a complete secondary axis in *Xenopus* embryos. How the five EGF-like domains contribute to WIF-1 function has remained unknown.

The three-dimensional structure of the WIF domain of WIF-1 was determined by NMR spectroscopy (Liepinsh et al., 2006). These structural studies were performed on a refolded WIF domain expressed in *Escherichia coli*. The refolded WIF domain consists of a nine-stranded β sandwich and two α helices, with two of the β strands interrupted by 4–6 residues of a disordered region. The N- and C-terminal regions are located at the opposite ends of the WIF domain indicating that the WIF domain may function as an independent structural module in different proteins. Alkyl chains of the detergent Brij-35 ($C_{12}H_{25}(OCH_2CH_2)_nOH$, n~23) used during refolding were found to bind tightly to the exterior of the domain. The observed detergent binding suggested that WIF-1 may bind to lipid modified Wnt proteins using the WIF domain as a lipid recognition module (Malinauskas, 2008; Bazan and de

Sauvage, 2009). Prior to starting my doctoral studies in Oxford, I found that computational docking methods indicated that the WIF domain binds to fatty acids with alkyl chain length of 16 to 18 carbon atoms with the highest affinity (Malinauskas, 2008). Whether WIF-1 is a lipid-binding protein and how refolding may have affected the WIF domain structure remained unknown.

1.5.1.2.2 Other WIF domain containing proteins: RYKs and Shifted

WIF domains serve as the extracellular modules of mammalian single-pass transmembrane RYK (*r*elated to t*y*rosine *k*inase) receptor tyrosine kinases and RYK homologues encoded by the *Drosophila* genes *doughnut*, *derailed*, and *derailed2* (Patthy, 2000; reviewed by Fradkin et al., 2010). It was shown that protein *Drosophila* Derailed binds to Wnt-5 and is required for axon guidance (Yoshikawa et al., 2003). Similarly, mammalian RYK functions as a Wnt co-receptor together with Frizzled and activates the Wnt/β-catenin pathway by direct binding to Dishevelled in neurites (Lu et al., 2004).

Notably, the *Drosophila* ortholog of WIF-1, called Shifted, binds to dually lipid-modified Hedgehog signalling proteins and modulates their diffusion without affecting the Wnt signalling pathway (Glise et al., 2005; Gorfinkiel et al., 2005). In contrast to human WIF-1, both WIF domain and EGF-like domains of Shifted are essential for its function. Structural differences between Shifted and WIF-1 that are responsible for their specificities are not known. Thus structural studies of human WIF-1 and their ligands may shed light on the molecular mechanisms behind modulation of two evolutionarily conserved, Wnt and Hedgehog, signalling pathways simultaneously.

1.5.2 Class II: LRP-binding antagonists

Members of the second class of Wnt antagonists include Dickkop (Dkk), Sclerostin and SERPINA3K that bind to LRPs and prevent Frizzled–Wnt–LRP5/6 complex formation.

Dkk was discovered as a protein secreted by Spemann's organizer during early development

as a head inducer (Glinka et al., 1998) in *Xenopus*. Dkk-1 gene deficient mice embryos lack head structures anterior of the midbrain, suggesting that Dkk-1 is essential for head formation (Mukhopadhyay et al., 2001). Members of the human Dkk protein family (Dkk-1, -2, -3, and -4) contain two cysteine-rich domains (Cys-1 and Cys-2) separated by a linker region. Antagonism of Wnt/β-catenin signalling requires binding of the Dkk Cys-2 to the third β-propeller domain of LRP5/6 (Chen et al., 2008). Cys-2 is highly conserved among Dkk-1, -2 and -4, and shares approximately 60 % amino acid sequence identity; Dkk-3 Cys-2 is only 27 % sequence identical to Cys-2 from Dkk-1, -2, and -4 (Semenov and He, 2000). The Dkk Cys-2 structure shares some features with colipase (Breg et al., 1995) and mamba intestinal toxin 1 (Boisbouvier et al., 1995); they lack extensive secondary structures and can be described as an assembly of extended loops held together by five disulfide bridges (Chen et al., 2008). Colipases interact with lipid micelles (van Tilbeurgh et al., 1999), suggesting a possible interaction between Wnt-transporting lipoprotein particles and Dkks. The function of Cys-1 remains poorly defined (Semenov and He, 2000).

Dkk-1 also forms a ternary complex with a single-pass transmembrane protein Kremen and LRP6 leading to endocytosis, which results in LRP6 removal from the cell surface and Wnt/β-catenin signalling inhibition (Mao et al., 2002).

Sclerostin is secreted by osteocytes and is an inhibitor of osteoblastogenesis (Silverman, 2010). Humans homozygous for a defective Sclerostin gene suffer from a disorder (sclerosteosis) characterized by high bone density (Hamersma et al., 2003). Deletion of the Sclerostin gene in mice leads to increased bone strength (Li et al., 2008). Thus therapies based on the inhibition of Sclerostin might be used to restore skeletal bone strength in patients suffering from osteoporosis, a disease characterized by bone loss.

Sclerostin contains approximately 190 amino acid residues and is a member of the DAN/Cerberus protein family (Veverka et al., 2009). Structures of Sclerostin determined by

NMR spectroscopy revealed a canonical cysteine-knot (CT) backbone topology observed in numerous secreted growth factors (TGF-β2, chorionic gonadotropin, BMP7, etc.) (Veverka et al., 2009; Weidauer et al., 2009). Two pairs of twisted antiparallel β strands form an extended rod like structure with three disulfide bonds at the bottom and one at the top. N- and C-terminal regions are flexible and dispensable for the Sclerostin-mediated Wnt/β-catenin signalling inhibition (Weidauer et al., 2009). Sclerostin binds to both LRP5 and 6 (Li et al., 2005; Ellies et al., 2006). However, LRP5–Sclerostin interactions at the atomic level remain poorly defined. The putative Sclerostin binding site is located on the first β-propeller of LRP5, predicted to form a negatively charged surface around Gly171 (Ellies et al., 2006). Interestingly, a prominent positively charged surface on Sclerostin is involved in heparin binding (Veverka et al., 2009), raising a question whether the negatively charged glycosaminoglycans and negatively charged region of LRP5 β-propeller 1 compete for the positively charged Sclerostin binding.

SERPINA3K is the least studied LRP-binding antagonist of Wnt/β-catenin signalling. SERPINA3K is a *ser*ine *p*roteinase *in*hibitor (SERPIN) discovered as a plasma proteinase kallikrein-binding protein (Chao et al., 1986). SERPINA3K functions as an angiogenesis (blood vessel formation) inhibitor independently of its interaction with kallikrein (Gao et al., 2003). It has been shown recently that the anti-angiogenetic activity of SERPINA3K might be linked to the inhibition of Wnt/β-catenin signalling (Zhang et al., 2010). SERPINA3K binds to the extracellular domain of LRP6 and blocks its association with the cysteine-rich domain of Frizzled8 via Wnt-1 (Zhang et al., 2010). Although less efficiently than Dkk1, SERPINA3K also inhibits *Xenopus* axis duplication induced by Wnt-8 (Zhang et al., 2010).

1.5.3 Class III: Frizzled- and Wnt-binding antagonists

Members of the third class of Wnt antagonists are secreted Frizzled-related proteins (sFRPs) that bind to Frizzled transmembrane receptors and Wnts. The family of sFRPs contains five

members in humans (sFRPs-1 to -5) and orthologs in other vertebrate species. Crescent (Pfeffer et al., 1997) and Sizzled (Salic et al., 1997) are two additional sFRP family members present in birds and amphibians but not in mammals (**Appendix 3**). Similarly to Dkk, DAN/Cerberus and WIF-1 Wnt antagonist families, sFRPs are not present in *Drosophila*. However, sFRPs are present in other invertebrates such as the purple sea urchin (Lapraz et al., 2006) or freshwater sponge *Lubomirskia baicalensis* (Adell et al., 2007), highlighting the ancient evolutionary origin of the sFRP family.

The first family member sFRP-3 (also named FrzB, Frz from Frizzled motif and B from bone development) was discovered in the long bones during embryonic and fetal development as a protein with a potential role in skeletal morphogenesis (Hoang et al., 1996). FrzB sequence similarity to *Drosophila* and rat Frizzleds was noted. Following identification of *Drosophila* Frizzled 2 as a Wingless receptor (Bhanot et al., 1996), it was found that *Xenopus* FrzB-1 antagonizes the effects of *Xenopus* Wnt-8 and co-immunoprecipitates with Wnts (Leyns et al., 1997; Wang et al., 1997); sFRP-2 and sFRP-3 derivatives, GPI-anchored to the cell surface, bind to *Drosophila* Wingless (Rattner et al., 1997).

sFRP family members contain approximately 300 amino acid residues, harboring an N-terminal secretion signal, followed by the cysteine-rich domain (CRD) and netrin-like domain. CRD are present in sFRPs, the extracellular region of transmembrane Frizzled receptors and other proteins (e.g. receptor tyrosine kinases RORs, MuSK, collagen XVIIIa, the Hh signalling pathway modulators Smoothened and *H*h-*i*nteracting *p*rotein (Hip)), and were identified as Wnt-binding domains (Lin et al., 1997). sFRPs could be classified into two groups based on the CRD similarity: (1) sFRP-3 and -4 contain CRDs that are similar to those present in Frizzled receptors, (2) sFRPs-1, -2 and -5 contain CRDs that are more similar to non-Frizzled proteins (Semenov and He, 2000). Classification based on the netrin-like domains yields the same two sFRP groups.

Crystal structures of the sFRP-3 and Frizzled-8 CRDs revealed a predominantly α helical fold with all 10 cysteines forming disulfide bonds (Dann et al., 2001). The sFRP-3 and Frizzled-8 CRDs form homologous dimers with a highly complementary dimer interface. However, both purified CRDs are monomeric in solution at 100 μM protein concentration, suggesting a relatively weak affinity between the two CRDs. Site-directed mutagenesis studies suggest a region of 950 $Å^2$ on the CRD surface as a putative Wnt binding site (Dann et al., 2001). A putative Wnt-binding surface overlaps with the interface of the CRD dimer observed in the crystals and suggests a mechanism in which CRD–CRD complex formation may interfere with Wnt binding (Dann et al., 2001).

1.5.4 Class IV: LRP- and Wnt-binding antagonist

The only member of the fourth class of Wnt antagonists is a protein encoded by a Wnt target gene (**Figure 1.3**), adenomatosis polyposis down-regulated 1 (APCDD1) (Takahashi et al., 2002). APCDD1 binds to Wnt-3a and LRP5 *in vitro* (Shimomura et al., 2010). Human APCDD1 is abundantly expressed in hair follicles and is predicted to be a single-pass transmembrane, cytoplasmic domain-lacking protein composed of 514 amino acid residues (Shimomura et al., 2010). Mutation Leu9Arg in the secretion signal peptide of APCDD1 is associated with hair loss and hair follicle miniaturization (Shimomura et al., 2010) in human whereas the wild-type exogenous APCDD1 promotes growth of colon cancer cells (Takahashi et al., 2002).

1.5.5 Class V: LRP- and Frizzled-binding antagonist

The only member of the fifth class of Wnt antagonists is insulin-like growth-factor-binding protein 4 (IGFBP-4) (Zhu et al., 2008). Members of the IGFBP family bind to insulin-like growth factors (IGFs) and sequester them from IGF receptors, leading to inhibition of mitogenesis, differentiation and other cellular processes (reviewed by Firth and Baxter, 2002). IGFBP-4 binds to IGFs I–II (Qin et al., 1998) but also inhibits the Wnt/β-catenin

signalling pathway. IGFBP-4 is required for the development of the heart in *Xenopus* and promotes mouse cardiomyocyte differentiation in an IGF-independent manner (Zhu et al., 2008). IGFBP-4 binds to the extracellular domains of LRP6 and Frizzled8, preventing Frizzled-8–Wnt-3a–LRP6 signalling complex assembly (Zhu et al., 2008).

1.5.6 *Drosophila* heparan sulfate proteoglycan-modifying protein Notum/Wingful

No homologues of the vertebrate secreted Wnt antagonists described above (class I to V) have been identified in *Drosophila* to date. A novel *Drosophila* Wingless signalling antagonist called Notum (also called Wingful) was identified in high throughput genetic screens (Giráldez et al., 2002; Gerlitz and Basler, 2002). In contrast to most vertebrate Wnt antagonists, Notum does not bind to Wingless and its receptors Frizzled-2 and Arrow (*Drosophila* ortholog of vertebrate LRP5/6) (Gerlitz and Basler, 2002). Loss of Notum function increases Wingless signalling while overexpression of Notum blocks it (Giráldez et al., 2002). Notum is an extracellular α/β hydrolase with amino acid sequence similarity to plant pectin acetyltransferases. Most likely, it deacetylates glycosaminoglycan (GAG) chains present on HSPGs called Dally and Dally-like protein (Dlp), thereby altering GAG-Wingless interactions and extracellular distribution of Wingless (Giráldez et al., 2002; Gerlitz and Basler, 2002; Baeg et al., 2001; **Figure 1.4**). In addition, Wingless signalling induces Notum expression. Thus it has been suggested that Wingless controls its own distribution by regulating expression of Notum that modulates Wingless–HSPG interactions on the cell surface (Giráldez et al., 2002; Gerlitz and Basler, 2002).

1.6 Aims

As discussed above, Wnt/β-catenin signalling orchestrates a variety of fundamental biological processes such as embryonic development, adult homeostasis and cancer. Many medically important molecular mechanisms of Wnt/β-catenin signalling and its antagonism remain

unknown. The main aim of this thesis is to clarify molecular mechanisms of Wnt/β-catenin signalling antagonism by Wnt inhibitory factor 1 and secreted Frizzled-related protein 1. X-ray crystallography in combination with a range of biophysical techniques and cellular assays serves as the major method to reveal molecular interactions.

Chapter 2 – Materials and methods

2.1 Molecular biology

2.1.1 Complementary DNA and expression vectors

Complementary DNA (cDNA) encoding human Wnt inhibitory factor 1 (WIF-1; Swiss-Prot identification number Q9Y5W5; isoform AAH18037, Gln166Lys) and zebrafish WIF-1 (Swiss-Prot Q9W6F9) clones were purchased from Source BioScience (human WIF-1: clone 4801033, Mammalian Gene Collection (MGC) number 26240, IMAGE 4801033; zebrafish WIF-1: clone 7150438, MGC number 110507, IMAGE 7150438) (Gerhard et al., 2004). A vector encoding human secreted Frizzled related protein 1 (sFRP-1; Swiss-Prot Q8N474, amino acid residues 32–176) was a kind gift from Drs. Radu Aricescu and Christian Siebold (University of Oxford, UK).

Vector pHLsec was used for protein expression in mammalian cell lines (Aricescu et al., 2006). DNA encoding different WIF-1 and sFRP-1 constructs (**Table 2.1**) were cloned into pHLsec using the AgeI and KpnI restriction sites as described in **Section 2.1.2** and standard protocols (Sambrook and Russell, 2000). The resulting constructs encode Kozak sequence at the 5’ end, followed by an N-terminal secretion signal, target protein, and a C-terminal hexahistidine (His_6) tag. During protein secretion, the N-terminal secretion signal is cleaved and the resulting protein contains the N-terminal Glu-Thr-Gly and the C-terminal Gly-Thr-Lys-His_6 tags. Vector maps are presented in the Supplementary information by Aricescu et al., 2006.

Table 2.1. Selected protein constructs and PCR primers.

Protein	Construct	Amino acids	Forward primer (5' to 3' direction)	Reverse primer (5' to 3' direction)	Restriction enzymes
Human WIF-1	WIF-$1_{\text{Full length}}$	Gly29–Trp379	CGCACCGGTGGGCCGCCGCAGGAGGAGA	GCGGGTACCCCAGATGTAATTGGATTCA	AgeI + KpnI
	WIF-$1_{\Delta C}$	Ser35–Ile346	CGCACCGGTAGCCTGTACCTATGGATCG	GCGGGTACCTATGAGGCTGGCTTCGTACC	AgeI + KpnI
	WIF-$1_{\text{WD-EGFs I-III}}$	Ser35–Glu274	CGCACCGGTAGCCTGTACCTATGGATCG	GCGGGTACCTTCACACTGCTCTCCCTC	AgeI + KpnI
	WIF-$1_{\text{WD-EGF I}}$	Ser35–Glu210	CGCACCGGTAGCCTGTACCTATGGATCG	GCGGGTACCCTCACAGTGAGGTCCGTGG	AgeI + KpnI
	WIF-1_{WD}	Ser35–Gln178	CGCACCGGTAGCCTGTACCTATGGATCG	GCGGGTACCTTGACATGTTTTAAAGAAGATAG	AgeI + KpnI
	WIF-$1_{\text{EGFs I-V}}$	Gln178–Ile346	CGCACCGGTAGCCTGTACCTATGGATCG	GCGGGTACCTATGAGGCTGGCTTCGTACC	AgeI + KpnI
	WIF-$1_{\Delta C}$ Asp145Asn/Val147Ser*	Ser35–Ile346	TCCCATGTCTTGGAAAACAGAATGGGTCCGCAGCATTTGAAGTGGATGT	ACATCCACTTCAAATGCTGCGGACCCATTCTGTTTTCCAAGACATGGGA	AgeI + KpnI
	WIF-$1_{\Delta C}$ Glu151Asn/Asp153Ser*	Ser35–Ile346	AGGATGGGGTGGCAGCATTTAATGTGTCCGTGATTGTTATGAATTCTGA	TCAGAATTCATAACAATCACGGACACATTAAATGCTGCCACCCCATCCT	AgeI + KpnI
	WIF-$1_{\Delta C}$ Phe51Asn/Glu53Ser*	Ser35–Ile346	AGGCAAGAGTACTCATAGGAAATGAATCAGATATCCTGATTGTTTCGGA	TCCGAAACAATCAGGATATCTGATTCATTTCCTATGAGTACTCTTGCCT	AgeI + KpnI
	WIF-$1_{\Delta C}$ Phe174Asn*	Ser35–Ile346	CACCTCAAAATGCTATCTTCAATAAAACATGTCAACAAGCTGA	TCAGCTTGTTGACATGTTTTATTGAAGATAGCATTTTGAGGTG	AgeI + KpnI
	WIF-$1_{\Delta C}$ Met77Trp*	Ser35–Ile346	TCAGAAAAGCGCAACAGAGATGGCCAGCTATTCCTGTCAATAT	ATATTGACAGGAATAGCTGGCCATCTCTGTTGCGCTTTTCTGA	AgeI + KpnI
Zebrafish WIF-1	Zebrafish WIF-$1_{\Delta C}$	Thr33–Asn336	CGCACCGGTACCATGTACATGTGGATTG	GCGGGTACCATTGCAGTGACGTCCGTGC	AgeI + KpnI
Human secreted Frizzled-related protein 1 (sFRP-1)	sFRP-$1_{32–176}$	Ser32–Glu176	CGCACCGGTAGCGAGTACGACTACGTGA	GCGGGTACCTTCGGTGGCATTGGGCGGC	AgeI + KpnI
	sFRP-$1_{52–176}$	Tyr52–Glu176	CGCACCGGTTACACCAAGCCACCTCAGTGCGTGG	GCGGGTACCTTCGGTGGCATTGGGCGGC	AgeI + KpnI

* Only internal primers to generate point mutations for these constructs are listed. External primers are the same as for the WIF-$1_{\Delta C}$ construct.
Enzyme restriction sites and codons encoding mutated amino acids are colour-coded: AgeI site – green, KpnI site – grey, Asn codon – yellow, Ser codon – blue, Trp codon – orange.

2.1.2 Construct design and polymerase chain reaction

Constructs for the sFRP-1 and WIF-1 proteins were designed taking into account domain boundaries listed on the UniProtKB/Swiss-Prot Internet protein database (UniProt Consortium, 2010) and data from the protein disorder prediction program RONN (Yang et al., 2005). Designed constructs were intended to encode the full length protein domains without disturbing their secondary and tertiary structure. Constructs designed for protein crystallisation omit any extensive regions predicted to be disordered by RONN.

DNA encoding the target protein was amplified using polymerase chain reaction (PCR) (Mullis et al., 1986). Typically, 50 µl PCR reaction mixture included: 1 µl (200 ng) template cDNA, 0.25 µl Pyrobest^{TM} DNA polymerase (5 units/µl, Takara Biotechnology (Dalian) Co.), 5 µl 10x Pyrobest^{TM} Buffer II, 5 µl dNTP mixture (2.5 mM each), 1 µl (100 pmol/µl) forward primer, 1 µl (100 pmol/µl) reverse primer, 36.75 µl H_2O. Pyrobest^{TM} is an α-type thermostable DNA polymerase from *Pyrococcus sp.*, which possesses 3'→5'exonuclease (proofreading) activity. PCR thermal cycling was performed using a Tetrad PTC-225 Thermo Cycler (MJ Research, Inc.):

1. 94 °C, 60 seconds to dissociate the template DNA duplex;
2. 94 °C, 45 seconds to dissociate the template DNA duplex;
3. 55 °C, 45 seconds to allow the primers to bind to the template DNA;
4. 72 °C, 45–120 seconds, DNA synthesis (elongation) time depending on the length of target PCR product;
5. Repeat steps 2–4 30 times;
6. 72 °C, 300 seconds, to ensure that the final DNA synthesis step is complete.

Site-directed mutagenesis was performed using the overlap extension PCR (Ho et al., 1989). PCR products were verified by agarose gel electrophoresis and purified using the QIAquick

Gel Extraction Kit (QIAGEN). Exemplary primers are listed in **Table 2.1**.

2.1.3 DNA restriction digest, ligation and transformation

Purified PCR products and appropriate vectors were digested for 1–2 hours with the corresponding restriction enzymes (New England Biolabs) according to the manufacturer's protocols. The digested DNA was verified by agarose gel electrophoresis and purified using the QIAquick Gel Extraction Kit (QIAGEN). Typical 20 µl ligation reaction mixture contained DNA vector and insert in approximately 1:3 ratio. DNA vector and insert were ligated for 10 minutes at 20–25 °C using the Quick Ligation™ Kit (New England Biolabs) according to the manufacturer's protocol. 1–2 µl of the ligation mixture was added to transform 20–50 µl of the Library Efficiency® DH5α™ competent cells (Invitrogen) and both were incubated on ice (0 °C) for 20–30 minutes. Then cells were heat shocked for 45 seconds at 42 °C, and incubated on ice for 2–5 minutes. 500 µl of S.O.C. medium (Invitrogen) was added to cells and incubated for 1 hour with 200 rpm shaking at 37 °C. Finally, 50 µl of cell suspension was spread on lysogeny broth (LB)-agar plates containing 100 µg/ml ampicillin (Melford Laboratories) and incubated for 16–20 hours at 37 °C. A single DH5α colony of each construct was picked up and used to inoculate 5 ml of LB-ampicillin cultures for 16–20 hours at 37 °C in order to purify plasmid DNA. The plasmid DNA was purified using QIAprep Spin Miniprep Kit (QIAGEN). The purified plasmid DNA was digested with appropriate enzymes and analyzed by agarose gel electrophoresis. The plasmid DNA containing an insert of the expected size was sequenced for final validation (Source BioScience DNA Sequencing facility at the Department of Biochemistry, University of Oxford, UK). Forward (Fwd) and reverse (Rev) external sequencing primers for the pHLsec vector are pHLsec-Fwd 5'-GCTGGTTGTTGTGCTGTCTCATC-3' and pHLsec-Rev 5'-CACCAGCCACCACCTTCTGATAG-3'.

Three Glypican 3 (GPC3, UniProtKB/Swiss-Prot number P51654) constructs ($GPC3_{31–493}$, residues Pro31–Phe493; $GPC3_{31–514}$, residues Pro31–Ile514; $GPC3_{31–538}$, residues Pro31–Asn538) in pHL-Avitag3 vector (Aricescu et al., 2006) for surface plasmon resonance (SPR) binding assays were provided by Benjamin Bishop and Dr. Christian Siebold (University of Oxford, UK). The pHL-Avitag3 vector is based on the pHLsec backbone. pHL-Avitag3 introduces a biotin-protein ligase recognition site and a hexahistidine tag at the C-terminus of the encoded protein. Biotin-protein ligase (enzyme commission (E.C.) number 6.3.4.15) BirA (biotin inducible repressor, BirA) was provided by Christian Bell (University of Oxford, UK). Biotin is transferred onto the ε-amino group of a target lysine residue by BirA in an ATP-dependent manner.

2.2 Protein expression and purification

2.2.1 Small scale protein expression and Western blot analysis

WIF-1 and sFRP-1 constructs were expressed and secreted by two human embryonic kidney (HEK) cell lines: HEK-293T and, N-acetylglucosaminyltransferase I-negative HEK-293T variant, HEK-293S. HEK-293S cells are unable to synthesize complex Asn-linked polysaccharides leaving posttranslational modification stalled at the simple $Man_5GlcNAc_2$ glycosylation stage; hence facilitating enzymatic deglycosylation and limiting a source of heterogeneity that may interfere with protein crystallisation (Reeves et al., 2002; Aricescu et al., 2006). Thus proteins produced by HEK-293S (but not HEK-293T) cells were preferred for crystallisation trials, generally after enzymatic deglycosylation. Cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM high glucose, Sigma), supplemented with L-glutamine, non-essential amino-acids (Gibco) and 10% foetal calf serum (FCS, Sigma) in a humidified 37 °C incubator with 5 % CO_2 (Aricescu et al., 2006).

Expression and secretion of all constructs was initially assessed in small scale expression tests in HEK-293T cells. Cells were seeded in 6-well plates (well volume 16.8 ml, Corning), ~0.5–2×10^6 cells and 2 ml media per well, and allowed to recover overnight. Then media was changed to DMEM containing 2 % FCS. 100 µl of FCS-free DMEM was gently mixed with 2–4 µg target construct plasmid DNA plus 10 µl Lipofectamine™ 2000 transfection reagent (Invitrogen), incubated at room temperature (~20 °C) for 10 minutes, and added to the well containing 2 ml DMEM, 2 % FCS. Cells were grown for three days in a humidified 37 °C incubator with 5 % CO_2. Then conditioned media and cells were collected and prepared for Western blot analysis to test expression and secretion levels of the target protein.

Conditioned media was centrifuged (15000g, 10 minutes, ~20 °C) to separate cell debris and 10 µl sample was taken for Western blot analysis. Cells from one well were suspended in 400 µl of lysis buffer (25 mM Tris-phosphate (pH 7.8), 2 mM DTT, 2 mM 1,2-diaminocyclohexane-N,N,N´,N´-tetraacetic acid, 10% glycerol, 1% Triton X-100), and 3 µl of sample was taken for Western blot analysis. Typically, these samples were boiled (5 minutes, 100 °C) in a reducing loading buffer and loaded onto a sodium dodecyl sulfate polyacrylamide gel, and proteins were separated according to their size by electrophoresis (SDS-PAGE) using standard methods (Sambrook and Russell, 2000). Polyhistidine-tagged molecular weight standard, BenchMark™ Protein Ladder (Invitrogen), was included as a sample during the SDS-PAGE. After SDS-PAGE, proteins from the gel were transferred onto a nitrocellulose membrane (0.45 µm pore size, Amersham Hybond™ ECL™ membrane from GE Healthcare) by electrophoresis using XCell SureLock™ Mini-Cell system (Invitrogen) in 25 mM Bicine, 25 mM Bis-Tris, 1 mM EDTA, pH 7.2 (NuPAGE transfer buffer, Invitrogen). The membrane was incubated in 5 % fat-free milk powder in PBS buffer for 10 minutes to reduce nonspecific binding of antibodies in the later steps. The membrane was (1) probed

with a mouse monoclonal anti-PentaHistidine primary antibody (1:1000 dilution, QIAGEN) for 1 hour at ~20 °C, (2) rinsed with PBS, 0.1 % Tween-20 buffer several times, (3) probed with a goat anti-mouse IgG horseradish peroxidase (HRP) conjugate for 1 hour at ~20 °C, (4) washed with PBS, 0.1 % Tween-20 five times (20–30 minutes in total). HRP was detected by chemiluminescence using the Amersham ECL™ Advance Western Blotting detection kit (GE Healthcare) and exposure to the light sensitive film (Amersham Hyperfilm ECL from GE Healthcare).

2.2.2 Large scale expression and purification of WIF-1, sFRP-1 CRD and GPC3

Constructs that were efficiently expressed and secreted by HEK-293T cells in small scale trials were taken forward for large scale expression. Large scale expression of secreted proteins was typically performed by Weixian Lu or Dr. Yuguang Zhao (University Of Oxford, UK) using HEK-293S cells. Cells were grown in expanded surface polystyrene roller bottles (surface area 2125 cm^2, Greiner Bio-One) (Aricescu et al., 2006). Cells were transfected using a mixture of 2 mg plasmid DNA (purified using endotoxin-free Plasmid Mega or Giga Kits from QIAGEN) plus 4 mg 25 kDa branched polyethylenimine (Aldrich) per litre of media. Typically, after 7–14 days, media (3 litres from 12 roller bottles) containing secreted protein was collected, centrifuged (7000g, 1 h, 16 °C), and filtered (polystyrene filter with 0.22 μm pore size, Steritop™ Filter Units, Millipore). Media was dialysed against 30 litres of PBS twice for ~48 hours at 4 °C with stirring (100 rpm). Proteins were further purified by immobilised metal ion affinity chromatography (IMAC). The dialysed media was incubated with 10 ml of Co^{2+}-charged Talon metal affinity beads (Clontech) (1–2 hours, 16 °C, 100 rpm). Beads with the protein bound were collected, washed with 500 ml PBS buffer and 250 ml of 10 mM imidazole in PBS buffer at room temperature (~20 °C). The Co^{2+}-bound protein was eluted with 50 ml of 350 mM imidazole in PBS buffer. Proteins used for crystallisation

were deglycosylated with endo-β-N-acetylglucosaminidase-F1 (Endo F1) (75 μg Endo F1/mg of target construct, 18 hours, 21 °C); otherwise, proteins used for functional or biophysical assays were kept fully glycosylated. The eluted protein-containing solution was concentrated to 1 ml and loaded onto a pre-equilibrated Superdex 75 16/60 or Superdex 200 16/60 (Amersham) column operated using the Äkta high performance liquid chromatography (HPLC) systems (GE Healthcare). Typically, size-exclusion chromatography (SEC) of WIF-1, sFRP-1 or GPC3 constructs was performed at a flow speed of 0.8 ml/min in 20 mM Tris-HCl pH 8.0, 300 mM NaCl buffer at room temperature (~20 °C). Protein fractions were analysed by SDS-PAGE. Fractions containing >95 % pure protein (based on the SDS-PAGE analysis) were pooled and immediately used for subsequent experiments (e.g. crystallisation or cell-based assays), or flash-frozen in liquid nitrogen and stored at -80 °C.

2.2.3 Large scale expression and purification of Wnt-3a

The mouse Wnt-3a purification protocol presented here is based on a previous publication (Willert et al., 2003). Mouse Wnt-3a was expressed by a mouse fibroblast cell line (American Type Culture Collection, catalogue number CRL-2647). The presence of mouse Wnt-3a was monitored by Western blot analysis throughout the purification with the following antibodies: primary rat monoclonal antibodies (rat IgG_{2A}) recognising mouse Wnt-3a regions Ser36–Gln75 and Trp219–Arg269 from R & D Systems; secondary goat antibodies (conjugated to horseradish peroxidase (HRP)) recognising rat IgG (H+L) from Invitrogen. HRP was detected and visualised as described by Aricescu et al., 2006. Cells were maintained in DMEM (Sigma) supplemented with L-Gln, non-essential amino acids (Gibco), 10 % foetal calf serum (FCS, Sigma) and aminoglycoside antibiotic G-418 disulfate (0.4 mg/ml, Melford Laboratories) in a humidified 37 °C incubator with 5 % CO_2. Cells were grown in expanded surface polystyrene roller bottles (surface area 2125 cm^2, Greiner Bio-One) (Aricescu et al.,

2006) for 7 days. Conditioned media (1 litre from 4 roller bottles) containing secreted Wnt-3a was collected, centrifuged (7000g, 1 hour, 16 °C) and filtered (polystyrene filter with 0.22 μm pore size, Steritop™ Filter Units, Millipore). Filtered media was dialysed against 15 litres of PBS twice for ~48 hours at 4 °C with stirring (100 rpm). Then media was warmed up to ~20 °C and supplemented with Triton X-100 (final concentration 1 %). Media was load onto a 5 ml HiTrap™ Blue HP column (pre-packed with Cibacron Blue F3G-A, GE Healthcare) at a flow rate 5 ml/min and the column was washed with 110 ml PBS, 1 % CHAPS. Wnt-3a was eluted with KCl gradient (1.5 M KCl, 20 mM Tris-HCl pH 8.0, 1 % CHAPS; flow rate 2 ml/min). Fractions containing Wnt-3a were pooled (total volume 15 ml), diluted in PBS, 1 % CHAPS (up to 50 ml) and loaded onto a Ni^{2+}-charged 5 ml HisTrap™ HP column (GE Healthcare). The column was washed with 50 ml PBS, 1 % CHAPS, then 25 ml 5 mM imidazole pH 8.0, PBS, 1 % CHAPS. Wnt-3a was eluted with imidazole gradient (from 5 mM to 400 mM imidazole pH 8.0 in PBS, 1 % CHAPS). Fractions containing Wnt-3a were pooled (total volume 15 ml), concentrated to 0.8 ml and loaded onto a pre-equilibrated (PBS, 1 % CHAPS) Superdex 200 10/300 column (GE Healthcare). SEC was performed at a flow rate of 0.4 ml/min. Fractions containing Wnt-3a were flash-frozen in liquid nitrogen and stored at -80 °C.

2.3 Protein crystallisation

For crystallisation trials, appropriate SEC protein peak fractions were concentrated in centrifugal concentrators (Vivascience) at 3000g, 16 °C. Proteins were crystallised using the sitting drop vapour-diffusion method (discussed in a textbook by B. Rupp, 2010, Chapter 3). To determine an appropriate protein concentration for crystallisation, a pre-crystallisation experiment using the PCT™ Pre-Crystallization Test kit (Hampton Research) was performed for each new protein sample. During initial crystallisation trials, sitting drop contained 100 nl

of concentrated protein solution and 100 nl reservoir solution (Walter et al., 2005). Crystallisation was performed in 96-well crystallisation plates (Greiner) and liquid nanolitre-scale pipetting was performed using robots (Cartesian Technologies). Crystallisation plates were placed into a temperature-controlled (21 °C) storage vault (TAP "Homebase") and sitting drops were regularly imaged using the Veeco visualisation system (Mayo et al., 2005). For the initial crystallisation trials, up to ~800 different reservoir solutions were tested from the commercially available crystallisation screening kits from three companies: Hampton Research, Emerald Biosystems, and Molecular Dimensions (Walter et al., 2005). Crystallisation conditions for WIF-1 and sFRP-1 constructs are provided in **Table 2.2** and corresponding chapters.

Table 2.2. Crystallisation conditions of the WIF-1 and sFRP-1 constructs.

Construct	**Space group**	**Protein conc., mg/ml**	**Crystallisation solution***	**Crystal cryo-protectant**
WIF-$1_{\Delta C}$	$I2_13$	6.5	0.75 M sodium succinate pH 5.7, 35 mg/ml sucrose octasulfate sodium salt (Toronto Research Chemicals Inc.)	25 % glycerol
WIF-$1_{WD\text{-}EGF\text{-}I}$	C222	7.8	1.3 M diammonium tartrate, 100 mM bis-tris propane pH 7.0	30 % glycerol
WIF-$1_{WD\text{-}EGF\text{-}I}$ Met77Trp	C222	10.3	1.1 M diammonium tartrate pH 7.0, 20 mM spermidine, 0.5 mM $CaCl_2$	30 % glycerol
WIF-1_{WD}	C222	9.2	30 % w/v PEG 8000, 200 mM sodium acetate, 100 mM sodium cacodylate pH 6.5	30 % glycerol
WIF-$1_{EGFs\ I–V}$	C222 (?)	10	30 % v/v polyethylene glycol monomethyl ether 2000, 0.2 M $(NH_4)_2SO_4$, 0.1 M sodium acetate pH 4.6, 0.5 mM $CaCl_2$	None
sFRP-$1_{32–176}$ CRD	$P4_32_12$	4.1	2.0 M $(NH_4)_2SO_4$, 0.1 M cacodylate pH 6.5, 0.2 M NaCl	30 % glycerol
sFRP-$1_{52–176}$ CRD Crystal I	$P6_322$	5.6	2.0 M $(NH_4)_2SO_4$, 0.1 M cacodylate pH 6.5, 0.2 M NaCl, 20 % v/v 1-propanol	30 % glycerol
sFRP-$1_{52–176}$ CRD Crystal II	P3	5.2	2 M $(NH_4)_2SO_4$, 4 % v/v 1-propanol, 5 mg/ml sucrose octasulfate sodium salt	Perfluoropolyether PFO-X125/03 (Alfa Aesar)
sFRP-$1_{52–176}$ CRD Crystal III	$P6_1$	5.2	10 mM $MgSO_4$, 1.8 M Li_2SO_4, 50 mM sodium cacodylate pH 6.0, 5 mg/ml sucrose octasulfate sodium salt	25 % glycerol
* Sitting-drop contained 100 nl of the listed crystallisation solution plus 100 nl protein, typically, in 300 mM NaCl, 20 mM Tris-HCl pH 8.0 buffer solution.				

2.3.1 Crystal optimisation

Multiple or poorly X-ray diffracting protein crystals obtained during initial crystallisation screening were further optimised using three procedures. The first procedure uses volatile and non-volatile compounds from Hampton Additive Screens 1–3 (Hampton Research) as

additives to the original crystallisation condition (Walter et al., 2005). The additional non-volatile chemicals are dispensed as 100 nl aliquots directly into the crystallisation sitting drop (containing 100 nl protein and 100 nl reservoir solution) (Walter et al., 2005). The volatile compounds are dispensed in 20 µl aliquots into the reservoir solution (after the 100 nl transfer of reservoir solution to the sitting drop) (Walter et al., 2005). For example, in the case of the WIF-$1_{WD\text{-}EGF\text{-}I}$ Met77Trp construct, the additive compound spermidine helped to avoid formation of multiple crystals and the corresponding electron density of spermidine is visible in the final electron density map. The second optimisation strategy was to vary the concentration of the compounds and pH for the reservoir solutions identified during initial crystallisation screening. The third strategy was to crystallise proteins with small molecule mimics of their natural ligands. For example, WIF-$1_{\Delta C}$, a heparan sulfate-binding protein, was crystallised in complex with sucrose octasulfate (SOS), a heparan sulfate analogue.

2.3.2 Reductive lysine methylation

The reductive lysine methylation procedure was used to optimize crystals of the WIF-1_{WD}, WIF-$1_{WD\text{-}EGF\text{-}I}$, and WIF-$1_{WD\text{-}EGF\text{-}I}$ Met77Trp proteins (Walter et al., 2006). This procedure was also used to evaluate the contribution of the ε-amino group of WIF-1 lysines to Wnt signalling inhibition in cellular assays (**Section 2.11**). The reductive lysine methylation reaction yields a dimethylated ε-amino group for lysine side chains. The additional methyl groups change the surface properties of the protein and may form hydrophobic interactions with the neighbouring molecules in the crystal. Lysine methylation was performed on proteins during the late steps of purification, typically, after IMAC or SEC, in PBS buffer. 20 mM borane dimethylamine complex ($(CH_3)_2NH \cdot BH_3$, a reducing agent), 40 mM formaldehyde and 0.1–1 mg/ml protein were mixed and incubated at 4 °C for 10 minutes. Then additional $(CH_3)_2NH \cdot BH_3$ and formaldehyde were added to the reaction to double their

concentration. Reagents were mixed and incubated at 4 °C for 10 minutes, $(CH_3)_2NH \cdot BH_3$ and formaldehyde were added again and the reaction mixture was incubated at 4 °C for 14–20 hours. Then 20 mM Tris-HCl pH 8.0 was added to react with the excess of formaldehyde and $(CH_3)_2NH \cdot BH_3$. The reaction mixture was centrifuged (5000g, 10 minutes, 16 °C) to remove precipitated protein or reagents, and concentrated before further purification by SEC in a Tris-based buffer (typically, 300 mM NaCl, 20 mM Tris-HCl pH 8.0).

2.4 Crystal cryo-protection and X-ray diffraction data collection

Crystal cryo-protection and subsequent X-ray diffraction data collection at low temperature (typically, -173 °C) are necessary for obtaining a complete data set by reducing the problem of radiation damage (reviewed by Garman, 2010). Cryo-protectants such as glycerol were used to protect crystals from the damaging effects of ice formation during the following cryo-cooling procedure: (1) crystals were transferred from the drops in the crystallisation plates to the drops containing reservoir solution and cryo-protectant (**Table 2.2**), (2) kept in the cryo-protectant-reservoir solution mixture for 5–60 seconds, and (3) flash frozen in liquid nitrogen (-196 °C).

Crystal X-ray diffraction data were collected using beamlines at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) and Diamond Light Source (Oxfordshire, UK) synchrotrons. The optimum orientation of the crystal for the complete data set collection was determined using a program iMosflm (Powell, 1999; Leslie, 2006). Typically, diffraction data corresponding to a minimum of 180° crystal rotation (with oscillation angle 1°) were collected using charge-coupled device (CCD)-based detectors (reviewed by Phillips et al., 2000). The crystal was kept in a nitrogen gas cryostream at 100 K (-173 °C) during the data collection.

2.5 X-ray diffraction data processing

X-ray diffraction data processing involved several steps: (1) determination of unit cell dimensions, (2) assignment of Miller indices (hkl) to each reflection, (3) integration of reflection intensities, (4) scaling and (5) merging reflection intensities. All diffraction data processing was performed with the HKL2000 program suite (Otwinowski and Minor, 1997), except WIF-$1_{WD\text{-}EGF\text{-}I}$ Met77Trp data (collected using a single-photon-counting PILATUS detector (Kraft et al., 2009)) that were processed and scaled with XDS (Kabsch, 2010) and Scala (Evans, 2006). Diffraction data statistics are given in the relevant results Chapters.

Structure factor amplitudes were calculated from the reflection intensities using a program TRUNCATE (French and Wilson, 1978); the intensity of a reflection is proportional to the square of the amplitude of the structure factor (Darwin's formula; Darwin, 1914; re-derivation can be found in Holton and Frankel, 2010) (Formula 1 in **Appendix 6**).

2.6 Structure determination, refinement and validation

2.6.1 Structure determination and refinement

Structure factor phases for each structure factor were determined using the molecular replacement method (Rossmann and Blow, 1962) and a program Phaser (McCoy et al., 2007). The crystal structure of WIF-1_{WD} was solved by molecular replacement using as a search model the WIF domain coordinates determined by NMR spectroscopy (Liepinsh et al., 2006). The initial electron density maps of the WIF domain (immediately after molecular replacement) revealed the architecture of a hydrophobic pocket which appeared to be occupied by lipid-like ligand (Formulae 2 and 3, **Appendix 6**). A selection of different phosphatidyl cholines were identified as the WIF-1 lipid ligands by mass spectrometry (**Section 2.8**). 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC) was built into the

electron density and included in the model for refinement. The crystal structures of the wild-type WIF-$1_{WD\text{-}EGF\text{-}I}$ and WIF-$1_{WD\text{-}EGF\text{-}I}$ Met77Trp were phased by molecular replacement in Phaser using the WIF-1_{WD} crystal structure (with DPPC omitted) and DPPC was again built into additional electron density. Structures were refined using iterative cycles in programs Coot (Emsley et al., 2010), autoBUSTER (Blanc et al., 2004), and phenix.refine (Adams et al., 2010). The composite omit electron density map (Terwilliger et al., 2008) was calculated using Phenix (Adams et al., 2010).

The wild-type WIF-$1_{WD\text{-}EGF\text{-}I}$ structure also reveals a potential metal-binding site. Based on coordination geometry, a sodium ion was built, which is coordinated by three water molecules (distance 2.5–2.8 Å) and the two protein backbone oxygens of Cys186 (distance 2.3 Å) and Gly190 (distance 2.5 Å) (Hsin et al., 2008).

Initial phases for the WIF-$1_{\Delta C}$ X-ray diffraction data were determined by molecular replacement using the high resolution crystal structure of WIF-1 WD as the search model in Phaser. Phases were improved by solvent flattening and histogram matching using the program Parrot (Zhang et al., 1997). The resulting electron density map was sufficiently clear to place models of the first three EGF-like domains. Structure refinement iterated cycles of restrained refinement in autoBUSTER, phenix.refine and manual rebuilding with Coot. No interpretable density was found for EGFs IV–V. These domains are accommodated within solvent channels and appear to retain significant flexibility within the crystal.

The sFRP-1 cysteine-rich domain diffraction data were processed with iMosflm (Powell, 1999; Leslie, 2006) and Scala (Evans, 2006), structure factor phases were obtained by molecular replacement using program Phaser and a homology model of the sFRP-1 CRD (based on the mouse sFRP-3 CRD crystal structure (Dann et al., 2001; PDB ID 1IJX). Non-crystallographic symmetry (NCS) restraints (discussed by Kleywegt, 1996) were imposed

during refinement using program phenix.refine.

2.6.2 Structure validation

Several parameters (R_{work} and R_{free}; Formulae 4 and 5, **Appendix 6**) were monitored throughout the refinement to ensure consistency between observed diffraction data and atomic coordinates of the protein model (Brünger, 1992). Throughout the refinement, stereochemical properties, bond lengths and angles in the protein model were monitored and compared to reference values obtained from high resolution structures (small molecules and proteins) using a program MolProbity (Davis et al., 2007). The composite omit electron density maps (Terwilliger et al., 2008) were calculated with Phenix (Adams et al., 2010) to validate specific features in atomic models, for example, presence of the bound DPPC. Final refinement and validation statistics for all crystal structures are presented in the corresponding chapters.

2.7 Surface plasmon resonance (SPR) binding assays

2.7.1 SPR binding studies of WIF-1 and Wnt-3a

Surface plasmon resonance (SPR) experiments were performed using a Biacore T100 machine (GE Healthcare) at 25 °C, typically, in PBS, 1 % CHAPS (w/v) buffer. All WIF-1 constructs underwent gel filtration in SPR running buffer immediately before use. WIF-1 construct concentrations were determined from the absorbance at 280 nm using calculated molar extinction coefficients from a program ProtParam (Gasteiger et al., 2005). Purified Wnt-3a and streptavidin (as a control protein, Thermo Fisher Scientific) were immobilised on a CM5 BIAcore sensor chip (Biacore Life Sciences) via primary amines using the following protocol:

1. Activate the sensor chip surface with a freshly prepared mixture of 0.4 M EDC (N-(3-

dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride from Aldrich or Biacore) plus 0.1 M NHS (N-hydroxysuccinimide from Fluka or Biacore) in water;

2. Immobilise streptavidin on the reference flow cell of the sensor (flow rate 10 µl/min, 5–10 minutes, immobilises ~5000–8000 response units (RUs) of streptavidin);
3. Dilute purified (~70 %) mouse Wnt-3a (0.01–0.05 mg/ml in PBS, 1 % CHAPS) 1:10 in 10 mM sodium acetate, 1 % CHAPS, pH 5.5 and inject into the experimental flow cells of the sensor (flow rate 10 µl/min, 5–10 minutes, immobilises >1000 response units (RUs) of the protein);
4. Deactivate sensor chip surface with 1 M 2-aminoethanol, 1 % CHAPS, pH 8.5 (flow rate 10 µl/min for 5–10 min or 100 µl/min for 2 minutes);
5. Wash all lanes with SPR running buffer (typically, PBS, 1 % CHAPS) until stable baseline is reached ((ΔRUs/Δtime) = 0), typically, 30–60 minutes, flow rate 5–10 µl/min.

For measurements of the WIF-1 construct binding to immobilised proteins, samples were injected at 10 µl/min. The signal from experimental flow cells was corrected by subtraction of a buffer and reference signal from a control protein coupled flow cell. After each injection and WIF-1–Wnt-3a binding, the sensor surface was regenerated using 1.5 M KCl, 20 mM Tris-HCl pH 8.0, 1 % CHAPS (100 µl/min, 1 min). WIF-$1_{\text{Full length}}$–Wnt-3a binding experiments in the presence of 2 mM $CaCl_2$ pH 8.0 or 2 mM EDTA pH 8.0 were performed in 150 mM NaCl, 20 mM Tris-HCl pH 8.0, 1 % CHAPS running buffer (buffer A); WIF-$1_{\text{WD-EGF-I}}$–Wnt-3a binding experiments in the presence of 2 mM phosphocholine chloride calcium salt tetrahydrate (Sigma) were performed in buffer A or in 150 mM NaCl, 20 mM HEPES pH 7.5, 1 % CHAPS running (buffer B). Note that Wnt3a must be always supplemented with 1 % CHAPS during all stages of the experiment (e.g. absence of CHAPS during steps 3 and 4

reduces signal/noise ratio in the subsequent binding measurements). In addition, I would recommend using HEPES-based (but not PBS- or Tris-based) buffers for WIF-1–Wnt-3a SPR binding studies. All experiments (including WIF-1 binding to glycosaminoglycans (GAGs), **Section 2.7.3**) were performed in triplicate (only the WIF-$1_{WD\text{-}EGF\text{-}I}$–Wnt-3a binding experiment in the presence of phosphocholine chloride in buffer B was performed in duplicate). K_d values were calculated using BIAevaluation 3.0 program (Biacore Life Sciences) and a 1:1 Langmuir binding isotherm model.

2.7.2 SPR binding studies of sFRP-1 CRD and Wnt-3a

For measurements of the sFRP-1 CRD binding to immobilised Wnt-3a, similar experimental settings as for WIF-1–Wnt-3a were used. 150 mM NaCl, 20 mM HEPES pH 7.5, 1 % CHAPS was used as a running buffer.

2.7.3 SPR binding studies of WIF-1 or sFRP-1 CRD and glycosaminoglycans

Heparin polymer (MW >9000 g/mol, Iduron), heparan sulfate from pig mucosa (Iduron), chondroitin sulfate sodium salt from shark cartilage (Sigma), and chondroitin 4-sulfate sodium salt (Fluka) were biotinylated with EZ-Link Biotin-LC-Hydrazide (Thermo Fisher Scientific) in aqueous 17 % DMSO solution (36 h, 20 °C), then extensively dialysed against water and SPR running buffer (120 mM NaCl, 10 mM HEPES pH 7.5, 0.05 % Tween 20). Biotinylated glycosaminoglycans (GAGs) were immobilised on the streptavidin-coated SPR sensor surface. For measurements of WIF-1 or sFRP-1 CRD construct binding to immobilised GAGs, the same experimental settings as for WIF-1–Wnt-3a were used. The sensor surface was regenerated using 1.5 M NaCl, 10 mM HEPES pH 7.5, 0.05 % Tween 20.

2.7.4 Preparation of GPC3 constructs

Three human GPC3 constructs were expressed in HEK-293T cells (to preserve native

glycosaminoglycan chains on the protein) and purified as described in **Section 2.2.2**. The biotinylation mixture contained:

1. 0.5 ml (~10–100 μg) purified GPC3 construct in 300 mM NaCl, 20 mM Tris-HCl pH 8.0;
2. 0.1 ml 500 mM bicine buffer pH 8.3;
3. 0.1 ml 100 mM ATP, 100 mM magnesium acetate, 500 μM D-biotin;
4. 50 μl (~2 mg/ml) biotin-protein ligase in 50 mM imidazole pH 6.8, 50 mM NaCl, 5 % glycerol, 6 mM 2-hydroxy-1-ethanethiol (β-mercaptoethanol).

The reaction mixture was incubated at 30 °C for 30 minutes, then extensively dialysed against 120 mM NaCl, 10 mM HEPES pH 7.5, 0.05 % Tween 20 with an Amicon Ultra centrifugal filter units (10 kDa molecular weight cut-off, Millipore), flash frozen in liquid nitrogen and stored at -80 °C before use.

2.7.5 SPR binding studies of sFRP-1 CRD and GPC3

For measurements of the sFRP-1 CRD binding to immobilised GPC3, the same experimental settings as for sFRP-1 CRD–GAG were used. 120 mM NaCl, 10 mM HEPES pH 7.5, 0.05 % Tween 20 was used as a running buffer.

2.8 Mass spectrometry analysis of WIF-1-bound lipids

WIF-1-bound lipids were extracted into $CHCl_3$/CH_3OH (2:1 ratio, v/v) at 65 °C, 30 min. The $CHCl_3$/CH_3OH-lipid fraction was washed with water three times and concentrated by evaporation at room temperature. The sample was analysed by matrix assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS) on a Bruker Daltonics Ultraflex TOF/TOF mass spectrometer with help from Drs. B. M. Kessler and K.

di Gleria (Universty of Oxford, UK). The sample solution was premixed with the matrix (20 mg/ml 2,5-dihydroxybenzoic acid in 0.1 % (v/v) trifluoroacetic acid, 30 % (v/v) acetonitrile) at a 1:1 ratio and 1 µl of mixture applied directly to the sample plate. The droplet was air-dried before analysis in the mass spectrometer. All positive mode MALDI-TOF-MS spectra were obtained in reflection mode. The ionization source was a nitrogen laser, emitting 337 nm electromagnetic radiation in a 3-ns pulse. The accelerating voltage in the ion source was 30 kV. Peaks of phosphatidylcholines in the MALDI-TOF mass spectrum were assigned according to Schiller et al., 1999.

2.9 WIF-$1_{\Delta C}$–lipid interactions

Purified WIF-$1_{\Delta C}$ (270 µl, final conc. 50 µM) was incubated with 16:0 Liss Rhod PE (1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine-N-(lissamine rhodamine B sulfonyl) ammonium salt, 10 µl of 1 mg/ml Liss Rhod PE in $CHCl_3$, Avanti Polar Lipids) or sulforhodamine B (final conc. 160 µM, Sigma-Aldrich) in PBS, 1 % CHAPS for 3 h at 37 °C. To remove the excess of 16:0 Liss Rhod PE, the sample was loaded onto a pre-equilibrated Superdex 200 10/300 column (GE Healthcare Life Sciences) after incubation and gel filtration was performed in PBS (Sigma), 1 % CHAPS at a flow rate 0.4 ml/min; WIF-$1_{\Delta C}$ peak fractions were pooled and gel filtration was performed again in PBS without CHAPS. Protein and fluorescent molecule elution was followed by absorption at 280, 557 and 571 nm simultaneously. The fluorescence intensity (excitation at 535 nm, emission at 595 nm) was measured on the Infinite F200 (Tecan) microplate reader. Experiments with membrane lipid strips (Echelon Biosciences) were performed according to manufacturer's instructions. Proteins on the strips were visualised by Western blotting and chemiluminescence using the Amersham ECL detection kit (GE Healthcare).

2.10 Analytical heparin affinity chromatography

Three purified WIF-1 constructs (WIF-$1_{WD\text{-}EGF\text{-}I}$, WIF-$1_{EGFs\ I–V}$, WIF-$1_{\Delta C}$) were dialyzed against 10 mM Tris-HCl pH 8.0, 10 mM NaCl and concentrated to a volume of 2 ml (~0.5 mg/ml). Each protein was loaded onto the pre-equilibrated HiTrap Heparin HP 1 ml column (Amersham) separately and eluted with a NaCl gradient (10 mM Tris-HCl pH 8.0, 10 mM NaCl to 10 mM Tris-HCl pH 8.0, 2 M NaCl). Protein elution was followed by absorption at 280 nm. The ligand in the HiTrap Heparin HP 1 ml column is a sulfated glycosaminoglycan which is extracted from the native proteoglycan of porcine intestinal mucosa. The molecular weight of the polymer is distributed over the range 5 000–30 000 Da (Amersham).

2.11 Cellular assays for Wnt signalling

HEK-293T cells were maintained in DMEM (Sigma) supplemented with L-Gln, non-essential amino acids (Gibco) and 10 % foetal calf serum (FCS, Sigma) in a humidified 37 °C incubator with 5 % CO_2. Cells were seeded in 6-well plates (Corning), ~2×10^6 cells/well and allowed to recover overnight. The cells were then transfected with Super-TOPFlash plasmid (10 µg/well; the plasmid contains 11 TCF/LEF-binding sites upstream of the firefly luciferase gene; DasGupta et al., 2005; **Figures 2.1** and **2.2**) using lipofectamine 2000 (10 µg/well, Invitrogen) in media containing 2 % FCS, incubated for 5 h. Then the media was changed to media containing 10 % FCS. Luciferase expression was induced by adding purified human Wnt-3a (350 ng/well, equal to 4.7 nM, R&D Systems) into the media. Cells were incubated for 72 h either in the presence or absence of purified WIF-1 constructs or bovine serum albumin (BSA, New England Biolabs) as a control. The cells were washed with PBS buffer, lysed and luciferase activity was measured using the Luciferase Assay System (Promega). The luciferase activity from Super-TOPFlash plasmid transfected cells

treated with Wnt-3a only was taken as 100 %. The effect of BSA on Wnt-3a signalling was negligible. Experiments were performed in triplicate.

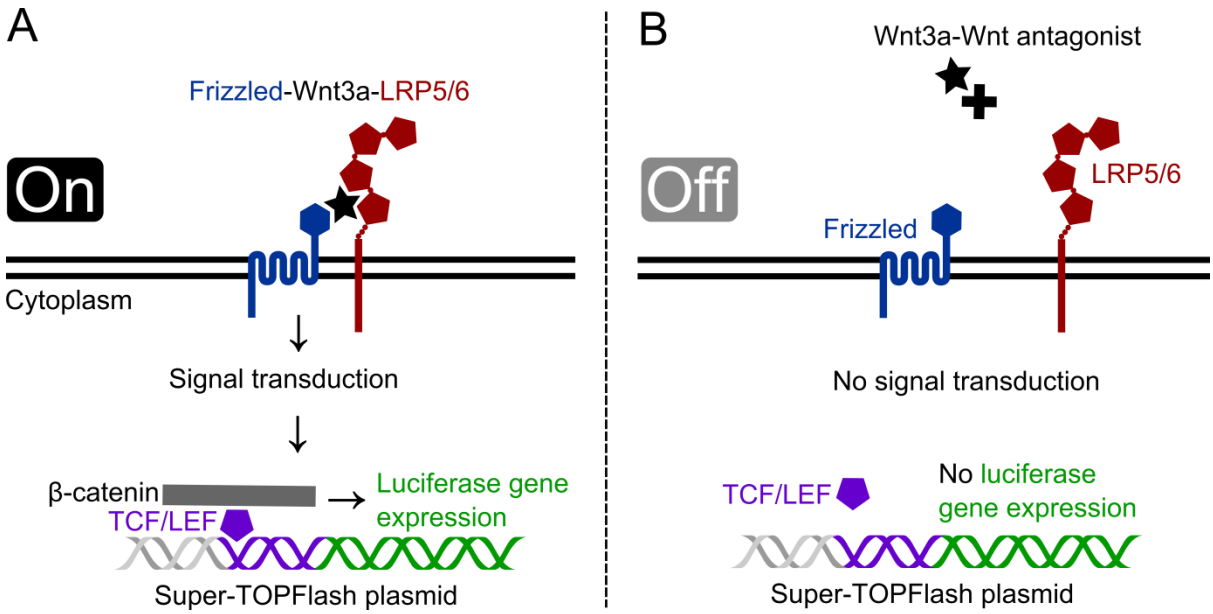


Figure 2.1. A cellular assay for measuring Wnt signalling and its inhibition. The canonical Wnt/β-catenin signalling pathway leads to the activation of TCF/LEF transcription factors, transcription and expression of target genes (**Figure 1.2**). **A**, in order to monitor Wnt signalling, HEK-293T cells were transfected with a plasmid encoding luciferase gene (shown as a green helix region) upstream of a minimal TA promoter and 11 TCF/LEF transcription factor binding sites (violet helix region) (**Figure 2.2**). After transfection, cells were treated with a human Wnt-3a protein (shown as a black star) leading to the activation of TCF/LEF transcription factors and thus expression of luciferase (Wnt signalling switched on). After three days, cells were harvested and the expression level of luciferase was quantified by measuring its activity. **B**, the inhibitory effect of different Wnt antagonists (typically, various WIF-1 constructs; shown as a black cross) was monitored by treating transfected cells with both Wnt-3a and antagonist variants. Wnt antagonist WIF-1 binds to Wnt-3a and thus inhibits signal transduction leading to the lower level of luciferase expression (Wnt signalling switched off).

5'-

TTCTCTATCGATAGGTACCGAGCTCTTACGCGAGATCAAAGGGGGTAAGATCAAAGGGGGTAAGATCAAAGGGGCGCGAGATCAA

AGGGGGTAAGATCAAAGGGGGTAAGATCAAAGGGGGTAAGATCAAAGGGGCGCGAGATCAAAGGGGGTAAGATCAAAGGGGGT

AAGATCAAAGGGGGTAAGATCAAAGGACGCGTCGCGCCCGCGTGCTAGCCCGGGCTCGAGATCTAGACTCTAGAGGGTATATAAT

GGAAGCTCGAATTCCAGCTTGGCATTCCGGTACTGTTGGTAAAAAGCTTGGCATTCCGGTACTGTTGGTAAAGCCACCATGGAAGA

CGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTATCCGCTGGAAGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAG

AGATACGCCCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGGACATCACTTACGCTGAGTACTTCGAAATGTCC

GTTCGGTTGGCAGAAGCTATGAAACGATATGGGCTGAATACAAATCACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTT

ATGCCGGTGTTGGGCGCGTTATTTATCGGAGTTGCAGTTGCGCCCGCGAACGACATTTATAATGAACGTGAATTGCTCAACAGTATG

GGCATTTCGCAGCCTACCGTGGTGTTCGTTTCCAAAAAGGGGTTGCAAAAAATTTTGAACGTGCAAAAAAAGCTCCCAATCATCCA

AAAAATTATTATCATGGATTCTAAAACGGATTACCAGGG-3'

Figure 2.2. DNA sequence region of Super-TOPFlash plasmid containing KpnI restriction site (grey), 11 TCF/LEF-binding sites (blue), MluI restriction site (red), TA minimal promoter from the herpes simplex virus thymidine kinase promoter (yellow, Clontech) and starting region of the firefly luciferase gene (luciferin 4-monooxygenase from *Photinus pyralis* (North American firefly), E.C. 1.13.12.7, green). The plasmid was a kind gift from Dr. Randall Moon (University of Washington, Seattle, USA) (DasGupta et al., 2005; also see **Figure 2.1**). Super-TOPFlash plasmid backbone is a pTA-Luc vector (Clontech). DNA sequencing primer used to obtain the above given sequence: 5'-CCTTATGCAGTTGCTCTCC-3'.

2.12 Binding studies between HEK-293T or CHO cells and WIF-1 constructs

HEK-293T, CHO-K1, and CHO-PgsA-745 cell lines were used to assess interactions between cell surface which is likely (HEK-293T, CHO-K1) or not (CHO-PgsA-745) to contain WIF-1-binding heparan sulfate proteoglycans. CHO-PgsA-745 is a Chinese hamster ovary cell mutant deficient in xylosyltransferase (UDP-D-xylose:serine-1,3-D-xylosyltransferase) (Esko et al., 1985). The CHO-PgsA-745 cell line (ATCC CRL-2242) was derived from CHO-K1 cells (ATCC CCL-61) treated with mutagen (ethylmethanesulfonate) and screened for mutants defective in proteoglycan synthesis (Esko et al., 1985). These cells have a defect in xylosyltransferase, the first sugar transfer enzyme in glycosaminoglycan synthesis, and do not produce glycosaminoglycans.

Cells were seeded at ~1–10×10^6 cells/well in 6-well plates and allowed to recover overnight in DMEM supplemented with L-Gln, non-essential amino-acids and 10 % FCS in a humidified 37 °C incubator with 5 % CO_2. Then the media was changed and supplemented with different, purified WIF-1 constructs (0.1 mg/well; 2 ml media/well). Cells were incubated for 6 hours and then washed three times with PBS buffer. The presence of the His_6-tagged WIF-1 constructs in the cell lysates was checked by Western blot analysis with anti-His antibodies as described previously (Aricescu et al., 2006).

2.13 Multi-angle laser light scattering (MALLS)

sFRP-1 CRD was purified as described in **Section 2.2.2**, and concentrated to 2.3 mg/ml in MALLS running buffer (150 mM NaCl, 10 mM Tris-HCl pH 8.0). 100 µl of the concentrated sample were loaded onto a 30 ml Superdex 75 HR 10/30 column (GE Healthcare) attached to an Agilent chromatography system. Size exclusion chromatography (SEC) was performed at

a flow speed of 0.5 ml/min at room temperature (~20 °C). Protein elution was monitored with multi-angle light scattering (DAWN® HELEOS™, Wyatt Technology), differential refractive index (Optilab-rEX, Wyatt Technology) and Agilent 1200 UV (Agilent Technologies) detectors. The DAWN® HELEOS™ II uses a 120mW solid-state laser operating at 658 nm as a light source. Data were analysed using the ASTRA V program suite (Wyatt Technology) with help from B. J. C. Janssen (University of Oxford, UK).

2.14 Analytical ultracentrifugation

Sedimentation velocity experiments were performed using an Optima Xl-I analytical ultracentrifuge (Beckman) with help from Christian Bell (University of Oxford, UK). Purified sFRP-1 CRD samples at different concentrations in 120 mM NaCl, 10 mM HEPES pH 7.5 and 0.05 % Tween 20 were centrifuged in double sector 12 mm centrepieces in a An-60 Ti rotor (Beckman) at 50000 rpm and 20 °C. Protein sedimentation was monitored by Rayleigh interference. Data were analysed using SEDFIT (Schuck, 2000).

Chapter 3 − Crystal structures of WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$

3.1 Summary

Wnt inhibitory factor 1 (WIF-1) is a six domain Wnt antagonist composed of the WIF domain (WD) and five epidermal growth factor (EGF)-like domains. The WD serves as the Wnt-binding module both in WIF-1 and Wnt receptor RYK across different species. The atomic details of these interactions are poorly defined and understood. Previous NMR spectroscopy studies on the refolded WD revealed a tight association between a detergent molecule used during the refolding procedure and the WD surface. This association raised two questions. First, is WD a lipid-binding domain and, second, does it bind lipid-modified Wnt proteins via their hydrophobic lipid modifications.

Crystal structures of the natively folded, glycosylated WD, with and without EGF-like domain I (constructs WIF-$1_{WD\text{-}EGF\text{-}I}$ and WIF-1_{WD}, respectively), are presented in **Chapter 3**. They reveal a novel 1,2-diacyl-phosphatidylcholine binding site in the WD. Two acyl chains of the lipid extend deep into the WD while the phosphocholine head group is surface exposed. We show that a tightly WD-bound lipid can be exchanged with its fluorescent lipid analogue. As revealed by the WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structure, the observed interdomain orientation of the WD to EGF-like domain I lacks stabilising non-covalent interactions, suggesting that it is possible for WIF-1 to adopt multiple conformations

3.2 Introduction

As discussed in **Section 1.5.1.2.1**, human Wnt inhibitory factor 1 (WIF-1) is a six domain Wnt antagonist composed of the WIF domain (WD) and five EGF-like domains. WIF-1 inhibits Wnt signal transduction by sequestering Wnt proteins from their receptors. WD is a Wnt-binding domain and is essential for the inhibitory effect of WIF-1 (Hsieh et al., 1999).

Moreover, WDs serve as the extracellular modules of mammalian single-pass transmembrane Ryk (*r*elated to t*y*rosine *k*inase) Wnt receptors and their homologues in *Drosophila melanogaster*. Mammalian RYK functions as a Wnt coreceptor together with Frizzled and activates the canonical Wnt/β-catenin signalling pathway (discussed in **Section 1.5.1.2.2**). The three dimensional structure of the human WIF-1 WD has previously been determined by NMR spectroscopy. NMR structural studies were performed on a refolded WD (henceforth referred to as $WD_{refolded}$) expressed in *Escherichia coli* (Liepinsh et al., 2006). Alkyl chains of the detergent used during refolding bound tightly to the exterior of the $WD_{refolded}$. The observed detergent binding raised the question whether WIF-1 may bind to lipid-modified Wnt proteins using the WD as a lipid recognition module (Malinauskas, 2008; Bazan and de Sauvage, 2009). *In silico* calculations suggested that the $WD_{refolded}$ binds with highest affinity to fatty acids with alkyl chain length of 16 to 18 carbon (Malinauskas, 2008).

3.3 Aims

The main aim of the work presented in **Chapter 3** is to determine crystal structures of the two WIF-1 constructs, WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$, in order to answer the following questions raised by the previous studies:

1. Is WD a lipid-binding domain?
2. Has expression in bacteria *Escherichia coli* and refolding affected the structure of the human WD ($WD_{refolded}$)?
3. How are the WD and the EGF-like domain I orientated relative to each other? What are implications of this conformation for Wnt protein recognition?
4. Is the WD surface structurally similar to that of the well characterized other Wnt-binding domain, the Frizzled cysteine-rich domain (CRD)?

3.4 Expression and purification of WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$

Constructs containing the WD (WIF-1_{WD}) or WD plus EGF-like domain I (WIF-$1_{WD\text{-}EGF\text{-}I}$) encompass amino acid residues Ser35–Gln178 and Ser35–Glu210 of human WIF-1, respectively (**Appendix 2**). The construct design strategy, including expression vector, PCR primers and restriction enzymes, is described in **Section 2.1**. WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ constructs encode the full length protein domains as suggested by:

1. the predicted domain boundaries listed on the UniProtKB/Swiss-Prot Internet protein database (UniProt Consortium, 2010);
2. the NMR structure of the $WD_{refolded}$ ($WD_{refolded}$ boundaries are defined by the first β strand which starts at Tyr37 and the last β strand which ends at Cys177; Figure 1 by Liepinsh et al., 2006);
3. the canonical EGF-like domain fold (reviewed by Bork et al., 1996) stabilised by three disulfide bonds (EGF-like domain I boundaries are defined by the first and the last cysteine residues, Cys182 and Cys209, respectively).

There is one predicted Asn-linked glycosylation site at Asn88 within the WD (the second predicted glycosylation site is located at Asn245, just before the first cysteine of the EGF-like domain III) (**Appendix 4**, **A**). Thus large scale expression was performed in HEK-293S cells because these cells, in contrast to HEK-293T cells, produce homogeneously glycosylated secreted proteins more suitable for subsequent enzymatic deglycosylation and crystallisation (Chang et al., 2007). There are no predicted mucin type GalNAc O-linked glycosylation sites within WIF-1 (**Appendix 4**, **B**) and no predicted disordered regions within the WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ constructs (**Appendix 4**, **C**) that could hinder crystallisation. Small scale expression tests in HEK-293T cells and subsequent Western blot analysis showed expression and secretion of these two constructs (**Figure 3.1**). Secretion of the constructs indicates that proteins are folded and are likely to possess the full biological activity (reviewed by

Trombetta and Parodi, 2003). Proteins from the HEK-293S conditioned media were purified using the standard three stage procedure (as described in **Section 2.2.2),** which includes dialysis, IMAC and SEC. Representative protein elution profiles from SEC and corresponding SDS-PAGE gels from the WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ purification are shown in **Figure 3.2**.

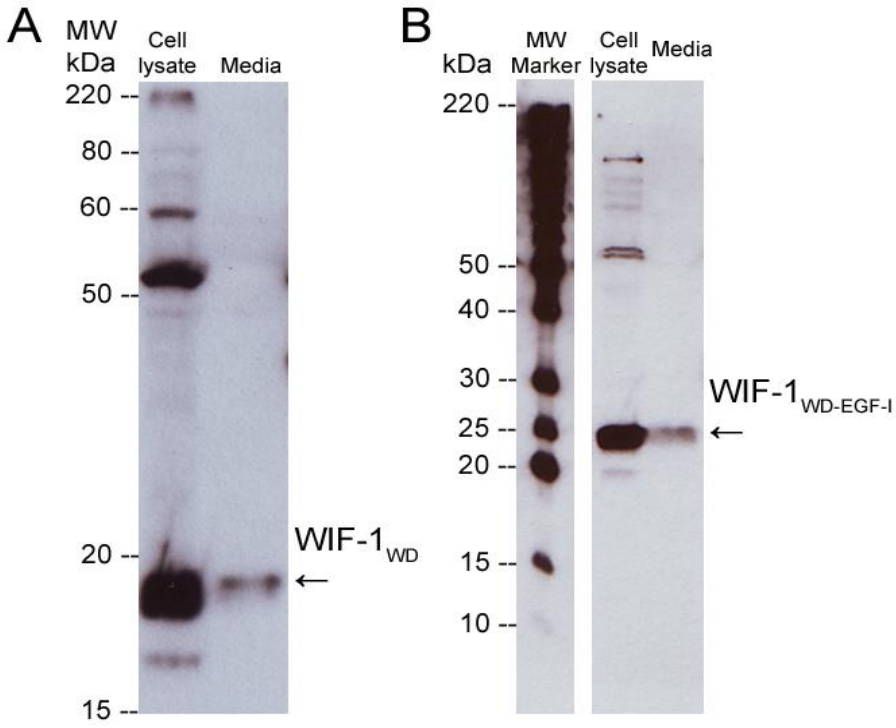


Figure 3.1. Western blot analysis of the WIF-1_{WD} (**A**) and WIF-$1_{WD\text{-}EGF\text{-}I}$ (**B**) expression and secretion by HEK-293T cells. Both His-tagged constructs (indicated with arrows) are present in cell lysates and are secreted into the conditioned medium. The Western blot was probed with a monoclonal antibody against the C-terminal His-tag as described in **Section 2.2.1**. Calculated molecular weight of the constructs: WIF-1_{WD}, 17424.9 g/mol, 156 amino acid residues; WIF-$1_{WD\text{-}EGF\text{-}I}$, 20886.7 g/mol, 188 amino acid residues. Both constructs carry an Asn88-linked glycosylation site (as observed in the crystal structures, see **Section 3.6**) which increases the observed molecular weight (MW) on the blot. MW standard BenchMark™ Protein Ladder from Invitrogen was used. The His-tagged protein detection threshold corresponds to approximately 1 mg/1itre of target protein in the conditioned medium (Aricescu et al., 2006). MW values are in kDa.

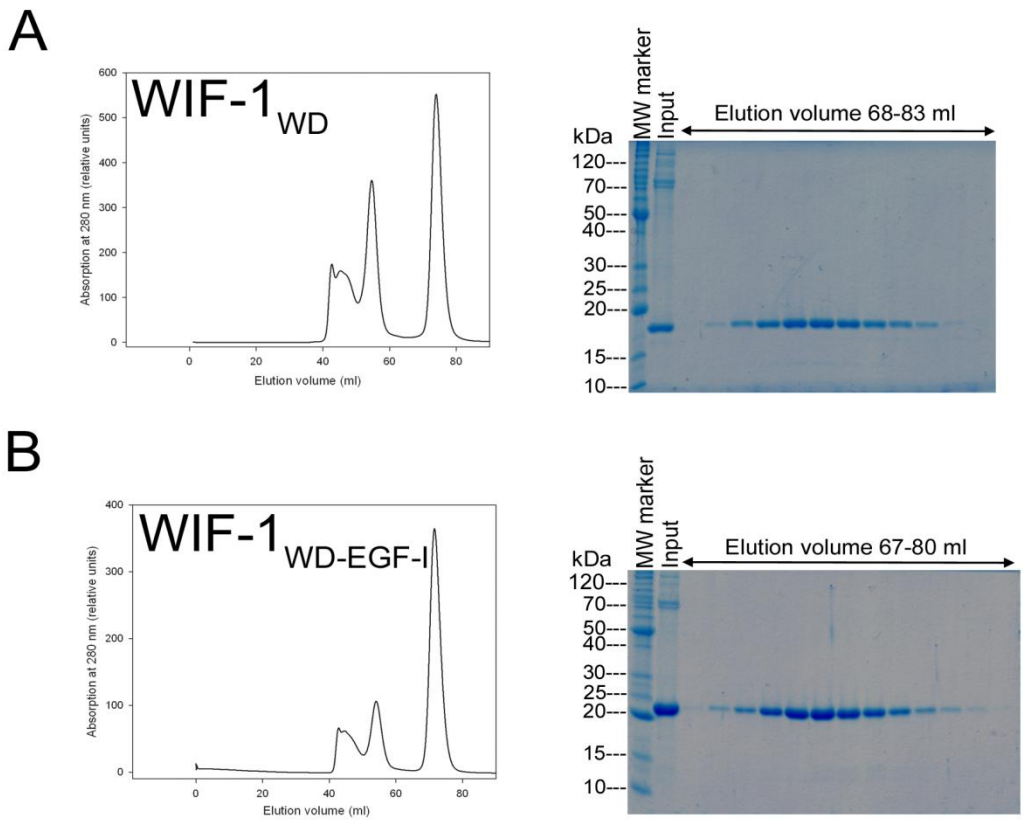


Figure 3.2. Size-exclusion chromatography (SEC) of WIF-1_{WD} (**A**) and WIF-$1_{WD\text{-}EGF\text{-}I}$ (**B**) constructs. Separation by SEC was performed using HiLoad 16/60 Superdex 75 column (GE Healthcare Life Sciences) at a flow speed of 0.8 ml/min in 20 mM Tris-HCl pH 8.0, 300 mM NaCl. Protein elution was monitored by absorbance at 280 nm (black curves, left). WIF-1 peak fractions were analysed on reducing SDS-polyacrylamide (15%) gels (right). Samples loaded onto SEC column are presented (designated as Input). Molecular weight (MW) marker: BenchMark Protein Ladder from Invitrogen. Gels were stained with SimplyBlue SafeStain from Invitrogen.

3.5 Crystallisation and structure determination of WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$

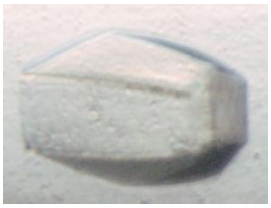
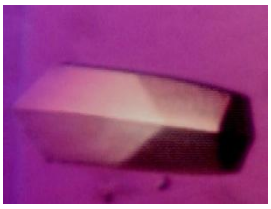
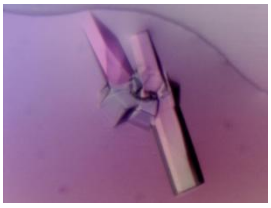
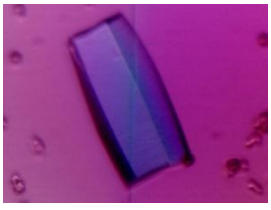
3.5.1 Crystallisation, and structure determination of WIF-1_{WD} and mass spectrometry analysis of WIF-1-bound ligands

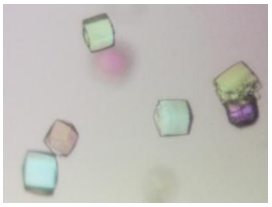
Prior to crystallisation, WIF-1_{WD} was enzymatically deglycosylated (Aricescu et al., 2006) and concentrated by ultrafiltration in SEC buffer (typically, 150–300 mM NaCl, 10–20 mM Tris-HCl pH 8.0) to 6–7 mg/ml. Sitting drop vapour diffusion crystallisation experiments were set up as described in **Section 2.3**. WIF-1_{WD} crystallised in various phosphate-

containing conditions (**Table 3.1**) but crystals diffracted X-rays relatively poorly, to 4–5 Å resolution. To improve the low diffraction-quality crystals, the primary amino groups (ε amino groups of lysines and an N-terminal amino group) of WIF-1_{WD} were methylated as described in **Section 2.3.2**. Crystals of the methylated WIF-1_{WD} formed during one day and diffracted X-rays to 1.9 Å resolution (**Figure 3.3**). A diffraction data set was collected from a single crystal (**Table 3.1**, crystal number eight) cryoprotected (as described in **Section 2.4**) at 100 K on beamline BM14 at the European Synchrotron Radiation Facility (ESRF). Statistical descriptors of the WIF-1_{WD} X-ray diffraction data set are presented in **Table 3.2**.

The WIF-1_{WD} structure was solved by molecular replacement using as a search model the WD coordinates determined by NMR spectroscopy (Liepinsh et al., 2006; **Section 2.6.1**). The initial electron density maps (after molecular replacement) of the WIF-1_{WD} structure revealed the architecture of a hydrophobic pocket which appeared to be occupied by lipid-like ligand. Matrix assisted laser desorption ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) was employed to identify any such small molecules (MW up to 3000 Da) bound to WIF-1 (as described in **Section 2.8**). A range of different phosphatidyl cholines were identified as the WIF-1 ligands (**Figure 3.4**). These MALDI-TOF-MS data were in excellent agreement with the X-ray crystallographic results. Cyclic rounds of model building and refinement resulted in a crystal structure of WIF-1_{WD} in complex with 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC) that has good values for crystallographic validation parameters (among the top 10 % best refined structures in the 1.85 Å ± 0.25 Å resolution range as evaluated with the program MolProbity (Davis et al., 2007)) (**Table 3.2**).

Table 3.1. Crystallisation of WIF-1_{WD}.

No.	Crystallisation conditions at 21 °C	Conc. (mg/ml)	Methylated ?	Crystal size (µm), in the longest dimension	Crystal image
1.	1 M $NH_4H_2PO_4$, 100 mM tri-sodium citrate pH 5.6	5.9	No	160	
2.	0.8 M NaH_2PO_4, 100 mM HEPES-Na pH 7.5, 0.8 M KH_2PO_4	5.9	No	100	
3.	Polyethylene glycol 3350 20 % w/v, 0.2 M KH_2PO_4	5.9	No	90	
4.	1 M $NH_4H_2PO_4$, 100 mM tri-sodium citrate pH 5.6, 8 % 1,3-propanediol	6.8	No	120	
5.	1 M $NH_4H_2PO_4$, 100 mM tri-sodium citrate pH 5.6, 8 % polypropylene glycol P400	6.8	No	200	
6.	1 M $NH_4H_2PO_4$, 100 mM tri-sodium citrate pH 5.6, 8 % polyethylene glycol 400	6.8	No	190	

7.	0.8 M $NH_4H_2PO_4$, 20 % ethylene glycol	6.1	No	80	
8.	30 % w/v PEG 8000, 200 mM sodium acetate, 100 mM sodium cacodylate pH 6.5	9.2	Yes	400	
9.	30 % w/v PEG 8000, 200 mM sodium acetate, 100 mM sodium cacodylate pH 6.5	9.2	Yes	430	

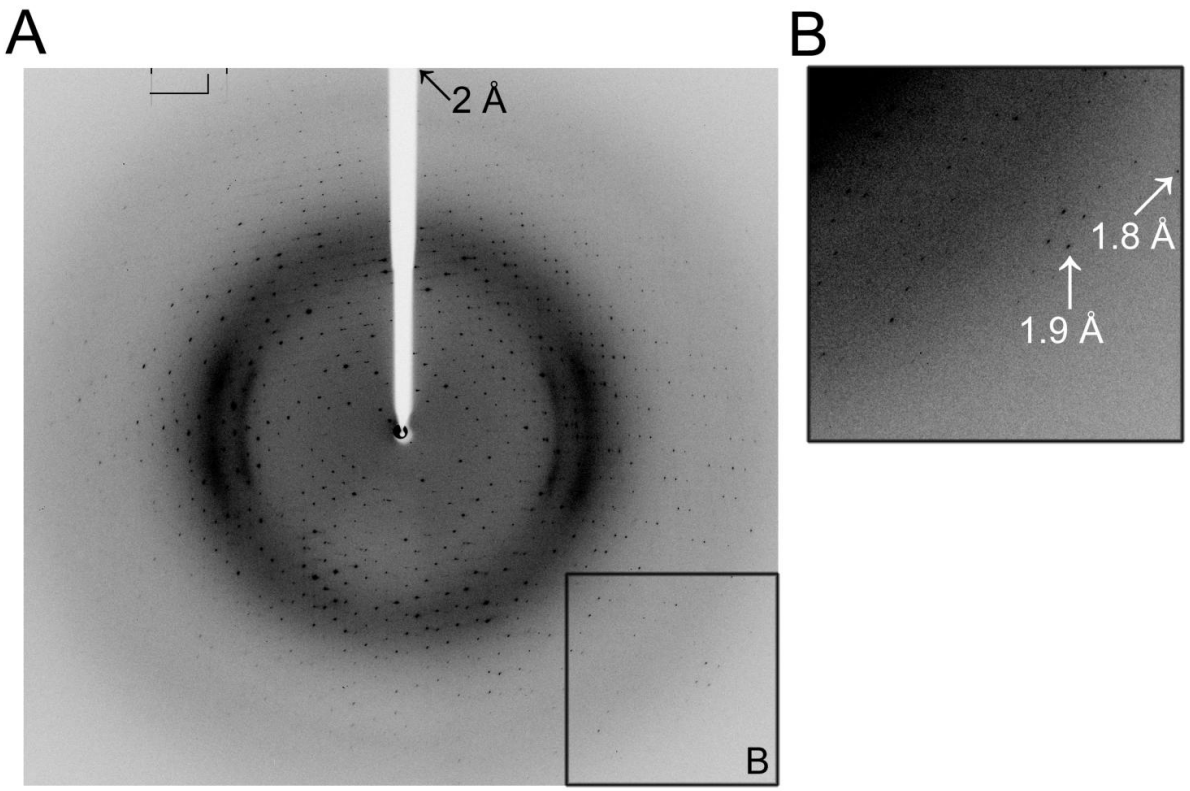


Figure 3.3. **A**, a representative X-ray diffraction image from a WIF-1_{WD} crystal (**Table 3.1**, number 8). The dimensions of the crystals were 200 × 50 × 50 µm. The crystal was exposed over a 0.7° oscillation range. The edge of the detector corresponds to a resolution of 2 Å (indicated with an arrow). **B**, high resolution reflections of the same diffraction image and corresponding resolution.

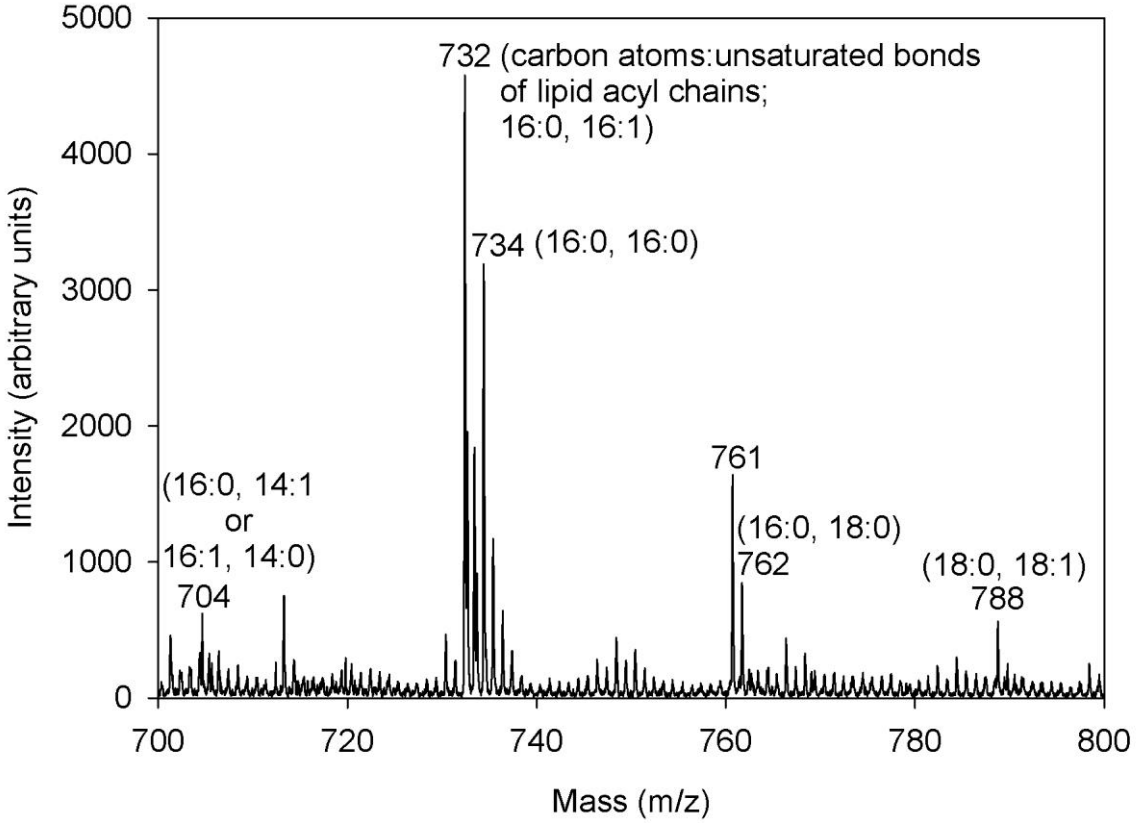


Figure 3.4. MALDI-TOF mass spectrum of WIF-1-bound ligands. The saturation and length of the 1,2-diacyl-phosphatidylcholine acyl chains are indicated in parentheses. Peaks of phosphatidylcholines in the MALDI-TOF mass spectrum were assigned according to Table 1 by Schiller et al., 1999.

Table 3.2. Crystallographic data collection and refinement statistics for the WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ crystals.

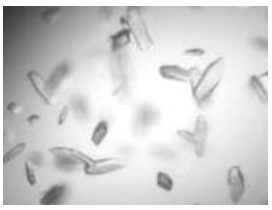
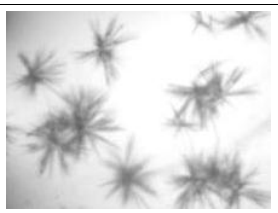
Construct	**WIF-1_{WD}**	**WIF-$1_{WD\text{-}EGF\text{-}I}$**
Data collection		
X-ray source	ESRF, BM14	ESRF, ID23.2
Resolution range (Å)	30.00–1.85 (1.93–1.85)	40.00–1.85 (1.99–1.85)
Space group	C222	C222
Cell dimensions (Å)	a = 49.9 b = 134.1 c = 60.4	a = 51.0 b = 134.3 c = 60.4
Solvent content (%) (molecules per asymmetric unit)	57 (1)	50 (1)
Wavelength (Å)	0.978	0.873
R_{merge} (%)	10.3 (70.0)	14.4 (94.5)
Total number of observations	93798 (7011)	133773 (25747)
Total number of unique observations	17655 (2003)	18293 (3576)
Mean ((I)/σ(I))	16.3 (1.4)	15.5 (1.9)
Completeness	98.8 (91.3)	100.0 (100.0)
Redundancy	5.3 (3.5)	7.3 (7.2)
Refinement		
Resolution range (Å)	29.16–1.85 (1.96–1.85)	37.42–1.85 (1.97–1.85)
Number of reflections	17575 (2652)	17107 (2507)
R_{work} (%)	20.4 (24.9)	18.3 (23.6)
R_{free} (%)	22.9 (27.6)	21.3 (28.7)
Number of atoms (protein/DPPC/GlcNAc, fucose/sucrose octasulfate/water)	1306 (1155/50/14/0/83)	1626 (1410/50/14/0/151)
Wilson B factor ($Å^2$)	30	18
Average B factor (protein/DPPC/GlcNAc, fucose/sucrose octasulfate/water) ($Å^2$)	42 (40/71/57/–/47)	37 (36/51/37/–/41)
RMSD bonds (Å)	0.010	0.015
RMSD angles (degrees)	1.1	2.1
Ramachandran statistics		
Favoured (%)	97.8	98.8
Number of outliers	0	0

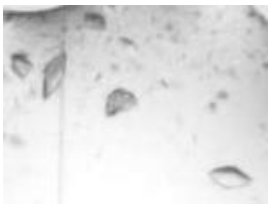
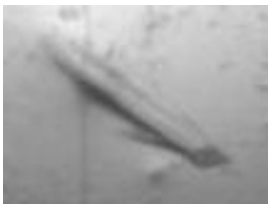
3.5.2 Crystallisation and structure determination of WIF-$1_{WD\text{-}EGF\text{-}I}$

Similarly to the WIF-1_{WD} construct (**Section 3.5.2**), WIF-$1_{WD\text{-}EGF\text{-}I}$ was enzymatically deglycosylated, methylated and concentrated by ultrafiltration in SEC buffer (300 mM NaCl, 20 mM Tris-HCl pH 8.0) to 7.8 mg/ml prior to crystallisation. Sitting drop vapour diffusion crystallisation experiments were set up as described in **Section 2.3**. WIF-$1_{WD\text{-}EGF\text{-}I}$ crystallised in various phosphate-, ammonium- and polyethylene glycol-containing conditions (**Table 3.3**). A diffraction data set was collected from a single crystal (**Table 3.3**, crystal number one) cryoprotected (as described in **Section 2.4**) at 100 K on beamline ID23.2 at the ESRF (**Figure 3.5**). Statistical descriptors of the WIF-$1_{WD\text{-}EGF\text{-}I}$ X-ray diffraction data set are presented in **Table 3.2**.

The crystal structure of WIF-$1_{WD\text{-}EGF\text{-}I}$ was phased by molecular replacement using the WIF-1_{WD} crystal structure with DPPC omitted. DPPC was again built into additional electron density during model building and refinement (as described in **Section 2.6.1**). The refined WIF-$1_{WD\text{-}EGF\text{-}I}$ structure has good values for crystallographic validation parameters (among the top 2 % best refined structures of 1.85 Å ± 0.25 Å resolution as evaluated with the program MolProbity (Davis et al., 2007)) (**Table 3.2**).

Table 3.3. Crystallisation of WIF-$1_{WD\text{-}EGF\text{-}I}$.

No.	Crystallisation conditions at 21 °C	Conc. (mg/ml)	Methylated ?	Crystal size (µm), in the longest dimension	Crystal image
1.	1.3 M di-ammonium tartrate, 100 mM bis-Tris propane pH 7.0	7.8	Yes	100–200	
2.	1.4 M di-ammonium tartrate, 100 mM Tris pH 8.5	7.8	Yes	100–120	
3.	1.5 M $(NH_4)_2SO_4$, 100 mM bis-Tris propane pH 7.0	7.8	Yes	200	
4.	1.6 M $(NH_4)_2SO_4$, 10 % v/v dioxane, 100 mM MES pH 6.5	7.8	Yes	200	
5.	100 mM acetate pH 4.5, 0.8 M NaH_2PO_4, 1.2 M K_2HPO_4, final pH 6.7	7.8	Yes	150	
6.	30 % w/v polyethylene glycol 6000, 100 mM MES pH 6.0	7.8	Yes	90	

7.	1.6 M $(NH_4)_2SO_4$, 10 mM $MgCl_2$, 50 mM HEPES-Na pH 7.0	7.8	Yes	290	
8.	1.17 M K_2HPO_4, 0.63 M NaH_2PO_4, final pH 6.9	7.8	Yes	60–230	
9.	0.5 % w/v polyethylene glycol monomethyl ether 5000, 0.8 M potassium sodium tartrate, 100 mM Tris pH 8.5	7.8	Yes	190	
10.	25 % w/v polyethylene glycol 3350, 200 mM $(NH_4)_2SO_4$, 100 mM bis-Tris pH 5.5	7.8	Yes	80	
11.	28 % w/v polyethylene glycol monomethyl ether 2000, 100 mM bis-Tris pH 6.5	7.8	Yes	250	

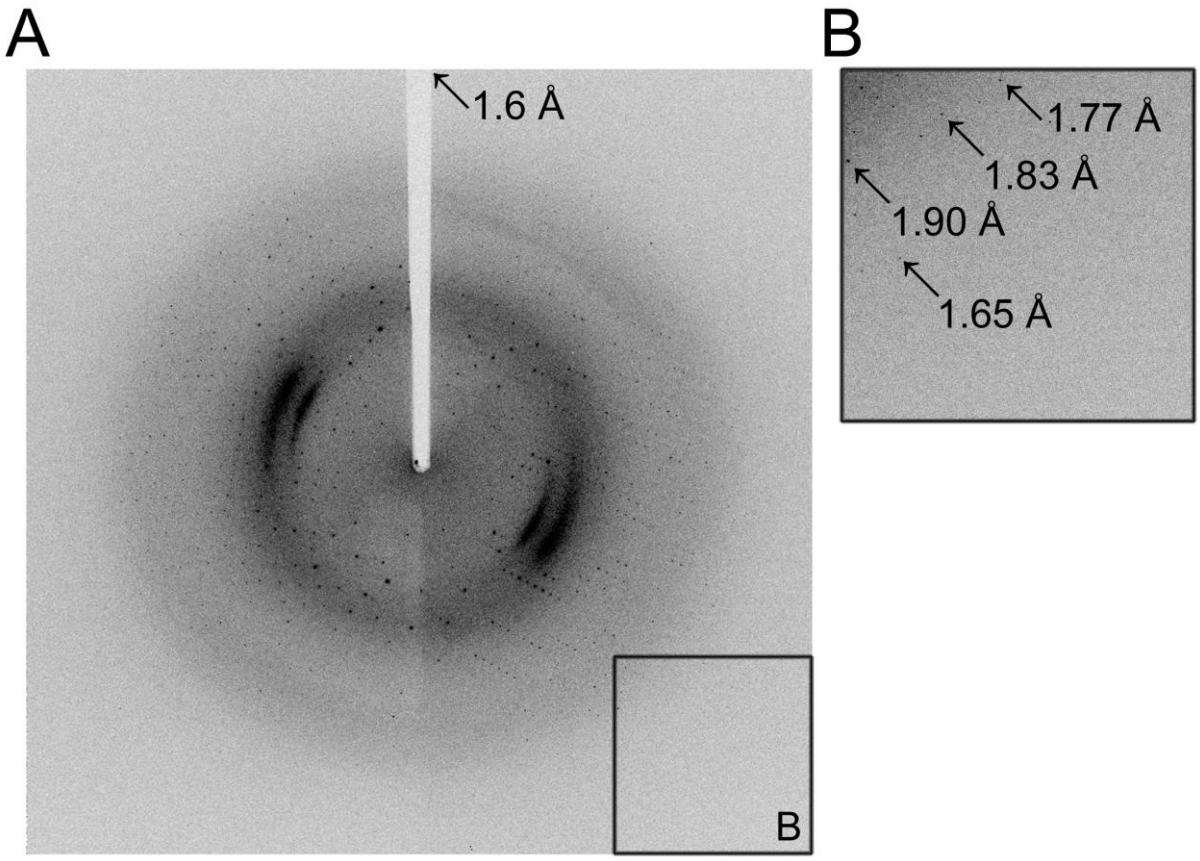


Figure 3.5. **A**, a representative X-ray diffraction image from a WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal (**Table 3.3**, crystal number one). The dimensions of the crystals were 100 × 50 × 20 µm. The crystal was exposed over a 1.0° oscillation range. The edge of the detector corresponds to a resolution of 1.6 Å (indicated with an arrow). **B**, high resolution reflections of the same diffraction image and corresponding resolution.

3.6 Key features of the $WIF_{WD\text{-}EGF\text{-}I}$ crystal structure

3.61 WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal packing

WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ constructs crystallised in the same C222 space group with similar unit cell dimensions (**Table 3.2**). Thus the solvent content in the less tightly packed WIF-1_{WD} crystals (57 % solvent, **Figure 3.6**, **A**) was higher compared to WIF-$1_{WD\text{-}EGF\text{-}I}$ crystals (50 % solvent, **Figure 3.6**, **B**). It has previously been commented that the more tightly packed crystals (containing lower solvent content) tend to diffract X-rays better, i.e. to higher resolution (Kantardjieff and Rupp, 2003). However, WIF-$1_{WD\text{-}EGF\text{-}I}$ crystals diffracted X-rays similar to the WIF-1_{WD} crystals as judged by X-ray diffraction data quality indicators (**Table 3.2**).

In both cases, the reductive lysine methylation was essential to obtain crystals that diffracted X-rays beyond 2 Å resolution. Notably, methylated (thus more hydrophobic) Lys130 contacts a hydrophobic region formed by Tyr37, Trp39 and Leu56 on the neighbouring molecule in the crystal (**Figure 3.6**, **C**). These additional interactions may contribute to the order of the molecules in the crystal, and thus better X-ray diffraction qualities (the non-methylated WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ crystals did not diffract X-rays beyond 3 Å resolution).

The structure of the WD is essentially identical in the WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structures (RMSD of 0.17 Å between equivalent Cα atoms for 144 residues). The WD from the WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structure is used for further analysis presented in **Chapter 3** because the WD is better ordered in the WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structure as judged by the final refinement statistics (R_{work} and R_{free}, **Table 3.2**) and manual inspection of the electron density maps.

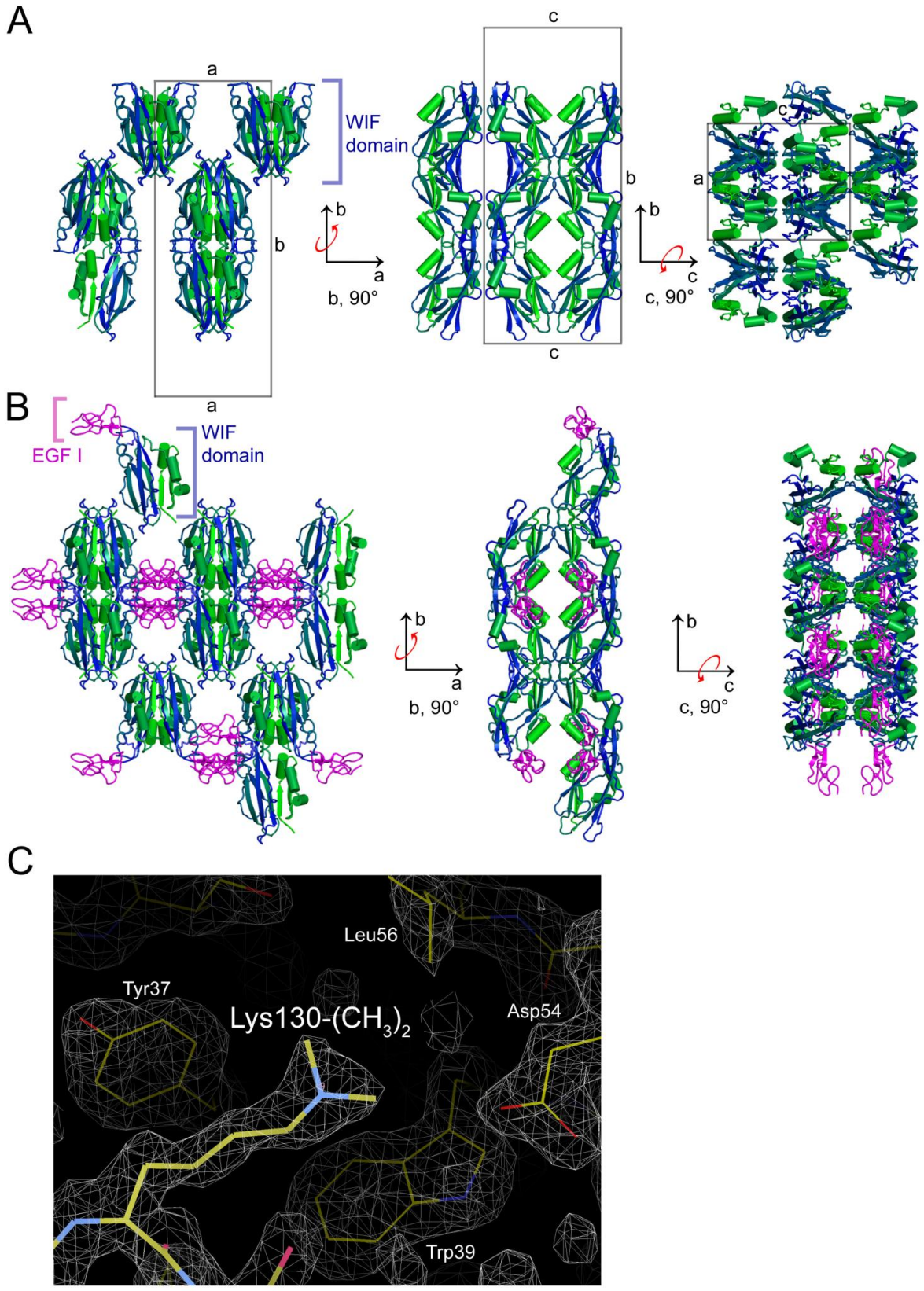


Figure 3.6. Crystal packing in the WIF-1$_{WD}$ (**A**) and WIF-1$_{WD-EGF-I}$ (**B**) crystals. WD is coloured in green-to-blue transition, EGF-like domain I is coloured in magenta. The WD helices are represented as green cylinders. Unit cell dimensions are indicated and correspond

to **Table 3.2**. Three unit lattice constants (a, b and c) are indicated. Three views of the crystal packing (both in **A** and in **B**) were obtained by 90° rotation of the unit cell around the indicated unit lattice vectors (**b** and **c**). **C**, an example of a well-defined electron density map ($2F_o$ - F_c, contoured at 1 σ level) for a dimethylated lysine (Lys130) involved in a crystal contact with the side chain residues of a neighbouring WIF-$1_{WD-EGF-I}$ molecule. Protein cartoons in **A** and **B** (and most other figures in this thesis) were prepared using program PyMOL (Schrödinger, LLC.), cartoons in **C** – with the program Coot (Emsley et al., 2010).

3.62 The WIF domain fold and surface properties

The WIF domain (WD) fold consists of a β sandwich which is splayed apart along one edge by an insertion of three short helices (one 3_{10} and two α helices proximal to the N-terminus) (**Figure 3.7**, **B**–**C**). The β sandwich consists of ten antiparallel β strands of variable length (3 to 13 amino acid residues) that are arranged into two β sheets. The two β sheets pack loosely against each other, forming a large cavity (1556 $Å^3$, calculated with CASTp using 1.4 Å probe radius; Dundas et al., 2006) which penetrates deep into the core of the WD. A gap of ~10 Å between the α1 and 3_{10} helices provides a gateway into this hydrophobic cavity. Thus, although the residues facing inward to the interior of the domain are primarily hydrophobic, the WD lacks the close packed core of a standard β sandwich. The two β sheets are locked together at the base of the sandwich by a disulfide bond between residues Cys140 and Cys177 (**Figure 3.7**, **C**). Sequence alignments suggest this disulfide is conserved between species and is a canonical feature of the WD fold when present in other proteins (**Figure 3.8**, **A**).

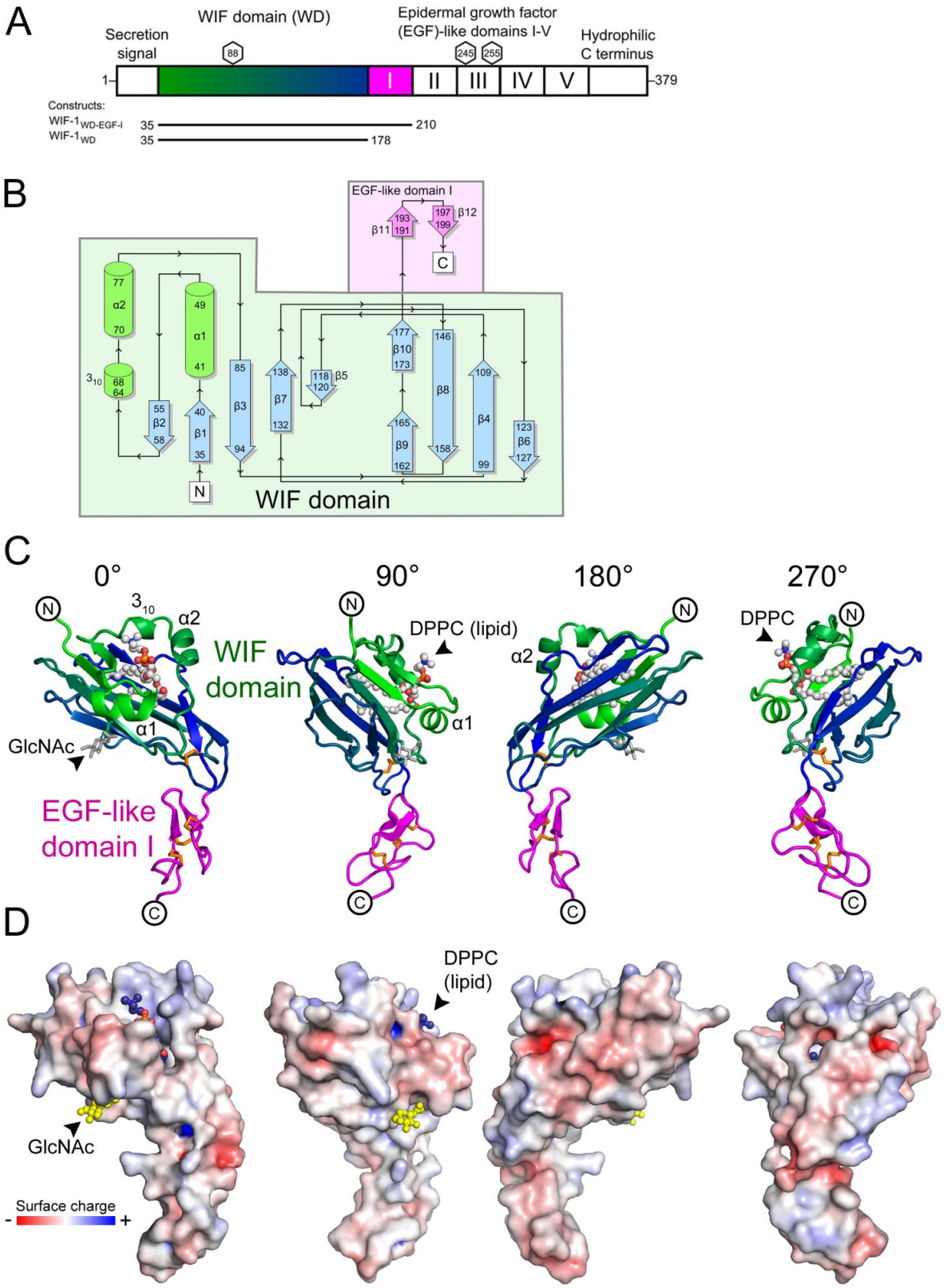
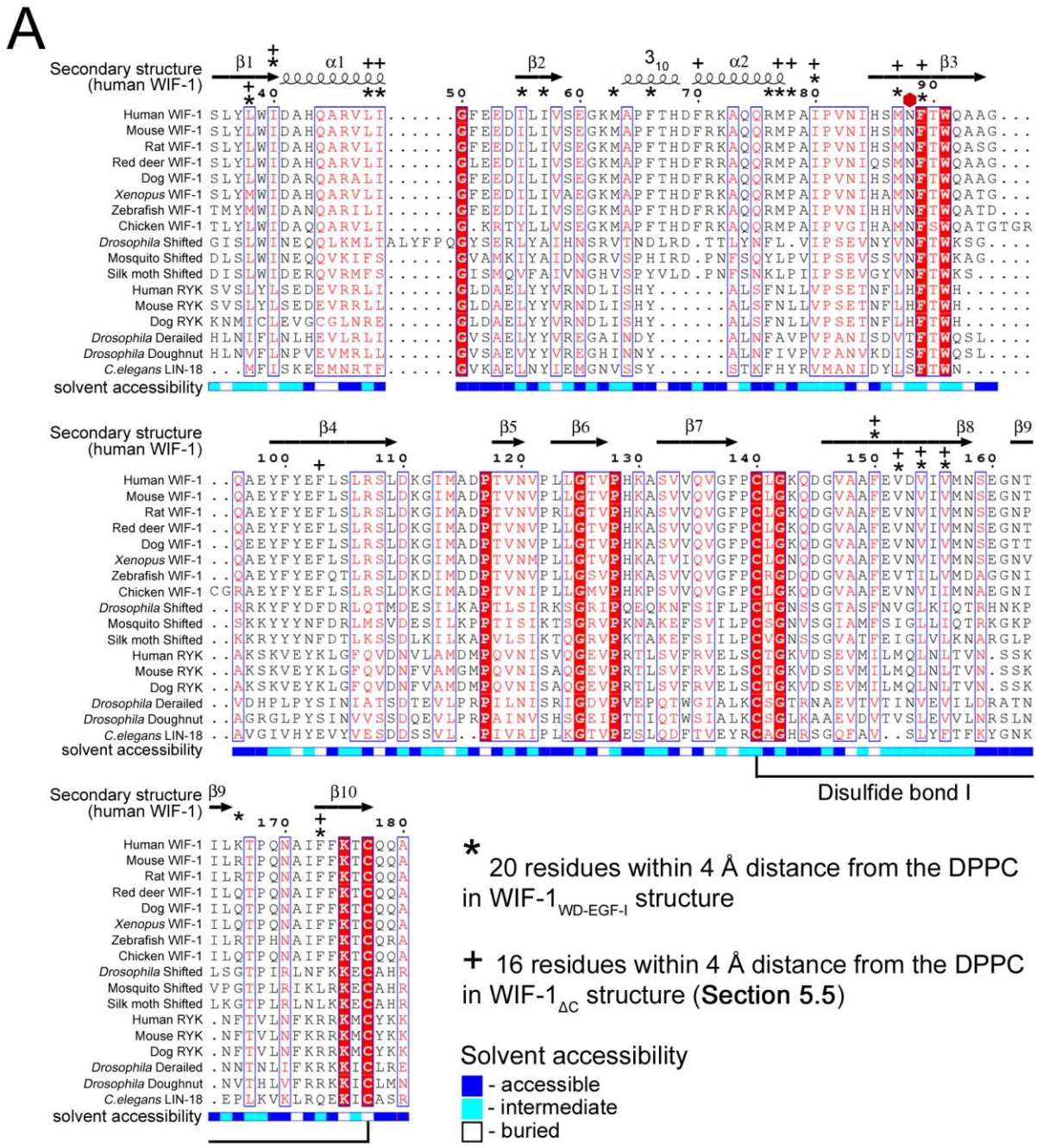

A
WIF domain (WD)
Secretion signal
Epidermal growth factor (EGF)-like domains I-V
Hydrophilic C terminus
88
245
255
1
379
I
II
III
IV
V
Constructs:
WIF-1WD-EGF-I 35 210
WIF-1WD 35 178
B
EGF-like domain I
WIF domain
N
C
C
0°
90°
180°
270°
N
C
3 10
α2
α1
WIF domain
DPPC (lipid)
GlcNAc
EGF-like domain I
D
DPPC (lipid)
GlcNAc
Surface charge
-
+

Figure 3.7. **A**, schematic domain organization of the human WIF-1. Asn88-, Asn245-, and Thr255-linked glycosylation sites are marked by hexagons. WD is coloured in green-to-blue transition, EGF-like domain I is coloured in magenta. Amino acid sequence boundaries of the WIF-$1_{WD\text{-}EGF\text{-}I}$ and WIF-1_{WD} constructs are shown below. **B**, topology diagram of the human WIF-$1_{WD\text{-}EGF\text{-}I}$ construct. α helices are shown as green cylinders, β strands are shown as light blue arrows. Amino acid sequence boundaries of each secondary structure element correspond to the full length human WIF-1 numbering (including secretion signal). The topology diagram was generated using program PDBsum (Laskowski, 2009) for the WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structure and modified using program Inkscape (http://www.inkscape.org/). Secondary structure boundaries defined by PDBsum for the end of α1 (residue 50) and the start of α2 (residue 73) were changed (to residues 49 and 70, respectively) after manual inspection of the structure. **C**, ribbon diagram of WIF-$1_{WD\text{-}EGF\text{-}I}$ in four views that differ by a 90° rotation about a vertical axis. WIF-$1_{WD\text{-}EGF\text{-}I}$ is colour coded as in **A**. Asn88-linked N-acetylglucosamine (GlcNAc, grey), and disulfide bonds (orange) are shown as sticks. DPPC is shown in spheres (carbons – grey; nitrogen – blue; oxygens – red; phosphorus – orange). Two α and one 3_{10} helix, N- and C-termini are indicated. **D**, electrostatic properties of the WIF-$1_{WD\text{-}EGF\text{-}I}$ surface. The protein is shown as solvent-accessible surface coloured by electrostatic potential contoured at ± 8 kT/e (red, acidic/negative; blue, basic/positive). The GlcNAc moiety on Asn88 (WD) is shown as yellow spheres (indicated with an arrow). DPPC is shown in colour coded spheres (the head group is indicated with an arrow in the 90° view).


A
Secondary structure (human WIF-1)
β1 α1 β2 3$_{10}$ α2 β3
Human WIF-1
Mouse WIF-1
Rat WIF-1
Red deer WIF-1
Dog WIF-1
Xenopus WIF-1
Zebrafish WIF-1
Chicken WIF-1
Drosophila Shifted
Mosquito Shifted
Silk moth Shifted
Human RYK
Mouse RYK
Dog RYK
Drosophila Derailed
Drosophila Doughnut
C.elegans LIN-18
solvent accessibility
β4 β5 β6 β7 β8 β9
Disulfide bond I
β9 β10
* 20 residues within 4 Å distance from the DPPC in WIF-1$_{WD\text{-}EGF\text{-}I}$ structure
+ 16 residues within 4 Å distance from the DPPC in WIF-1$_{\Delta C}$ structure (Section 5.5)
Solvent accessibility
- accessible
- intermediate
- buried

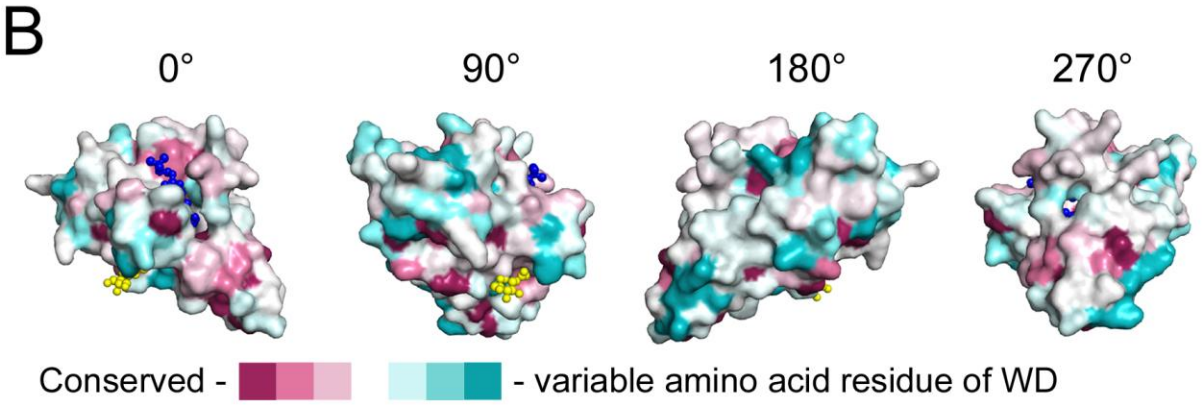

B
0°
90°
180°
270°
Conserved -
- variable amino acid residue of WD

Figure 3.8. **A**, Amino acid sequence alignment of the WIF domains. Alignments of human (residues 35–180, SwissProt identification accession number Q9Y5W5), mouse (residues 35–180, SwissProt Q9WUA1), rat (residues 35–180, SwissProt Q6IN38), red deer (*Cervus elaphus*, residues 3–148, UniProtKB/TrEMBL A6YLN5), dog (residues 35–180, National Center for Biotechnology Information, U.S. National Library of Medicine (NCBI) identification accession number XP_538269.2), *Xenopus laevis* (residues 30–175, Swiss-Prot Q9W6F8), zebrafish (residues 33–178, SwissProt Q9W6F9), chicken (residues 30–179, NCBI XP_416072.2) WIF domains of WIF-1 proteins, and WIF domains of *Drosophila* Shifted (residues 116–264, SwissProt Q9W3W5), African malaria mosquito Shifted (*Anopheles gambiae*, residues 46–189, XP_311028.4), silk moth Shifted (*Bombyx mori*, residues 38–180, NCBI NP_001040227.1), human RYK (residues 60–194, SwissProt P34925), mouse RYK (residues 47–181, SwissProt Q01887), dog RYK (residues 36–170, NCBI XP_534269.2), *Drosophila* Derailed (residues 21–158, SwissProt Q27324) and Doughnut (residues 46–183, SwissProt Q9V422), *Caenorhabditis elegans* Abnormal cell lineage protein 18 (residues 1–129, TrEMBL Q18041), respectively. The sequences were aligned using MultAlin (Corpet, 1988; http://multalin.toulouse.inra.fr/) and formatted with ESPript (Gouet et al., 1999; http://espript.ibcp.fr/). Numbering corresponds to the full length human WIF-1. Secondary structure assignments (calculated using program PDBsum (Laskowski, 2009; http://www.ebi.ac.uk/pdbsum/)) for the WD of WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structure are displayed above the alignment and correspond to **Figure 3.7**, **B**. Disulfide bridge I (between Cys140 and Cys177) and the solvent accessibilities (calculated with ESPript) of WIF-1 residues are presented below the alignments (blue, accessible; cyan, intermediate; white, buried). An N-linked N-acetylglucosamine is marked by a filled red hexagon above glycosylated Asn88. Twenty residues within 4 Å of the DPPC in the WIF-$1_{WD\text{-}EGF\text{-}I}$ structure are marked with an asterisk (*). Sixteen residues within 4 Å of the DPPC in the WIF-$1_{\Delta C}$ structure are marked with a plus sign (+). WIF-$1_{\Delta C}$ is described in **Section 5.5**. **B**, The WD surface coloured by residue conservation (conserved – magenta, variable – cyan). Fourteen Wnt-binding WDs (**A**) are included in the sequence conservation analysis. WDs from Hh-binding Shifted are excluded.

The WD surface lacks extensive positively or negatively charged areas. Ten surface exposed side chains of lysine and arginine residues are relatively evenly distributed as well as fourteen aspartate and glutamate side chains (**Figure 3.7**, **D**). The only exception is a consecutive acidic Glu52-Glu53-Asp54 stretch between the helix α1 and strand β2. The significance of this negatively charged stretch remains unknown. Charge distribution on the WD is further discussed in the context of WIF-1 binding studies to the negatively charged glycosaminoglycans in **Chapter 5**.

The function of the WDs present in RYKs (single-pass transmembrane Wnt receptors) and vertebrate WIF-1 proteins is thought to be evolutionarily conserved: they all bind different Wnt proteins. Thus, it is possible that the WD surface-exposed residues involved in Wnt binding are evolutionarily conserved too. We have analysed which surface exposed residues of the WD are conserved among the WDs that are present in fourteen, presumably, Wnt-binding proteins, RYK receptors and vertebrate WIF-1 proteins (**Figure 3.8**). Most of the conserved surface exposed residues are located in the approximately rectangular region defined by four points: (1 and 2) the terminal residues of helices α1 and α2, (3) C-terminus of the WD (Gln179), and (4) middle of the longest WD strand β8 (**Figure 3.8**, **B**, 0° and 270° views). On the opposite side of the WD (**Figure 3.8**, **B**, 90° and 180° views), the surface exposed residues are less conserved and are accompanied by the Asn88-linked glycosylation site. The glycosylated WD surface surrounded by the evolutionarily divergent surface exposed residues is less likely to serve as a surface which is recognised by Wnts. Conservation of the WD residues is further discussed in the context of WIF-1–Wnt3a surface plasmon resonance binding studies and cellular Wnt3a/β-catenin signalling inhibition assays in **Chapters 4**.

In addition, we could not identify any obvious similarities (e.g. conserved regions of the surface exposed amino acid residues) between the WIF-1 WD and two Frizzled cysteine-rich

domain structures (Dann et al., 2001; these domains serve as Wnt-binding modules in Frizzled receptors), suggesting different molecular surfaces are used for Wnt protein recognition.

3.63 The WIF domain lipid-binding pocket

Unexpectedly, the crystal structures of the WD described here differ markedly (RMSD of 2.64 Å for 144 equivalent Cα atoms) from the previously reported WIF-1 WD structure determined by solution NMR spectroscopy (Liepinsh et al., 2006; see **Section 3.64**). The topology of the WD is largely identical between the crystal and solution structures, but there is no cavity within the β sandwich in the NMR structure, where two β sheets form a relatively close packed protein core. In contrast, the electron density maps for the WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structures show clear evidence for a lipid-like molecule bound as an integral part of the WD core (**Figure 3.7**, **C**–**D**). This additional molecule fills the large hydrophobic cavity, stabilizing the splayed out conformation of the WD β sandwich. No supplemental lipids were used during protein production and crystallization, therefore the molecule must have been incorporated into the WIF-1 WD during mammalian expression (the NMR structure ($WD_{refolded}$) was determined from protein expressed as inclusion bodies in *E. coli* and refolded; Liepinsh et al., 2006). Hydrophobic ligands were extracted from HEK-293 cells-expressed WIF-$1_{\Delta C}$ and identified by mass spectrometry as a series of phosphatidylcholines differing in acyl chain length (C14–C18) and degree of saturation (C16:0–1, C18:0–1) (see **Section 3.5.1** and **Figure 3.4**). The major species consists of two acyl chain lipids, each chain of 16 carbon atoms, a chemistry typified by the common mammalian cell membrane lipid 1,2-dipalmitoylphosphatidylcholine (DPPC). DPPC was built into the lipid-like density and refined successfully as an integral part of the WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structures (**Figures 3.7**–**9**). The acyl chains of DPPC fit in an extended conformation snuggly within the WD core, filling the available volume of the

pocket, whilst the phosphatidylcholine head group is surface exposed making interactions with three residues (**Figure 3.9**). The ethylene group of the head group forms a hydrophobic interaction with Phe66, the glycerol backbone packs against Ile49 and the phosphatidyl oxygen makes a hydrogen bond with Arg76 (**Figure 3.9**).

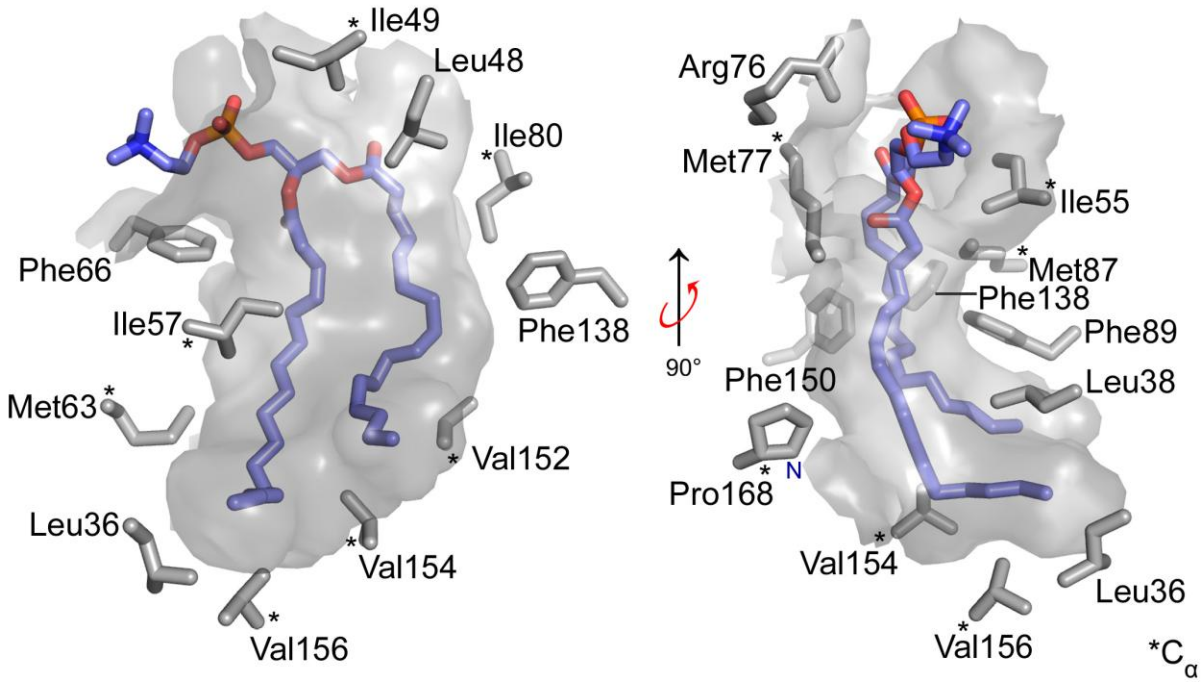


Figure 3.9. The WD lipid-binding pocket (grey surface) in two views that differ by a 90° rotation about a vertical axis. DPPC is shown as sticks (carbons – violet; nitrogen – blue; oxygens – red; phosphorus – orange). Residues within 4.5 Å from DPPC are shown as grey sticks. Cα atoms are marked with asterisks where necessary.

3.64 Exchange of the WIF domain-bound lipids

To assess the lipid-binding properties of mammalian cell expressed WIF-1 I incubated purified WIF-$1_{\Delta C}$ with a fluorescent lipid, 16:0 Liss Rhod PE (1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine-N-(lissamine rhodamine B sulfonyl) ammonium salt). Gel filtration revealed WIF-$1_{\Delta C}$ and 16:0 Liss Rhod PE co-eluted for a sample incubated in the presence of 1 % CHAPS suggesting that phosphatidylcholines bound within the core of the WD, such as DPPC, could be exchanged in the presence of detergent (**Figure 3.10**). Under similar conditions (including 1 % CHAPS), no binding was observed between WIF-$1_{\Delta C}$ (50 μM) and

a fluorescent head group of 16:0 Liss Rhod PE alone (sulforhodamine B, 160 μM) consistent with the diacyl lipid chains of 16:0 Liss Rhod PE mediating the binding to WIF-$1_{\Delta C}$. Finally, in the light of the detergent interaction mapped to the WIF-1 WD surface in the previously reported solution NMR analysis (Liepinsh et al., 2006), I probed WIF-$1_{\Delta C}$ for surface exposed lipid-binding sites using a lipid strip. I observed no binding of WIF-$1_{\Delta C}$ to 15 different, biologically abundant, lipids in the absence of 1 % CHAPS (**Figure 3.10**, **F**–**G**). These analyses of the lipid-binding properties, in combination with the WD structures of mammalian cell expressed constructs, reveal DPPC, and closely related PCs, act as physiological components of WIF-1 and saturate the lipid-binding capacity of the WD. Sequence alignments indicate that the hydrophobic nature, and, in general, the size of the residues which line the lipid-binding pocket in the WIF-1 WD are conserved in other WD containing proteins (**Figure 3.8)** suggesting that PCs such as DPPC may be integral to WDs in general (see also **Section 3.7** regarding putative RYK WD–lipid interactions).

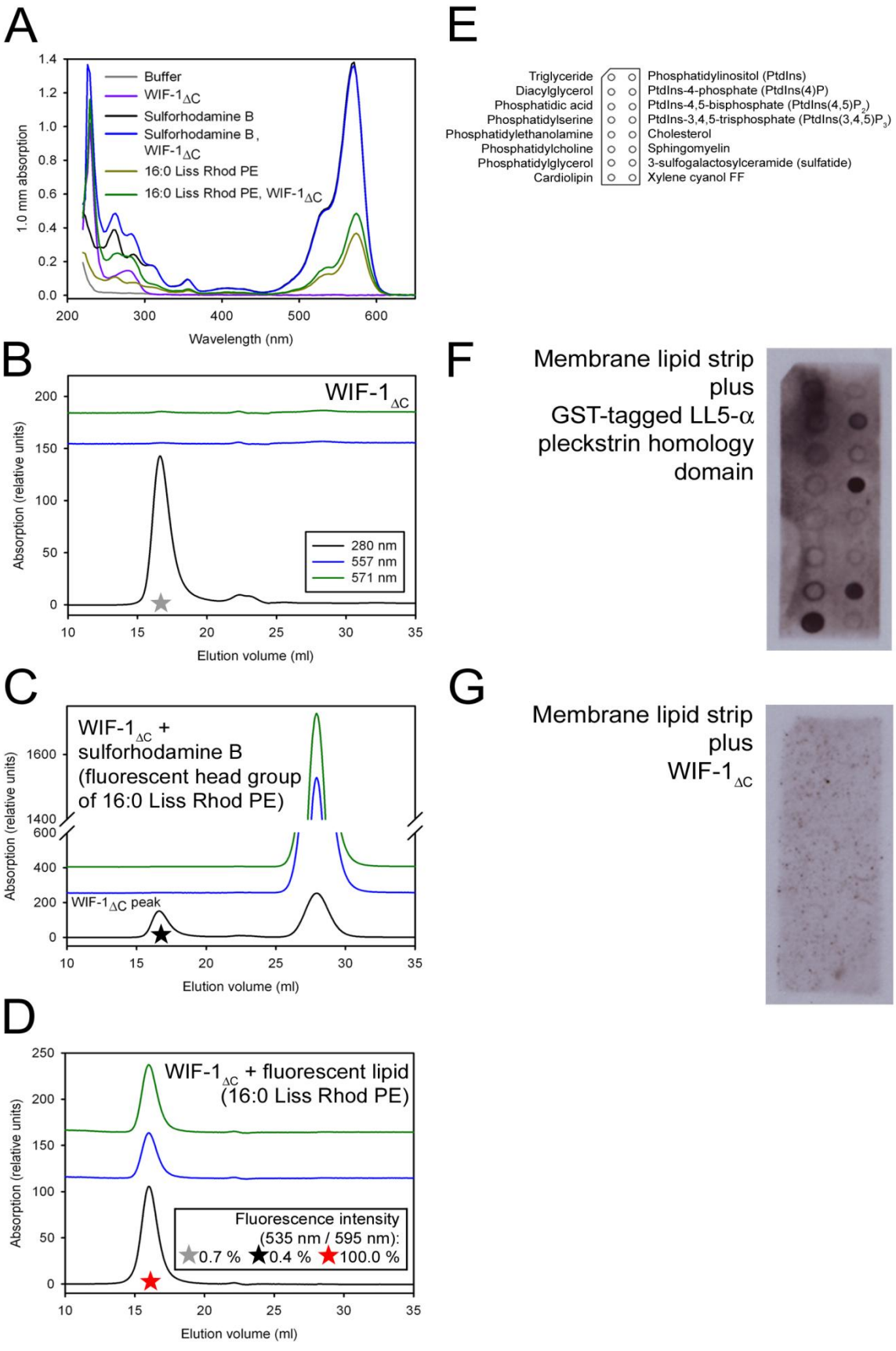

A
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1.0 mm absorption
200
300
400
500
600
Wavelength (nm)
Buffer
WIF-1ΔC
Sulforhodamine B
Sulforhodamine B, WIF-1ΔC
16:0 Liss Rhod PE
16:0 Liss Rhod PE, WIF-1ΔC
B
WIF-1ΔC
200
150
100
50
0
Absorption (relative units)
280 nm
557 nm
571 nm
10
15
20
25
30
35
Elution volume (ml)
C
WIF-1ΔC + sulforhodamine B (fluorescent head group of 16:0 Liss Rhod PE)
1600
1400
600
400
200
0
Absorption (relative units)
WIF-1ΔC peak
10
15
20
25
30
35
Elution volume (ml)
D
WIF-1ΔC + fluorescent lipid (16:0 Liss Rhod PE)
250
200
150
100
50
0
Absorption (relative units)
Fluorescence intensity (535 nm / 595 nm):
0.7 %
0.4 %
100.0 %
10
15
20
25
30
35
Elution volume (ml)
E
Triglyceride
Diacylglycerol
Phosphatidic acid
Phosphatidylserine
Phosphatidylethanolamine
Phosphatidylcholine
Phosphatidylglycerol
Cardiolipin
Phosphatidylinositol (PtdIns)
PtdIns-4-phosphate (PtdIns(4)P)
PtdIns-4,5-bisphosphate (PtdIns(4,5)P2)
PtdIns-3,4,5-trisphosphate (PtdIns(3,4,5)P3)
Cholesterol
Sphingomyelin
3-sulfogalactosylceramide (sulfatide)
Xylene cyanol FF
F
Membrane lipid strip plus GST-tagged LL5-α pleckstrin homology domain
G
Membrane lipid strip plus WIF-1ΔC

Figure 3.10. WIF-1 interactions with a fluorescent DPPC analogue and biologically abundant lipids. **A**, absorption spectra of solutions used in WIF-1–fluorescent molecule binding studies: buffer (PBS, 1 % CHAPS; grey curve), WIF-$1_{\Delta C}$ (50 μM, purple), sulforhodamine B (160 μM, black), sulforhodamine B (160 μM) plus WIF-$1_{\Delta C}$ (50 μM) (blue), 16:0 Liss Rhod PE (10 μl (1 mg/ml in $CHCl_3$) in 270 μl buffer; dark yellow), 16:0 Liss Rhod PE plus WIF-$1_{\Delta C}$ (50 μM) (10 μl (1 mg/ml in $CHCl_3$) in 270 μl buffer/WIF-$1_{\Delta C}$; green). **B–D**, size-exclusion chromatography of the purified WIF-$1_{\Delta C}$ incubated: without fluorescent molecules (**B**), with sulforhodamine B (fluorescent head group of 16:0 Liss Rhod PE) (**C**) and with 16:0 Liss Rhod PE (**D**). Separation by size exclusion was performed using Superdex 200 10/300 column (GE Healthcare Life Sciences) at a flow speed of 0.4 ml/min in PBS at room temperature. Protein and fluorescent molecule elution was monitored by absorbance at 280, 557 and 571 nm (black, blue and green curves, respectively) simultaneously. Binding between WIF-$1_{\Delta C}$ and 16:0 Liss Rhod PE was confirmed by fluorescence (excitation at 535 nm, emission at 595 nm) intensity measurements: intensity of WIF-$1_{\Delta C}$ peak fractions containing 16:0 Liss Rhod PE (indicated with a red star in **D**) was hundred fold (100 %) higher than WIF-$1_{\Delta C}$ alone (0.7 %; indicated with a grey star in **B**) or WIF-$1_{\Delta C}$ incubated with sulforhodamine B (0.4 %; indicated with a black star in **C**). **E**, details of the 15 commercially available, pre-spotted lipids on the hydrophobic membrane strip (Echelon Biosciences Inc.). **F**, Western blot of the membrane strip (**E**) after incubation with a glutathione S-transferase (GST)-tagged LL5-α pleckstrin homology domain (MultiPIP Grip, Echelon Biosciences) and anti-GST horseradish peroxidase (HRP) conjugate (GE Healthcare Life Sciences) as a positive control experiment. HRP was detected by chemiluminescence using the Amersham ECL detection kit (GE Healthcare Life Sciences). Dark round spots on the strip indicate protein–lipid interactions. **G**, Western blot of the membrane strip (**E**) after incubation with a His-tagged WIF-$1_{\Delta C}$, anti-PentaHis monoclonal primary antibody (Qiagen) and goat anti-mouse IgG HRP conjugate (Sigma).

3.65 Comparison of the WIF domain structures determined by NMR spectroscopy and X-ray crystallography

As noted in **Section 3.63**, the structure of the refolded WIF domain ($WD_{refolded}$) of human WIF-1 differs markedly from the WD crystal structures presented here in three interrelated aspects.

First, the human WIF-1 WD for NMR studies was overexpressed in inclusion bodies of bacteria *Escherichia coli* and refolded. Thus, the WD used for NMR studies was not exposed to its potential natural ligands that would be accessible in human cells. As a result, in contrast to the crystal WD–lipid complex, the $WD_{refolded}$ lacks the bound lipid (**Figure 3.11**, **A**–**C**). Lack of the bound lipid may possibly explain why the lipid-binding pocket is three times smaller in the $WD_{refolded}$ compared to the WD crystal structure (438 $Å^3$ versus 1556 $Å^3$; calculated with CASTp; Dundas et al., 2006): in order to shield hydrophobic residues lining

the exterior of the lipid-binding pocket from the hydrophilic solvent, the $WD_{refolded}$ adopts a "collapsed" conformation where energetically unfavourable interactions between hydrophobic lipid-binding pocket residues and hydrophilic solvent are largely avoided.

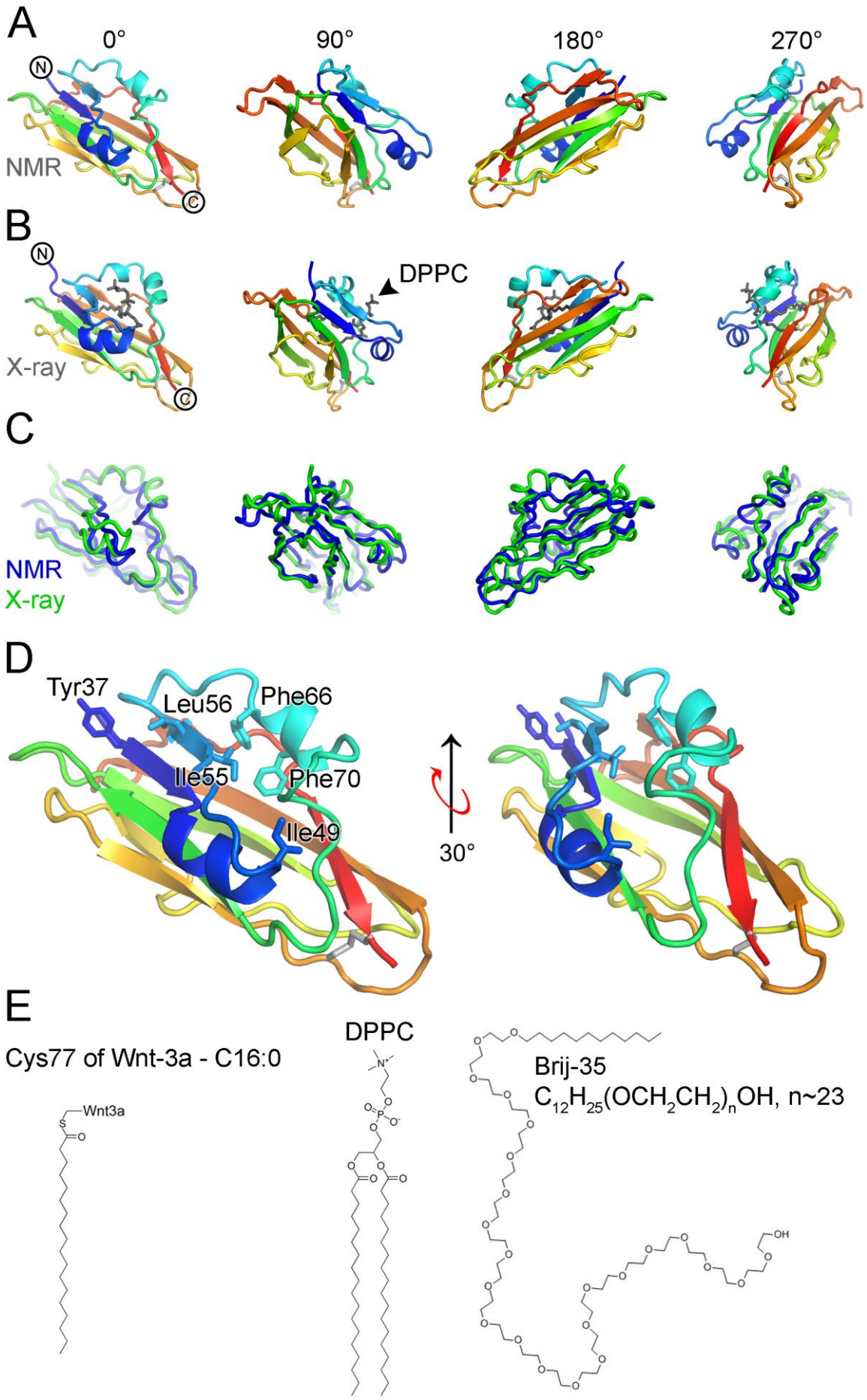

A
0°
90°
180°
270°
N
C
NMR
B
N
DPPC
X-ray
C
C
NMR
X-ray
D
Tyr37
Leu56
Phe66
Ile55
Phe70
Ile49
30°
E
Cys77 of Wnt-3a - C16:0
Wnt3a
DPPC
Brij-35
$C_{12}H_{25}(OCH_2CH_2)_nOH$, n~23
OH

Figure 3.11. Comparison of the human WIF-1 WD structures. Comparison of the WD structures solved by X-ray crystallography (**A**) and NMR spectroscopy (**B**) (Liepinsh et al., 2006; PDB ID 2D3J, model 1). The orientation of the WD is the same as in **Figure 3.7**, **C**. Ribbon diagrams of WDs in four views that differ by a 90° rotation about a vertical axis are presented. WD is colour coded as a rainbow with N- and C-termini shown in blue and red, respectively. The WD-bound DPPC (grey), N- and C-termini are indicated. **C**, superposition of **A** (green) and **B** (blue). **D**, residues of the $WD_{refolded}$ that interact with the CH_2 (Ile55, Phe66, Phe70) or ethylene glycol part (Tyr37, Ile49, Leu56) of the detergent Brij-35 used during protein refolding (Liepinsh et al., 2006). **E**, The skeletal formulae left to right: palmitate moiety covalently linked to the Cys77 of mouse Wnt-3a, DPPC (1,2-dipalmitoyl-glycero-3-phosphocholine) and Brij-35 ($C_{12}H_{25}(OCH_2CH_2)_nOH$, n~23; figure adapted from Santa Cruz Biotechnology, Inc. website).

Interestingly, $WD_{refolded}$ provided evidence of an association between the WD and the alkyl chain-containing detergent Brij-35 ($C_{12}H_{25}(OCH_2CH_2)_nOH$, n~23) used in the refolding protocol (Liepinsh et al., 2006). Intermolecular nuclear Overhauser effects were observed between (1) the CH_2 groups of Brij-35 and Ile55, Phe66, Phe70 of $WD_{refolded}$, and (2) OCH_2CH_2 groups of Brij-35 and Trp29, Tyr37, Ile 49 and Leu56 (**Figure 3.11**, **D**–**E**). In agreement with the $WD_{refolded}$ structure, interactions between the CH_2 group of the lipid head group and Phe66, and Ile49–(glycerol backbone of the lipid) were observed in the crystal (**Figure 3.9**). However, in contrast to the WD–lipid crystal structure, no deep penetration of the single Brij-35 alkyl chain into the WD lipid-binding pocket was observed (Liepinsh et al., 2006). Thus, it is tempting to speculate that (1) interactions between the hydrophobic single chain of the detergent and the lipid-binding pocket residues are less energetically favourable than interactions between lipid-binding pocket residues in the "collapsed" $WD_{refolded}$ conformation or (2) that the hydrophobic detergent alkyl chain ($C_{12}H_{25}$) is too short to fill the whole lipid-binding pocket and thus is similarly energetically unfavourable. In agreement with the second speculation (2), no diacyl phospholipids with twelve carbon atom-acyl chain length (or shorter) were identified as WIF-1 ligands in our mass spectrometry analysis (see **Section 3.5.1** and **Figure 3.4**); the shortest acyl chain present in the WIF-1-bound ligands contains fourteen carbon atoms (**Figure 3.4**).

The second notable difference between the $WD_{refolded}$ and the WD crystal structures is presence of the Asn88-linked glycosylation in the WD crystal structures. The *E.coli* expression system used to obtain the $WD_{refolded}$ structure lacks a series of enzymes required for Asn-linked glycosylation. Eukaryote protein folding and Asn-linked glycosylation are tightly linked processes (reviewed by Roth et al., 2010; Trombetta and Parodi, 2003). Asn-linked glycosylation increases solubility of the target proteins and lack of it may explain the tendency of $WD_{refolded}$ to precipitate (Liepinsh et al., 2006). To conclude, it remains an open question how the absence of the Asn88-linked glycosylation may have affected folding and lipid-binding properties of the $WD_{refolded}$.

The third difference between the $WD_{refolded}$ and WD crystal structures is probably a consequence of refolding, absence of the Asn-linked glycosylation and the bound lipid: the $WD_{refolded}$ lacks several secondary structure elements observed in the WD crystal structure and some of the amino acid residues are in the energetically unfavourable conformations. For example, the 3_{10} (residues Ala64–His68) and α2 helices (residues Phe70–Met77; **Figures 3.7**, **B**–**C** and **3.8**, **A**), observed in the crystal structures, are present as a single helix in the $WD_{refolded}$ (residues Thr67–Arg71; Liepinsh et al., 2006, and Figure 1 therein). We have compared the dihedral angles ψ and φ of amino acid residues (Ramachandran et al., 1963) in the WD crystal and $WD_{refolded}$ structures with MolProbity (Davis et al., 2007). The WD crystal structures contain no Ramachandran outliers; in contrast, the $WD_{refolded}$ structure contains 3 to 6 outliers per NMR model (all 20 $WD_{refolded}$ NMR models were analysed with MolProbity, PDB ID 2D3J). In addition, 98 % of the WD crystal structure residues belong to the most favoured region of Ramachandran plot (**Table 3.2**) (in contrast to 82.6 % of the $WD_{refolded}$, Liepinsh et al., 2006, and Table 1 therein), suggesting that the quality of the WD model obtained by X-ray crystallography is higher compared to the WD models obtained by NMR spectroscopy (Kleywegt and Jones, 1996).

3.66 Proteins structurally similar to the WIF domain crystal structure

Proteins with high three-dimensional structural similarity often have similar function. Thus, we have compared the WD crystal structure with all experimentally determined macromolecular structures deposited in the Protein Data Bank (69266 entries as of April 31, 2011) in order to identify structurally similar proteins that could potentially provide further clues to WIF-1 function. The secondary structure matching (SSM) algorithm of protein three-dimensional structure comparison was used (Krissinel and Henrick, 2004). The SSM algorithm includes matching graphs built on the proteins' secondary structure elements, followed by an iterative three dimensional alignment of protein backbone Cα atoms (Krissinel and Henrick, 2004). The quality of the alignment of two structures was assessed taking into account both the number of aligned residues and root mean square deviation (RMSD) of aligned Cα atoms. A high number of aligned residues and low RMSD values indicate structural similarity between the two aligned structures. Structures similar to the WD crystal structure are listed in **Appendix 5**. None of the identified proteins (except $WD_{refolded}$, **Figure 3.12**, **A**) are known as Wnt-binding molecules. Most of the structurally similar proteins, such as Filamin A and Ca^{2+}-binding domain from the Na^{+}/Ca^{2+} exchanger (**Figure 3.12**, **B**–**C**), contain the immunoglobulin-like fold consisting of anti-parallel β sheets similar to the WD. However, the functional roles of Filamin A or Ca^{2+} binding domain from the Na^{+}/Ca^{2+} exchanger and WD seem to be different. Interestingly, another protein structurally matching the WD, the guanine nucleotide dissociation inhibitor RhoGDI, is also composed of two anti-parallel β sheets (**Figure 3.12**, **D**), but, in addition, contains a similar short 3_{10} helix on the edge of the β sandwich and binds a hydrophobic ligand in the internal pocket. These three WD-resembling domains are briefly described below.

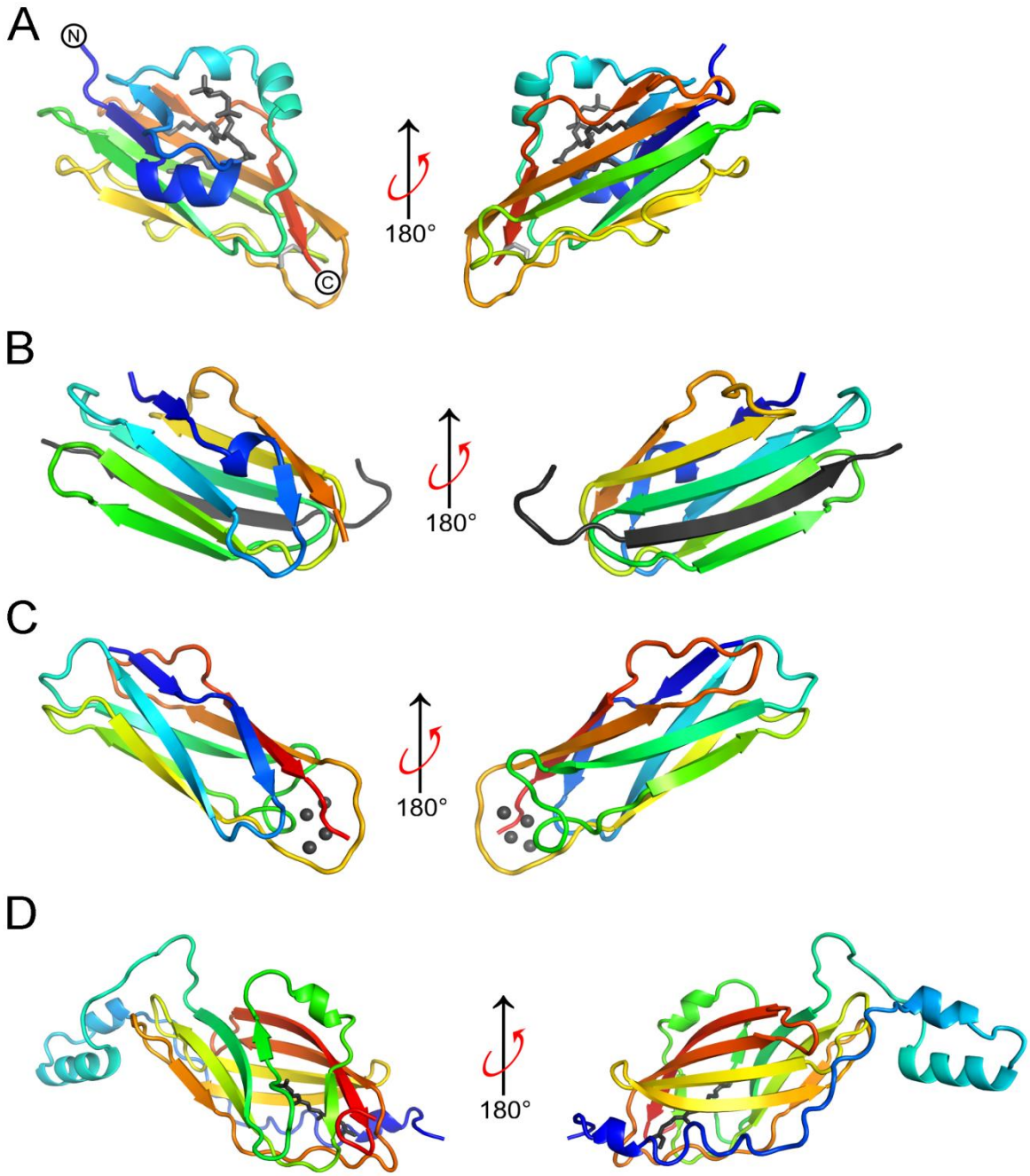


Figure 3.12. Comparison of the human WIF-1 WD crystal structure (**A**) with the Filamin A (FLNa) domain 17 and Ib α peptide complex (PDB ID 2BP3, chain B) (**B**), Ca^{2+}-binding domain 1 (CBD1) from the CALX Na^{+}/Ca^{2+} exchanger (PDB ID 3EAD, chain D) (**C**), and RhoGDI and geranyl-geranyl complex (PDB ID 1DOA, chain B) (**D**). Domains are colour coded as a rainbow with N- and C-termini shown in blue and red, respectively. DPPC and a disulfide bond in **A**, peptide in **B**, four Ca^{2+} ions in **C**, and a geranyl-geranyl moiety in **D** are shown in grey. Two views of each structure that differ by a 180° rotation about a vertical axis are presented.

3.66.1 Crystal structure of the Filamin A (FLNa) domain 17 and platelet adhesion glycoprotein Ib α complex

Filamin A is a dimeric protein that anchors the platelet adhesion glycoprotein (GP) Ib-IX-V receptor to the actin cytoskeleton (Nakamura et al., 2006). The structure of the free FLNa domain 17 was solved by NMR spectroscopy (Nakamura et al., 2006). The structure consists of a β sandwich of relatively straight 3- and 4-strand β sheets, and resembles FLN-type immunoglobulin-like domains. The RMSD of 3.0 Å for 79 aligned residues between FLNa domain 17 (PDB ID 2BP3, chain B) and the WD crystal structure suggests only limited structural similarity between the two structures (**Figure 3.12**, **B**). In contrast to the WD crystal structure, FLNa domain 17 forms a tightly packed β sandwich which lacks an internal cavity suitable for the lipid binding. FLNa domain 17 has only one short α helix (in contrast to two α and one 3_{10} helix of the WD crystal structure) on the edge of β sandwich. The FLNa domain 17 structure in complex with the glycoprotein Ib α peptide was determined by X-ray crystallography (Nakamura et al., 2006). It reveals that the Ib α peptide binds in a groove formed between β strands of the FLNa domain, on the β sandwich edge lacking helices. Whether the WD binds any of the proteins in a similar way between β strands remains an open question. However, we have assessed the potential of the corresponding WD region to serve as a Wnt3a-binding site by site-directed mutagenesis coupled to the Wnt/β-catenin cellular assays (see **Section 4.8**).

3.66.2 Crystal structure of the Ca^{2+} binding domain 1 (CBD1) from the CALX Na^{+}/Ca^{2+} exchanger

The transmembrane Na^{+}/Ca^{2+} exchanger is a Ca^{2+} export protein that facilitates the re-establishment of cytosolic Ca^{2+} levels. Ca^{2+} interactions at its Ca^{2+} binding domains are essential for the allosteric regulation of Na^{+}/Ca^{2+} exchange activity. The Ca^{2+}-bound CALX-

CBD1 crystal was determined at 2.25 Å resolution (Wu et al., 2010). The CALX-CBD1 structure has an immunoglobulin-like fold formed by two anti-parallel β sheets. The RMSD of 3.8 Å for 99 aligned residues between CALX-CBD1 (PDB ID 3EAD, chain D) and the WD crystal structure suggests relatively limited structural similarity between the two structures (**Figure 3.12**, **C**). In contrast to the WD, CALX-CBD1 has no α helices and binds to four Ca^{2+} ions proximal to its C-terminus. This renders drawing functional parallels between the WD and CBD1 difficult.

3.66.3 Crystal structure of the RhoGDI and geranyl-geranyl complex

RhoGDI is a guanine nucleotide dissociation inhibitor (GDI) which binds to the Rho family guanosine-5'-triphosphate (GTP)-binding protein Cdc42. RhoGDI regulates Cdc42 activity by preventing GTP dissociation and modulating Cdc42-membrane interactions. The RMSD of 3.7–3.9 Å for 105–109 aligned residues between RhoGDI (**Appendix 5**) and the WD crystal structure suggests only limited structural similarity between the two structures (**Figure 3.12**, **D**). However, a closer inspection of the two structures, and, in particular, similar hydrophobic ligands, is intriguing.

The structure of RhoGDI consists of nine β strands in two antiparallel sheets, forming a β sandwich, and a short 3_{10} helix (Keep et al., 1997; **Figure 3.12**, **D**). The two-turn 3_{10} helix is located at the edge of the β sheet (Keep et al., 1997). Similarly, the WD crystal structure (but not $WD_{refolded}$) revealed a one-turn 3_{10} helix at the edge of the β sheet, proximal to the surface-exposed lipid head group. Although amino acid sequence similarity between the 3_{10} helices is limited (WD, residues 64–68, AlaProPheThrHis; RhoGDI (PDB ID 1DOA, chain B, residues LeuGluSerPheLysLysGln), both helices have a single phenylalanine with its side chain orientated towards the interior of the β sandwich. In the case of WD, the aromatic ring of Phe66 packs against the CH_2 atoms of the lipid-head group (**Figure 3.9**).

Moreover, the RhoGDI–Cdc42 complex crystal structure shows that RhoGDI buries a

hydrophobic geranyl-geranyl chain from Cdc42 in its interior (Hoffman et al., 2000; **Figure 3.12**, **D**). The transfer of the geranyl-geranyl moiety of Cdc42 from the lipid bilayer to a binding pocket within GDI results in the dissociation of the Cdc42–GDI complex from the membrane (Hoffman et al., 2000). Similarly to the WD, the geranyl-geranyl moiety at the C-terminus of Cdc42 is inserted into a hydrophobic pocket between the two β sheets. However, the entrance to the hydrophobic pockets of RhoGDI and WD is located at different relative positions of the β sandwich; in the case of WD, it is located on the β sandwich edge proximal to the 3_{10} helix; in the case of RhoGDI, it is located on the opposite β sandwich edge lacking the 3_{10} helix.

The hydrophobic ligand binding capabilities of RhoGDI and WD again highlight several questions, for example, whether (1) the WD binds to the lipid moieties present in certain Wnt proteins, and (2) can the single-chain lipid linked to Wnts displace the double-chain phosphatidyl choline buried inside the WD.

3.67 Structure of the WIF-1 EGF-like domain I

EGF-like domain I is linked to the WD via a short linker of approximately four amino acid residues (between Cys177 of the WD and Cys182 of the EGF-I) (**Figure 3.7**). The EGF-like domain I has the canonical pattern of disulfide bonding (Cys182–Cys192, Cys186–Cys198, Cys200–Cys209). These three disulfide bonds stabilise the EGF-like domain architecture which lacks extensive secondary structure. EGF-like domain I is most structurally similar to the EGF-like domain from human clotting factor IX (Rao et al., 1995; PDB ID 1EDM, chain B, RMSD of 2.0 Å over 30 aligned Cα atoms as calculated with DaliLite program (Holm and Rosenström, 2010)), despite the absence of Ca^{2+} binding residues in the EGF-like domain I. This domain extends away from the WD into solvent and is packed against the EGF domain from the neighbouring molecule in the crystal (**Figure 3.6**, **B**). Despite this packing, the EGF I structure appears relatively disordered compared to the WD (as judged from

crystallographic B-factors and quality of electron density), probably as a result of orientational flexibility relative to the WD (**Figure 3.13**).

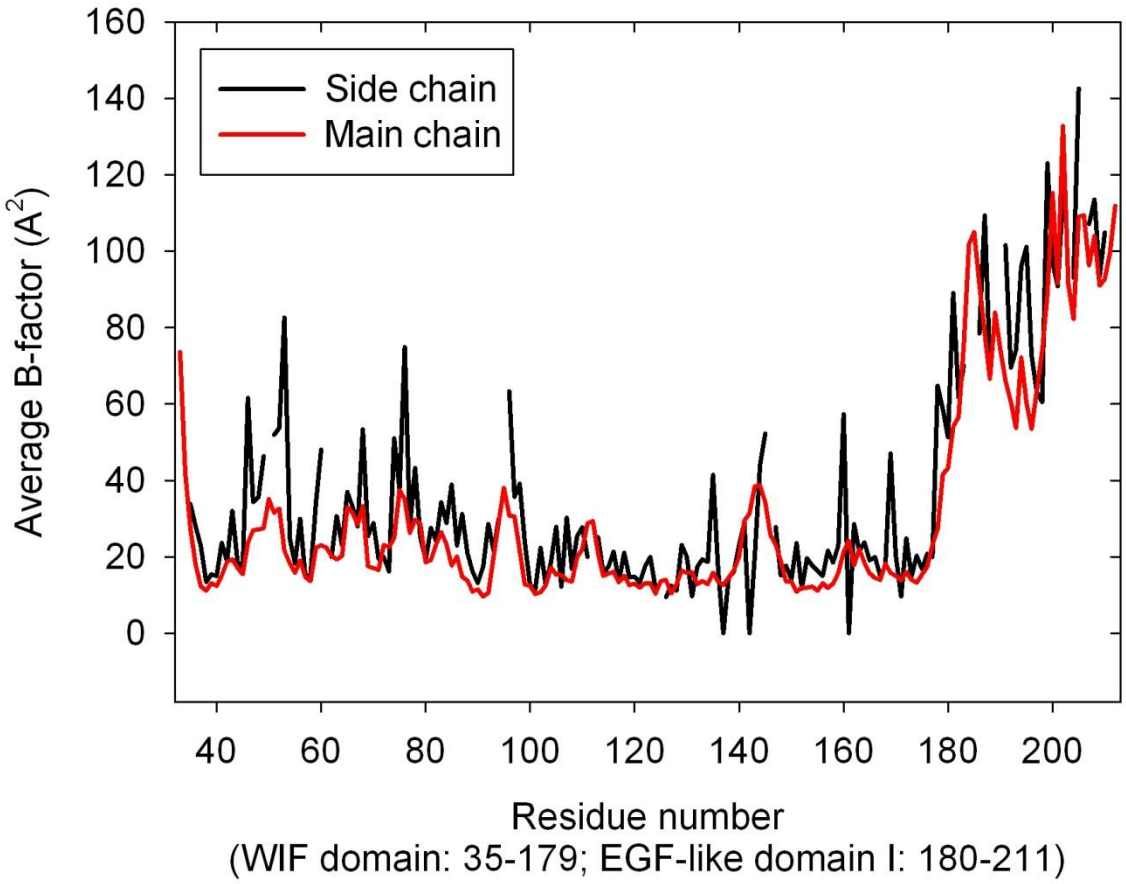


Figure 3.13. Side chain (black) and main chain (red) per-residue crystallographic B-factor (also called atomic displacement parameter; see Formula 6 in **Appendix 6**) for the final WIF-$1_{WD\text{-}EGF\text{-}I}$ model. Note that B-factors of the EGF-like domain I are higher compared to the WD. Gaps in the curve (black) corresponding to the side chains are due to the presence of glycine residues (lacking side chains).

3.7 Homology models of WIF domains from *Drosophila* WIF-1 (Shifted) and human RYK

The presence of the DPPC in the WD of human WIF-1 raises a question whether WDs present in other proteins are likely to bind 1,2-diacyl-phosphatidylcholines, and do they have suitable residues to form the internal lipid-binding pocket. The sequence alignment of WDs (**Figure 3.8**, **A**) indicates high evolutionary conservation and suggests the possibility of the lipid-binding pocket conservation too. We have extended this sequence conservation analysis by building homology models (using the WD (Ser35–Gln179) from the WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structure as a template) of human RYK (Ser60–Lys193, **Figure 3.8**, **A**) and *Drosophila melanogaster* WIF-1 (also called Shifted, a Hedgehog, rather than Wnt, binding protein, residues Gly116–His263, **Figure 3.8**, **A**) WDs. A homology modelling program Modeller 9.9 (Marti-Renom et al., 2000) with a graphical user interface EasyModeller 2.0 was used (Kuntal et al., 2010; sequence alignment as presented in **Figure 3.8**, **A** was used). High sequence similarity between the human WIF-1, Shifted and RYK WDs is reflected in the modelled structures too (**Figure 3.14**, **B**–**C**). All three WDs retain the lipid-binding pocket of similar size as calculated with CASTp (probe radius 1.4 Å, Dundas et al., 2006; http://sts.bioengr.uic.edu/castp/), suggesting that all proteins may have lipid-binding properties:

1. WIF-1: lipid-binding pocket volume 1556 $Å^3$, pocket surface area 1044 $Å^2$;
2. RYK: 1625 $Å^3$, 1342 $Å^2$;
3. Shifted: 1821 $Å^3$, 1272 $Å^2$;

However, a closer inspection of the human RYK WD model shows the side chain of the Arg187 pointing to the interior of the putative lipid-binding pocket. All vertebrate WIF-1 homologues (Phe173 in human WIF-1, **Figure 3.8**, **A**) and *Drosophila* Shifted have

phenylalanines at the corresponding position. In the WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structure, the distance between the closest DPPC and Phe173 carbon atoms is 3.7 Å and the arginine side chain at the corresponding position would sterically clash with the hydrophobic acyl chains of the lipid. This arginine residue is highly conserved among RYK homologues (**Figure 3.8**, **A**) suggesting that the WDs of RYK receptors are not lipid-binding modules or have different architecture that accommodates the corresponding arginine in a lipid-binding compatible way. Furthermore, assuming that, first, RYK WDs are not lipid-binding modules, and, second, they are Wnt-binding domains, this renders the possibility of the Wnt-linked lipid moiety insertion into the WD unlikely.

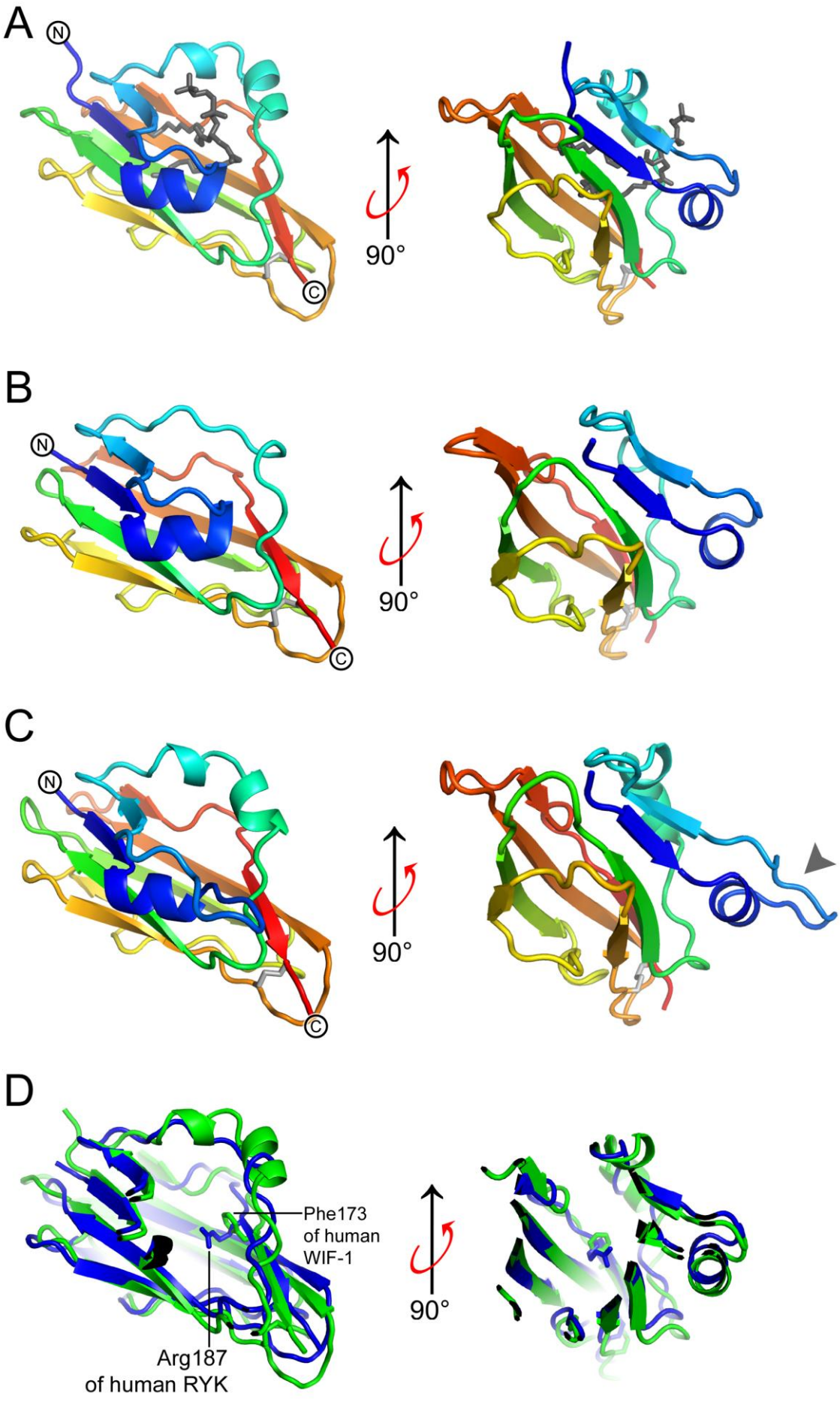

A
N
C
90°
B
N
C
90°
C
N
C
90°
D
Phe173
of human
WIF-1
Arg187
of human RYK
90°

Figure 3.14. Based on the human WD crystal structure (**A**), WD homology models of human Wnt receptor RYK (**B**) and *Drosophila melanogaster* WIF-1 (also called Shifted, a Hedgehog, but not Wnt, binding protein) (**C**) were built. The orientation of the human WIF-1 WD is the same as in **Figure 3.7**, **C**. Ribbon diagrams of WD in two views that differ by a 90° rotation about a vertical axis are presented. WD is colour coded as a rainbow with N- and C-termini shown in blue and red (**A**–**C**), respectively. The WD-bound DPPC is shown as grey sticks in **A**, N- and C-termini are indicated with capital letters. A hydrophobic amino acid stretch (AlaLeuTyrPheProGln) of Shifted WD is indicated with a grey arrow in **C** (90° view). **D**, superposition of the human WIF-1 (green) and RYK (blue) WDs. Two residues facing the interior of the lipid-binding pocket (Phe173 of WIF-1 and Arg187 of RYK) are shown as sticks. DPPC is omited for clarity.

3.8 Future perspectives

The WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structures presented in **Chapter 3** reveal a novel 1,2-diacyl-phosphatidylcholine binding pocket inside the WIF domain (WD) of the human WIF-1. We show that the WD-bound lipid can be exchanged with the fluorescent analogue, suggesting the dynamic nature of WD-lipid interactions. These observations raise a question: what is the role of the WD-bound lipid in WIF-1–Wnt interactions? **Chapter 4** will investigate WIF-1–Wnt3a interactions through a series of cellular and surface plasmon resonance binding assays, coupled to site-directed mutagenesis and further structural analysis of WIF-1. The flexible linkage between the WD and EGF-like domain I, observed in the WIF-$1_{WD\text{-}EGF\text{-}I}$ structure, highlights the question what is the function of the five EGF-like domains of WIF-1? This question will be addressed in **Chapters 5** and **6**, using a series of biophysical and cellular assays, coupled to further crystallographic studies.

Chapter 4 – Cellular and binding studies of WIF-1–Wnt-3a interactions

4.1 Summary

Interactions between Wnt inhibitory factor 1 (WIF-1) and Wnt proteins have been shown to be important for the control of Wnt signalling and human cancer cell growth. The medically important molecular mechanisms of Wnt signalling inhibition by WIF-1 remain poorly understood because it is difficult to produce pure Wnt and six-domain WIF-1 proteins in sufficient quantities for biochemical and biophysical studies. Both glycoproteins are problematic to produce using conventional, bacterial (*E.coli*-based) protein expression systems that lack eukaryotic protein folding and glycosylation machinery. In **Chapter 4**, I set out to dissect the mechanisms by which WIF-1 interacts with Wnt-3a using a series of human WIF-1 constructs and mouse Wnt-3a expressed in mammalian cells. In this chapter I report biophysical and cellular assays, combined with the site-directed mutagenesis, which suggest that the Wnt-3a-binding surface is located on the N-terminal WIF domain (WD) of WIF-1, proximal to the lipid head group. The results also implicate the five C-terminal epidermal growth factor (EGF)-like domains of WIF-1 in Wnt-3a binding and signalling inhibition. In addition, comparison of the human WIF-1 WD residues involved in Wnt-3a binding with the corresponding residues of *Drosophila* WIF-1 (also known as Shifted) provides possible clues why Shifted does not bind to Wnt proteins.

4.2 Introduction

WIF-1 is present across vertebrates. Deletion of the mouse WIF-1 gene accelerates development of radiation-induced osteosarcomas *in vivo* (Kansara et al., 2009). Downregulation of WIF-1 expression by promoter methylation has been reported in human

cancers (Mazieres et al., 2004), while overexpression of WIF-1 inhibits lung and bladder cancer cell growth (Kim J. et al., 2007; Tang et al., 2009) (also see **Section 1.5.1.2.1**). Although association between downregulation of WIF-1 expression (as well as upregulation of Wnt signalling; Logan and Nusse, 2004) and cancer development is well established, the molecular mechanisms of WIF-1–Wnt interactions remain poorly defined. No studies on these interactions have been performed using purified WIF-1 and Wnt proteins.

Studies of WIF-1–Wnt interactions are hindered by the lack of suitable, pure protein samples. Wnt proteins are notoriously difficult to express and purify largely because of multiple posttranslational modifications. For example, in addition to Asn-linked glycosylation (Komekado et al., 2007), mouse Wnt-3a has been shown to be palmitoylated (Willert et al., 2003), palmitoleoylated (Takada et al., 2006) and glycophosphatidylinositol (GPI)-anchored to the endoplasmic reticulum (Zoltewicz et al., 2009). These three lipid modifications increase hydrophobicity and complicate purification of Wnt-3a (Willert et al., 2003). In contrast to Wnt-3a, the glycoprotein WIF-1 lacks lipid modification. The recent development of mammalian protein expression systems allows production of secreted glycosylated proteins, such as human WIF-1, in sufficient quantities for biophysical studies, typically, 1–10 milligrams of purified protein per litre of conditioned media (Aricescu et al., 2006). I took advantage of the previously published Wnt-3a purification methods (Willert et al., 2003) and recent developments of mammalian protein expression system (Aricescu et al., 2006) to express and purify mouse Wnt-3a and a series of human WIF-1 constructs, respectively. This allowed me to start to dissect the medically important molecular mechanisms by which WIF-1 inhibits Wnt-3a signalling.

WIF-1 consists of an N-terminal secretion signal sequence, the WIF domain (WD, 143 amino acid residues), five EGF-like domains (31–33 amino acid residues each) and a hydrophilic C-terminus (**Appendix 2**). Human WIF-1 has been shown to bind via the WD to eight different

Wnts (3a, 4, 5a, 7a, 9a, 11 (Surmann-Schmitt et al., 2009), Wingless, *Xenopus* Wnt-8 (Hsieh et al., 1999)) and a protein involved in neuronal differentiation, Olfactomedin 1 (Nakaya et al., 2008)). The role of the WIF-1 EGF-like domains is unknown although they were shown to be required for the full protein activity *in vivo* (Hsieh et al., 1999).

Knowledge of WIF-1–Wnt interaction at atomic resolution is required to understand which of the 19 Wnts encoded in the human genome are inhibited by WIF-1. Selective inhibition of specific Wnt proteins involved in cancerogenesis, while preserving signalling by other Wnt family members (e.g. that are essential for stem cell self-renewal (Reya et al., 2003; Zeng and Nusse, 2010) and tissue regeneration (Kim et al., 2007)), could contribute to disease treatment and is important for Wnt pathway-targeted drug development (Garber, 2009).

Atomic details of WIF-1–Wnt interactions are also required to understand why *Drosophila* WIF-1 (also known as Shifted) does not interact with Wnt proteins but is required for the stabilisation and diffusion of lipid-modified Hedgehog morphogen proteins (Glise et al., 2005; Gorfinkiel et al., 2005). Both human WIF-1 and *Drosophila* Shifted share the same domain architecture, comprised of the N-terminal WD and five C-terminal EGF-like domains, and exhibit substantial amino acid sequence similarity (26 % sequence identity).

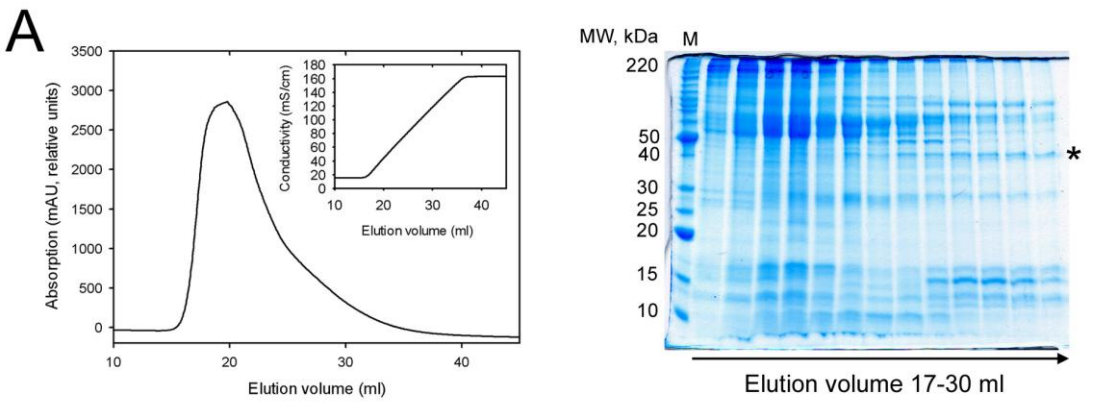
4.3 Aims

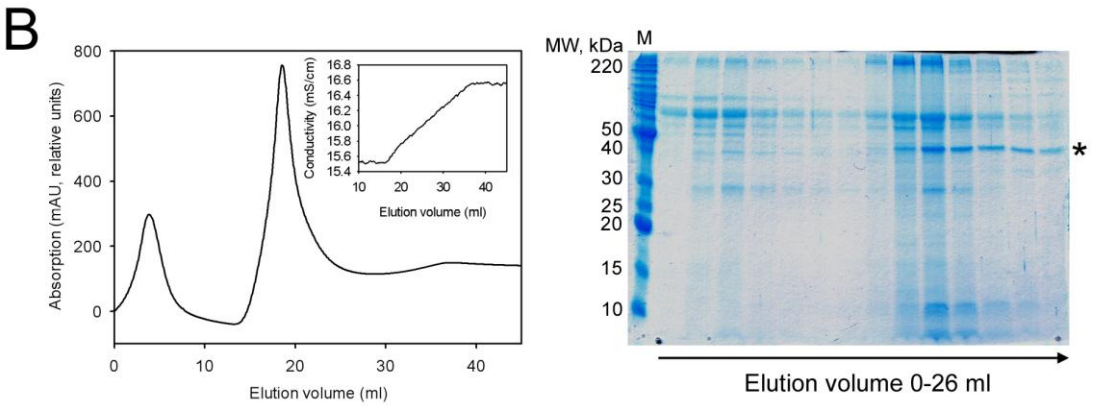
The main two aims of the work presented in **Chapter 4** are, first, to determine the contribution of separate WIF-1 domains to Wnt-3a binding and signalling inhibition; second, to determine the Wnt-binding surface on the WD using a range of biophysical and cellular assays, combined with the site-directed mutagenesis, selective chemical modification and phylogenetic analysis of WIF-1 family members. The two inter-related aims are directed towards understanding, first, how WDs present in WIF-1 and Wnt *r*eceptors related to t*y*rosine *k*inase (RYK proteins) recognise Wnt proteins and, second, how *Drosophila* WIF-1

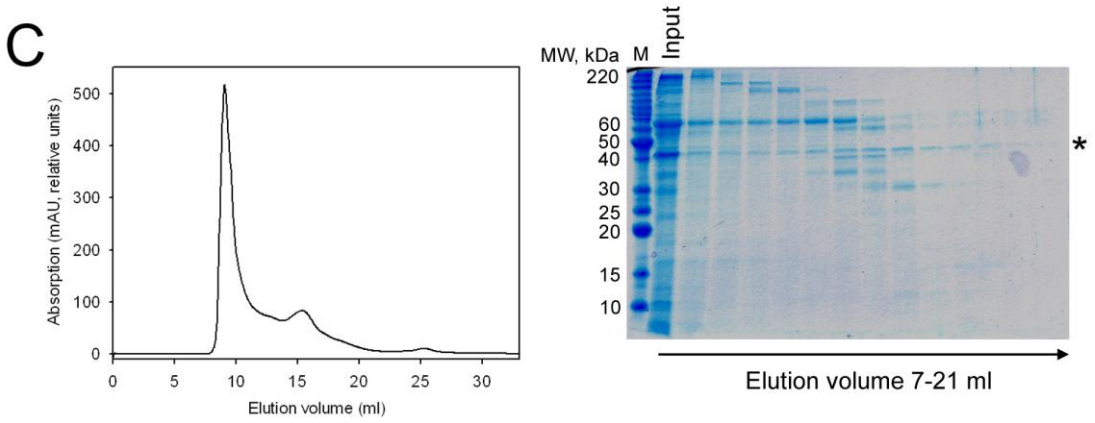
(Shifted) avoids binding to Wnts.

4.4 Expression and purification of mouse Wnt-3a

WIF-1–Wnt interactions have never been studied using the two purified proteins. Typically, studies were carried out using conditioned medium (CM) from the cells overexpressing WIF-1 and Wnts as a source of protein (Hsieh et al., 1999; Surmann-Schmitt et al., 2009, respectively). Thus these studies cannot rule out the possibility that other, unknown molecules (that are present in the CM) may participate and affect WIF-1–Wnt interactions. To rule out this possibility, we set out to study WIF-1–Wnt-3a interactions using purified proteins. Mouse Wnt-3a was expressed by a mouse fibroblast cell line and purified as described previously (Willert et al., 2003) with minor modifications (**Section 2.2.3**). Briefly, CM containing Wnt-3a was dialysed and underwent three-stage fractionation (blue sepharose, IMAC and SEC as summarised in **Figure 4.1**). As a result, a sample which contains ~70 % pure mouse Wnt-3a was obtained and used for the subsequent WIF-1–Wnt-3a binding studies using surface plasmon resonance spectroscopy (**Section 4.7**).


A
Absorption (mAU, relative units)
Conductivity (mS/cm)
Elution volume (ml)
MW, kDa
M
220
50
40
30
25
20
15
10
*
Elution volume 17-30 ml


B
Absorption (mAU, relative units)
Conductivity (mS/cm)
Elution volume (ml)
MW, kDa
M
220
50
40
30
25
20
15
10
*
Elution volume 0-26 ml


C
Absorption (mAU, relative units)
Elution volume (ml)
MW, kDa
M
Input
220
60
50
40
30
25
20
15
10
*
Elution volume 7-21 ml

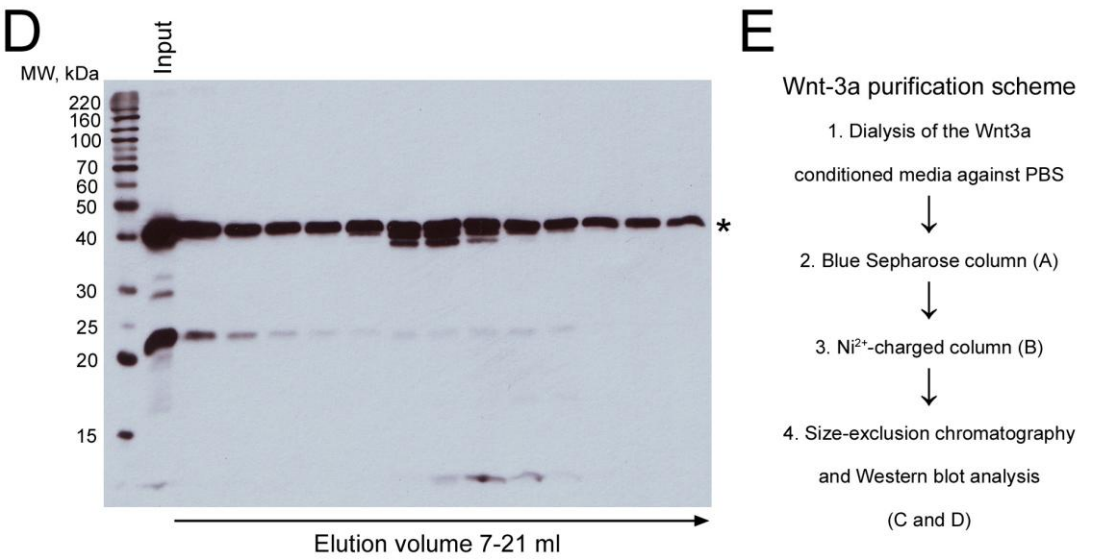

D
MW, kDa
Input
220
160
100
70
60
50
40
30
25
20
15
*
Elution volume 7-21 ml
E
Wnt-3a purification scheme
1. Dialysis of the Wnt3a
conditioned media against PBS
2. Blue Sepharose column (A)
3. Ni2+-charged column (B)
4. Size-exclusion chromatography
and Western blot analysis
(C and D)

Figure 4.1. Purification of mouse Wnt-3a. (**A**) Blue Sepharose fractionation of Wnt-3a conditioned medium (CM). Before fractionation, 1 litre of the CM was dialysed against 15 litres of PBS buffer at 4 °C twice using dialysis membrane which allowed molecules with a molecular weight lower than 12 kDa to diffuse into the PBS buffer out of the CM (see **Section 2.2.3** for further details). The dialysed CM was loaded onto a 5 ml HiTrap Blue Sepharose column at a flow speed of 5–10 ml/min at room temperature. Wnt-3a binds to Cibacron Blue F3G-A molecules covalently attached to Sepharose CL-6B by the triazine coupling method (GE Healthcare). The column with the bound Wnt-3a was washed with 100 ml of PBS, 1 % w/v CHAPS buffer. Wnt-3a was eluted from the column with a salt gradient: starting with PBS, 1 % w/v CHAPS up to 1.5 M KCl, 20 mM Tris-HCl pH 8.0. The increase of the elution buffer ionic strength was monitored by the increase of its conductivity (**A**, left, inset). Fractions containing protein were analysed by reducing SDS-PAGE (**A**, right). (**B**) Fractions containing Wnt-3a (marked with an asterisk in **A**, right) were loaded onto a 5 ml HisTrap HP column charged with Ni^{2+} at a flow speed of 2 ml/min. It has been previously shown that Wnt-3a binds to divalent cations (Cu^{2+}) and this property was exploited in the purification method described here. After Wnt-3a loading, the column was washed with approximately 100 ml of 10 mM imidazole, PBS, 1 % w/v CHAPS buffer. Wnt-3a was eluted from the column with the imidazole gradient: 10 mM to 350 mM imidazole in PBS, 1 % w/v CHAPS (**B**, left). The imidazole gradient was monitored by the increase of buffer conductivity (**B**, inset) and light absorbance by imidazole at 280 nm. Fractions containing Wnt-3a (as assessed by SDS-PAGE analysis, **B**, right, Wnt-3a is marked with an asterisk) were concentrated to 0.4 ml total volume by ultracentrifugation using centrifugal filter units (Millipore). (**C**) The sample was loaded onto a Superdex 200 10/30 gel filtration column (GE Healthcare). The size-exclusion chromatography (SEC) was performed at a flow speed of 0.4 ml/min in PBS, 1 % w/v CHAPS buffer. Protein elution from the column was monitored by absorption at 280 nm (**C**, left). Fractions containing protein were analysed by SDS-PAGE (**C**, right) and Western blotting (**D**). The primary anti-mouse Wnt-3a antibody from rat (R&D Systems) and the secondary anti-rat antibody from goat conjugated to horseradish peroxidase (Invitrogen) were used for Wnt-3a detection during Western blotting (**D**) (see **Section 2.2.3** for further details). Fractions containing Wnt-3a were filtered (using 0.22 μm pore centrifugal filters from Millipore) and used for cellular assays (see **Section 4.8**) or flash-frozen in liquid nitrogen and stored at -80 °C until further use.

4.5 Expression and purification of WIF-1 constructs

4.5.1 Expression and purification of wild-type WIF-1 constructs

Constructs containing the full length human WIF-1 (WIF domain, five EGF-like domains and a hydrophilic C-terminus; WIF-$1_{\text{Full length}}$; residues Gly29–Trp379), WIF-1 lacking the hydrophilic C-terminus (WIF-$1_{\Delta C}$; residues Ser35–Ile346), WIF domain plus EGF-like domains I–III (WIF-$1_{\text{WD-EGFs I–III}}$; residues Ser35–Glu274), EGF-like domains I–V (WIF-$1_{\text{EGFs I–V}}$; residues Gln178–Ile346) (**Figure 4.2**) and zebrafish WIF-1 lacking the hydrophilic C-

terminus (residues Thr33–Asn336; **Appendix 2**) were cloned into the pHLsec expression vector as described in **Section 2.1**. Proteins were expressed in HEK-293S cells and purified as described in **Sections 2.2.2**. Briefly, proteins from the HEK-293S conditioned medium were purified using the standard three stage procedure which includes dialysis, IMAC and SEC. Representative protein elution profiles from SEC and corresponding SDS-PAGE gels are shown in **Figure 4.2**.

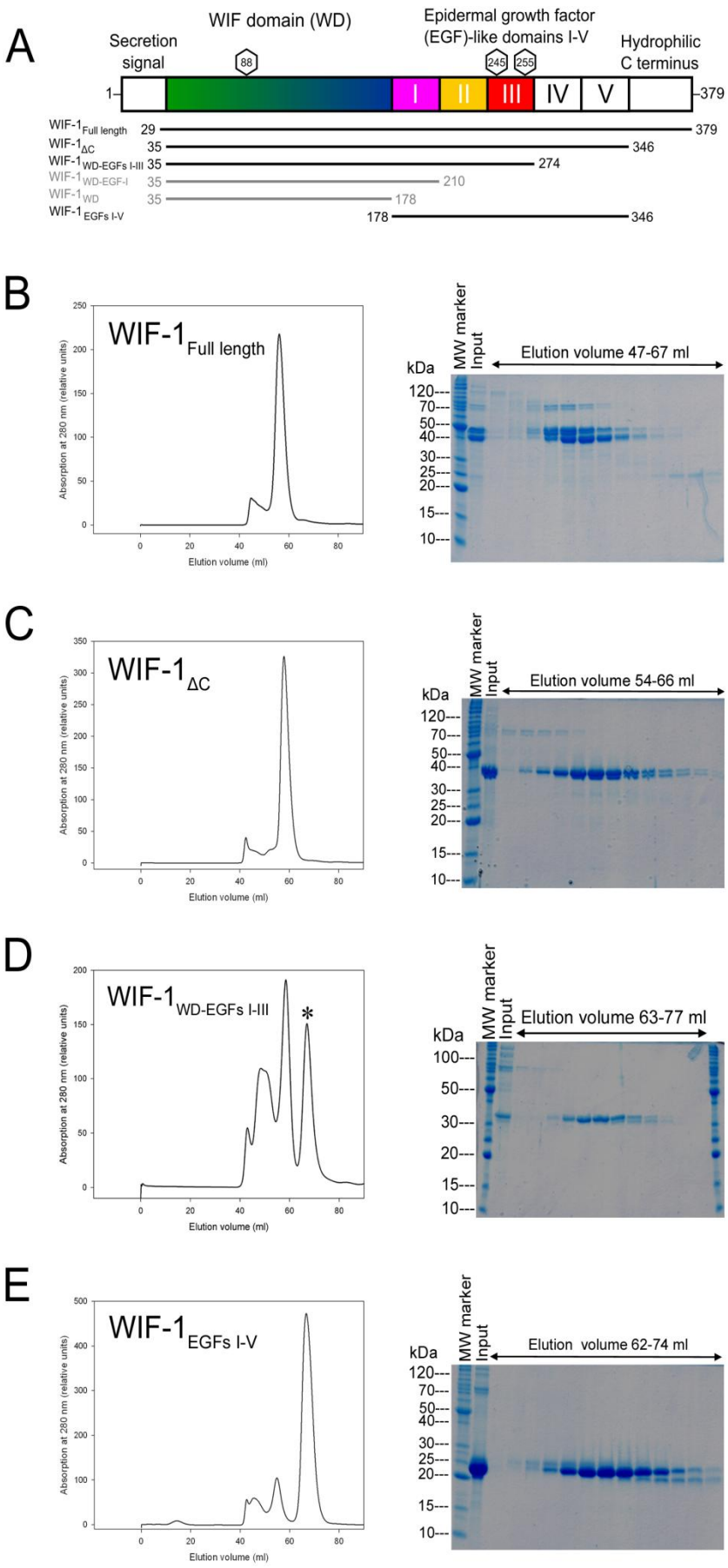

A
WIF domain (WD)
Secretion signal
Epidermal growth factor (EGF)-like domains I-V
Hydrophilic C terminus
88
245
255
1
379
I
II
III
IV
V
WIF-1 Full length 29 379
WIF-1 ΔC 35 346
WIF-1 WD-EGFs I-III 35 274
WIF-1 WD-EGF-I 35 210
WIF-1 WD 35 178
WIF-1 EGFs I-V 178 346
B
WIF-1 Full length
Absorption at 280 nm (relative units)
Elution volume (ml)
MW marker
Input
Elution volume 47-67 ml
kDa
120---
70---
50---
40---
30---
25---
20---
15---
10---
C
WIF-1 ΔC
Elution volume 54-66 ml
D
WIF-1 WD-EGFs I-III
*
Elution volume 63-77 ml
100---
E
WIF-1 EGFs I-V
Elution volume 62-74 ml

Figure 4.2. Domain organization and size-exclusion chromatography (SEC) of the human WIF-1 constructs. **A**, schematic domain organization of the human WIF-1. Asn88-, Asn245-, and Thr255-linked glycosylation sites are marked by hexagons. WD is coloured in green-to-blue transition, EGF-like domains I to III are coloured in magenta, yellow, and red, respectively. Amino acid sequence boundaries of the WIF-1 constructs are shown below; WIF-$1_{WD\text{-}EGF\text{-}I}$ and WIF-1_{WD} constructs, described in **Chapter 3**, are represented in grey. **B**–**E**, SEC of four human WIF-1 constructs: WIF-$1_{Full\ length}$ (**B**), WIF-$1_{\Delta C}$ (**C**), WIF-$1_{WD\text{-}EGFs\ I–III}$ (**D**), and WIF-$1_{EGFs\ I–V}$ (**E**). Separation by SEC was performed using a HiLoad 16/60 Superdex 75 column (GE Healthcare Life Sciences) at a flow speed of 0.8 ml/min in 20 mM Tris-HCl pH 8.0, 300 mM NaCl. Protein elution was monitored by absorbance at 280 nm (black curves, left). WIF-1 peak fractions were analysed on reducing SDS-polyacrylamide (15 % w/v) gels (right). WIF-$1_{WD\text{-}EGFs\ I–III}$ (residues 35–274) expresses at significantly lower levels (**D**). The expressed protein is present in the peak marked with an asterisk (*); three slightly longer WIF-$1_{WD\text{-}EGFs\ I–III}$ constructs (residues 35–275, 35–276, 35–277) were not secreted by HEK-293T cells (data not shown). Samples loaded onto SEC column are presented (designated as Input). Molecular weight (MW) marker: BenchMark Protein Ladder from Invitrogen. Gels were stained with SimplyBlue SafeStain from Invitrogen.

4.5.2 Expression and purification of mutant WIF-1 constructs

Site-directed mutagenesis was performed using overlap extension PCR (Ho et al., 1989). A selection of PCR primers and restriction enzymes are described in **Section 2.1** and listed in **Table 2.1**. PCR products were cloned into the pHLsec expression vector. All constructs were verified by DNA sequencing. Proteins were expressed and purified as described for the wild-type constructs. SDS-PAGE and Western blot analysis of the WIF-$1_{\Delta C}$ constructs with an introduced Asn-linked glycosylation site is presented in **Figure 4.3**.

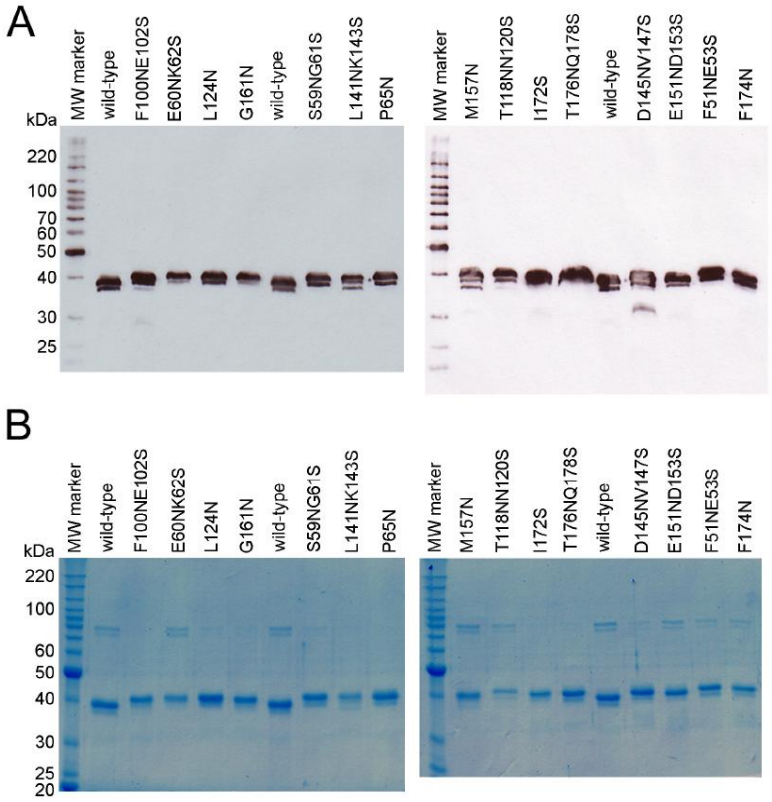


Figure 4.3. Western blot (**A**) and SDS-polyacrylamide gel electrophoretic (**B**) analysis of the purified WIF-$1_{\Delta C}$ constructs. Mutations (indicated above each lane) introduce an Asn-linked glycosylation site and increase molecular weight (MW) of WIF-$1_{\Delta C}$. Reducing SDS-polyacrylamide (12 % w/v) gels were used (**A** and **B**). Western blots were probed with mouse anti-PentaHis monoclonal primary antibody (Qiagen) and goat anti-mouse IgG horseradish peroxidase (HRP) conjugate (Sigma). HRP was detected by chemiluminescence using the Amersham ECL detection kit (GE Healthcare Life Sciences) (**A**). Gels (**B**) were stained with SimplyBlue SafeStain from Invitrogen. MW marker: BenchMark Protein Ladder from Invitrogen.

4.6 Both WIF and EGF-like domains contribute to Wnt-3a signalling inhibition

It has been shown previously that the WD of WIF-1 is sufficient for Wnt signalling inhibition whilst the function of the five EGF-like domains (EGFs I–V) has remained unknown (Hsieh et al., 1999). Thus I used a series of WIF-1 constructs (**Figures 3.2** and **4.2**) in a cellular Wnt/β-catenin response assay to begin to dissect the roles of the various domains of WIF-1 in Wnt-3a inhibition. HEK-293T cells were transfected with a luciferase reporter of β-catenin-mediated transcriptional activation, a canonical Wnt signalling pathway response (DasGupta et al., 2005; **Section 2.11**). The reporter plasmid Super-TOPFlash contains 11 TCF/LEF

transcription factor binding sites upstream of the firefly luciferase gene. Upon cell treatment with Wnt-3a, β-catenin translocates to the nucleus, binds to TCF/LEF proteins and activates transcription of the luciferase gene. Luciferase expression, and the corresponding Wnt-3a signalling level, was quantified by measuring enzymatic activity of the luciferase as described in **Section 2.11**.

Wnt signalling was readily activated by adding purified human Wnt-3a (4.7 nM; from R&D Systems) to the growth medium. The signalling was inhibited in a dose dependent manner (33–500 nM; **Figure 4.4**) by the WIF-$1_{\Delta C}$ construct. WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ were also inhibitory, consistent with the previous report that WD in isolation is able to inhibit Wnt signalling (Hsieh et al., 1999), but both these constructs were significantly less effective than the molecules containing both WD and EGFs I–V. Thus, although EGFs I–V in isolation (construct WIF-$1_{EGFs\ I–V}$) showed no pronounced inhibitory activity (**Figure 4.4**), the results of the cellular assays point to an involvement of the EGF-like domains in Wnt-3a/β-catenin signalling inhibition.

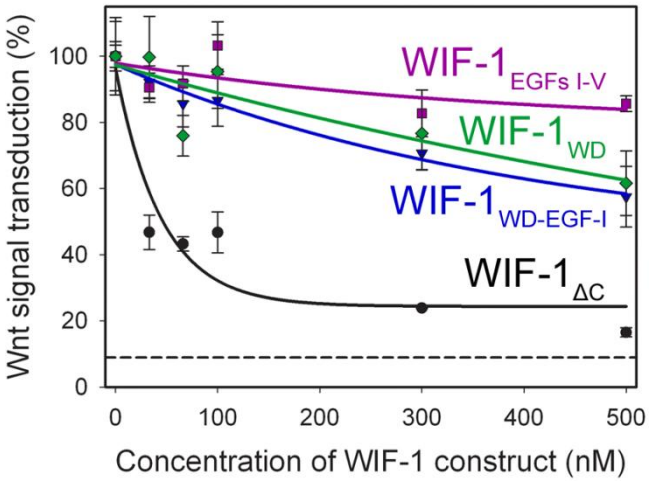


Figure 4.4. The concentration-dependent inhibition of human Wnt-3a signalling in a cellular assay by four constructs of human WIF-1: WIF-$1_{\Delta C}$ (black), WIF-$1_{WD\text{-}EGF\text{-}I}$ (blue), WIF-1_{WD} (green), and WIF-$1_{EGFs\ I–V}$ (magenta). The concentration of Wnt-3a was kept constant at 4.7 nM. The dashed line indicates endogenous Wnt signalling in HEK-293T cells without added Wnt-3a. Experiments were performed in triplicate, the height of the error bar is equal to the standard deviation.

4.7 Both WIF and EGF-like domains bind to Wnt-3a

To analyse the biophysical basis for the observed difference in efficiency of Wnt-3a signalling inhibition by WD and molecules containing WD plus EGFs I–V, I probed the interactions between the different WIF-1 domains and purified mouse Wnt-3a (**Section 4.4**) using a surface plasmon resonance (SPR) binding assay. Mouse Wnt-3a purified in-house was chosen over commercially available human Wnt-3a (R&D Systems) which is relatively expensive (10 µg for £490, protein purity >90%) and might be not suitable for detailed biophysical characterization (it is lyophilized and might be misfolded and/or proteolytically cleaved; Cajanek et al., 2010). The sequence identity between mouse and human Wnt-3a is 96.6 % (i.e. 340 out of 352 amino acid residues at the equivalent positions are identical).

In-house purified mouse Wnt-3a (**Figure 4.1**) was immobilized on the SPR sensor surface via primary amines and binding was measured for analytes comprising a standard series of WIF-1 constructs (**Figure 4.2**). All the constructs gave measureable binding affinities (**Figure 4.5**

and **Table 4.1**). In agreement with the cellular assay, the tightest binding was observed between Wnt-3a and WIF-$1_{\Delta C}$ with a dissociation constant (K_d) of 21.6 ± 8.0 nM. The full length WIF-1 construct has a comparable binding affinity (K_d 62.8 ± 23.3 nM), indicating that the hydrophilic C-terminal region makes no contribution to Wnt-3a binding, consistent with a sequence-based prediction of flexibility (**Appendix 4**, **C**) and evolutionary divergence (**Appendix 2**). These SPR based dissociation constants are in good agreement with a previously reported K_d of 16 nM derived from plate-based affinity assays for the interaction of *Xenopus* Wnt8 and human WIF-1 (Hsieh et al., 1999). WIF-1_{WD} and WIF-$1_{WD\text{-}EGF\text{-}I}$ showed significantly weaker binding to Wnt-3a (K_d 3.4 ± 1.6 and 5.6 ± 2.4 µM, respectively). Notably, I measured significant binding affinity between WIF-$1_{EGFs\ I–V}$ and Wnt-3a (K_d 0.9 ± 0.2 µM), suggesting that the differences I observed in the cellular assays are the result of a direct interaction between EGFs I–V and Wnt-3a, an interaction which acts in addition to that of the WD. To provide further evidence of this additive effect, I tested an additional construct, WIF-$1_{WD\text{-}EGFs\ I–III}$, in the cellular and SPR assays (**Table 4.1**). These additional data again showed reduced Wnt-3a binding and inhibition compared to the results for WIF-$1_{\Delta C}$. This suggests that EGFs IV and V are necessary for the contribution of the EGF I–V unit to Wnt-3a binding, however, I cannot exclude that the reduced Wnt-3a binding is an indirect effect arising from perturbation of EGF III because of the truncation at the EGF III–EGF IV interface (**Figure 4.2**, **D**). In summary, the cellular and SPR binding assays are consistent with a modular model in which both the WD and the EGF-like domains of WIF-1 interact with Wnt-3a and contribute to the signalling inhibition.

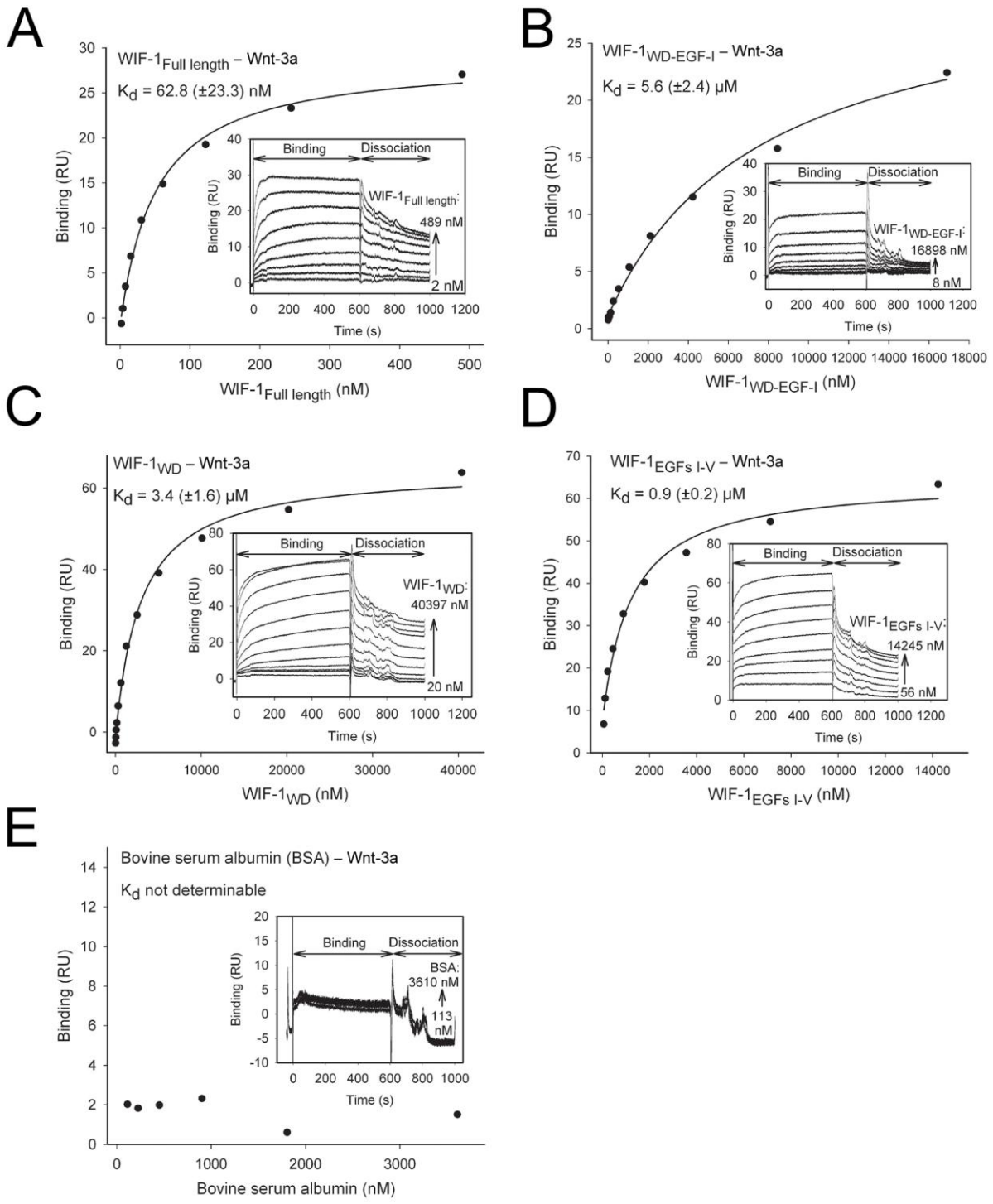


Figure 4.5. Binding of human WIF-1 constructs and BSA (control) to mouse Wnt-3a: WIF-$1_{\text{Full length}}$ (**A**), WIF-$1_{\text{WD-EGF-I}}$ (**B**), WIF-1_{WD} (**C**), WIF-$1_{\text{EGFs I–V}}$ (**D**), and BSA (**E**). Different concentrations of WIF-1 constructs were injected over surface coupled with Wnt-3a. Representative plots of the binding response (response units (RU)) derived from the sensorgrams (insets) as a function of WIF-1 construct (or BSA) concentration. K_d values (± standard deviation) calculated from three independent experiments are shown above the curves.

Table 4.1. Dissociation constants (K_ds) measured by SPR between mouse Wnt-3a and human WIF-1 constructs in PBS, 1 % CHAPS buffer unless stated otherwise (see **Section 2.7.1**).

WIF-1 construct	**Dissociation constant K_d, nM**
WIF-1_{WD}	3400 ± 1600
WIF-$1_{WD\text{-}EGF\text{-}I}$	5600 ± 2400
WIF-$1_{WD\text{-}EGFs\ I–III}$	3210 ± 670
WIF-$1_{\Delta C}$	21.6 ± 8.0
WIF-$1_{EGFs\ I–V}$	900 ± 200
WIF-$1_{Full\ length}$	62.8 ± 23.3
WIF-$1_{Full\ length}$ with 2 mM $CaCl_2$ pH 8.0 (buffer A)	190.0 ± 34.9
WIF-$1_{Full\ length}$ with 2 mM EDTA pH 8.0 (buffer A)	167.6 ± 12.4
WIF-$1_{WD\text{-}EGF\text{-}I}$ (buffer B)	9860 ± 1150
WIF-$1_{WD\text{-}EGF\text{-}I}$ with 2 mM phosphocholine chloride (buffer A)	15900 ± 930
WIF-$1_{WD\text{-}EGF\text{-}I}$ with 2 mM phosphocholine chloride (buffer B)	16900 ± 4400
WIF-$1_{WD\text{-}EGF\text{-}I}$ with 2 mM $CaCl_2$ pH 7.8 (buffer B)	16800 ± 1970
WIF-$1_{\Delta C}$ Met77Trp	217.0 ± 26.0
WIF-$1_{WD\text{-}EGF\text{-}I}$ Met77Trp	2860.0 ± 530.0
Bovine serum albumin	Not determinable

4.8 The Wnt-3a-binding site on the WIF domain is in the proximity of the lipid head group

In order to map out the areas of the WIF domain (WD) surface involved in human Wnt-3a recognition I used the high resolution crystal structures to identify suitable residues for site-directed mutagenesis. Residues with surface exposed side chains were treated as suitable. A series of 15 surface exposed sequences X-Z-Ser, Asn-Z-X and X-Z-X (where X and Z is any amino acid, but Z is not Pro) were mutated to Asn-Z-Ser in order to introduce an Asn-linked glycosylation site to obstruct WIF-1–Wnt-3a recognition (**Figure 4.3**). The introduction of the Asn-linked glycosylation site resulted in an increase of WIF-$1_{\Delta C}$ variant molecular weight compared to the wild-type WIF-$1_{\Delta C}$ as judged by SDS-PAGE and Western blot analysis (**Figure 4.3**), consistent with the presence of additional Asn-linked glycan. The effect of the introduced Asn-linked glycan on WIF-1 function was evaluated using the cellular assay (**Figure 4.6**, **A**). For finer scale dissection of potential binding surfaces 6 additional single residue mutations were tested (**Figures 4.6**, **A** and **4.7**, **B**). Out of the total of 21 WIF-$1_{\Delta C}$ variants assayed Phe174Asn, Phe51Asn/Glu53Ser, Glu151Asn/Asp153Ser and Asp145Asn/Val147Ser for which the mutations introduced Asn-linked glycosylation sites, and Met77Trp, a single point mutation, were the weakest inhibitors of Wnt-3a signalling (**Figures 4.6**, **A** and **4.7**, **A**–**B**). In the context of WIF-$1_{\Delta C}$, the Met77Trp mutation also showed a 10-fold decrease of binding affinity to mouse Wnt-3a in the SPR assay (**Table 4.1**).

I supplemented the site-directed mutagenesis studies by evaluating the role of surface exposed ε-amino groups of lysines in Wnt-3a recognition. WIF-$1_{\Delta C}$ was subjected to reductive lysine methylation (Walter et al., 2006) and tested in the Wnt-3a signalling inhibition assay (**Figures 4.6**, **A** and **4.7**, **C**). The chemical conversion of lysines to dimethyllysines did not impair WIF-1 function suggesting that modification of ε-amino

groups of lysines does not interfere with Wnt-3a binding. I also assessed the effect of sequence variation between species and found that zebrafish WIF-1 inhibited Wnt-3a signalling as efficiently as the human ortholog (**Figures 4.6**, **A** and **4.7**, **D**), consistent with the Wnt-3a-binding surface on the vertebrate WDs being highly conserved. The combined results of the site-directed mutagenesis, lysine methylation and species variation based analyses highlight a discontinuous but discrete area of the WIF-1 WD surface in the proximity of the lipid head group as a putative human Wnt-3a-binding site (**Figure 4.6**, **C**).

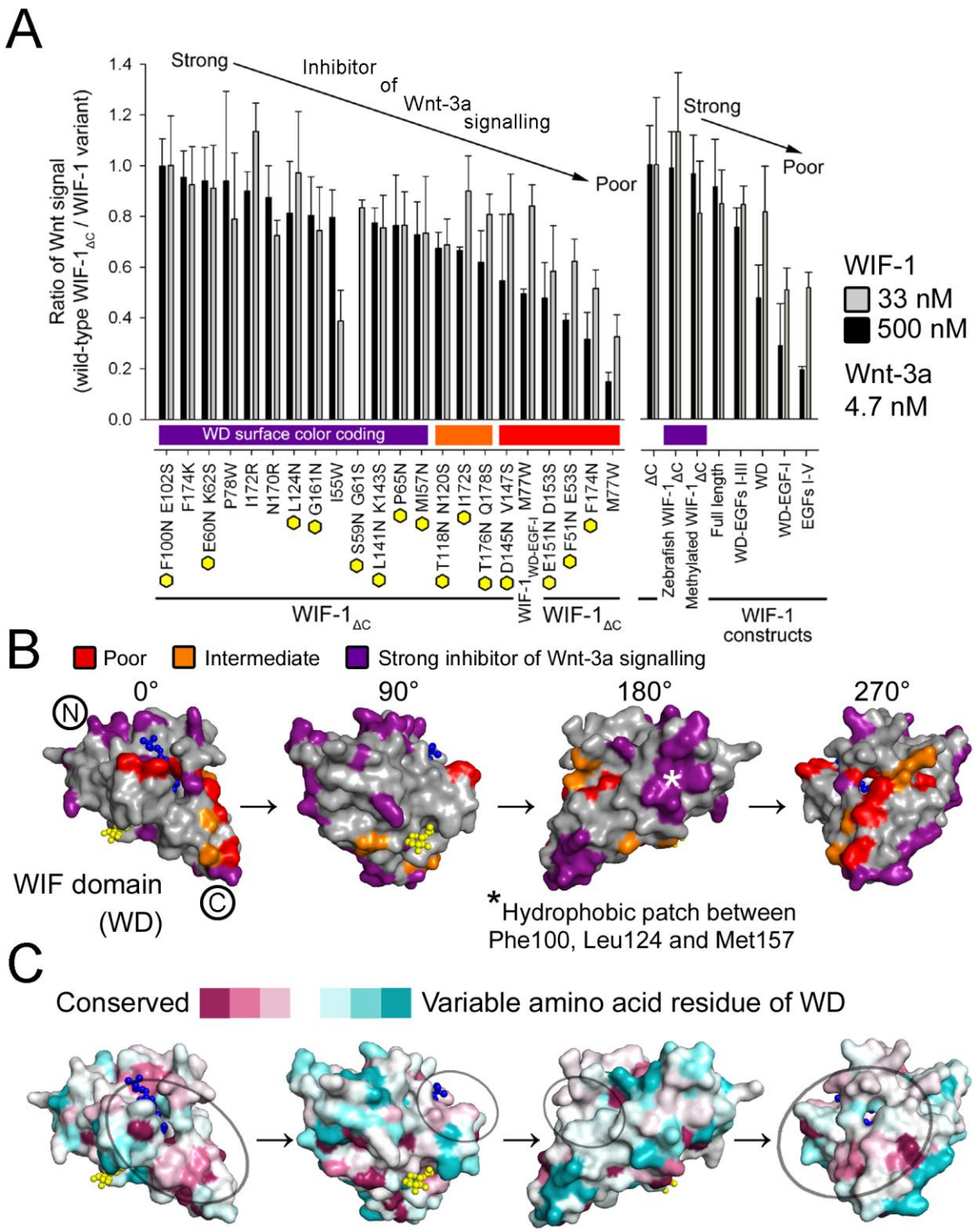

A
Strong
Inhibitor
of
Wnt-3a
signalling
Poor
Strong
Poor
Ratio of Wnt signal
(wild-type WIF-1ΔC / WIF-1 variant)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
WD surface color coding
F100N E102S
F174K
E60N K62S
P78W
I172R
N170R
L124N
G161N
I55W
S59N G61S
L141N K143S
P65N
MI57N
T118N N120S
I172S
T176N Q178S
D145N V147S
WIF-1WD-EGF-I M77W
E151N D153S
F51N E53S
F174N
M77W
ΔC
Zebrafish WIF-1ΔC
Methylated WIF-1ΔC
Full length
WD-EGFs I-III
WD
WD-EGF-I
EGFs I-V
WIF-1ΔC
WIF-1ΔC
WIF-1
constructs
WIF-1
33 nM
500 nM
Wnt-3a
4.7 nM
B
Poor
Intermediate
Strong inhibitor of Wnt-3a signalling
0°
90°
180°
270°
N
C
WIF domain
(WD)
*Hydrophobic patch between
Phe100, Leu124 and Met157
C
Conserved
Variable amino acid residue of WD

Figure 4.6. A discontinuous human Wnt-3a-binding site on the WIF domain (WD). **A**, inhibition of Wnt-3a signalling by mutation-bearing (left bar chart) and wild-type WIF-1 constructs (right bar chart). Inhibition of each WIF-1 construct was compared to the wild-type WIF-$1_{\Delta C}$ in the cellular assay. Higher ratio (0.8–1) of Wnt signal (wild-type WIF-$1_{\Delta C}$ / WIF-1 variant) corresponds to strong, while lower (0.2–0.5) to poor inhibition. Inhibition of human Wnt-3a (4.7 nM) signalling was tested in the presence of 33 nM (grey bar) and 500 nM (black bar) of the WIF-1 variant. Yellow hexagons indicate mutations which introduced Asn-linked glycosylation sites. **B**, residues tested in the cellular assay (**A**) for Wnt-3a signalling inhibition mapped on the WD surface. DPPC (blue) and Asn88-linked GlcNAc (yellow) are shown as ball-and-sticks. Residues for which mutations resulted in poor inhibition of Wnt-3a signalling are coloured in red, residues that were mutated and resulted in the intermediate and strong inhibition of Wnt-3a signalling are coloured in orange and purple, respectively. Phe174 and Ile172 are coloured in red and orange, respectively. A hydrophobic patch on the WD surface (180° view) is marked with an asterisk (*). **C**, the WD surface coloured by amino acid residue conservation (conserved – magenta, variable – cyan). Fourteen Wnt-binding WDs (**Figure 3.8**, **A** and **Figure 4.11**) are included in the sequence conservation analysis. WDs from Hh-binding *Drosophila* WIF-1 (Shifted) are excluded. A potential Wnt-3a-binding surface is circled.

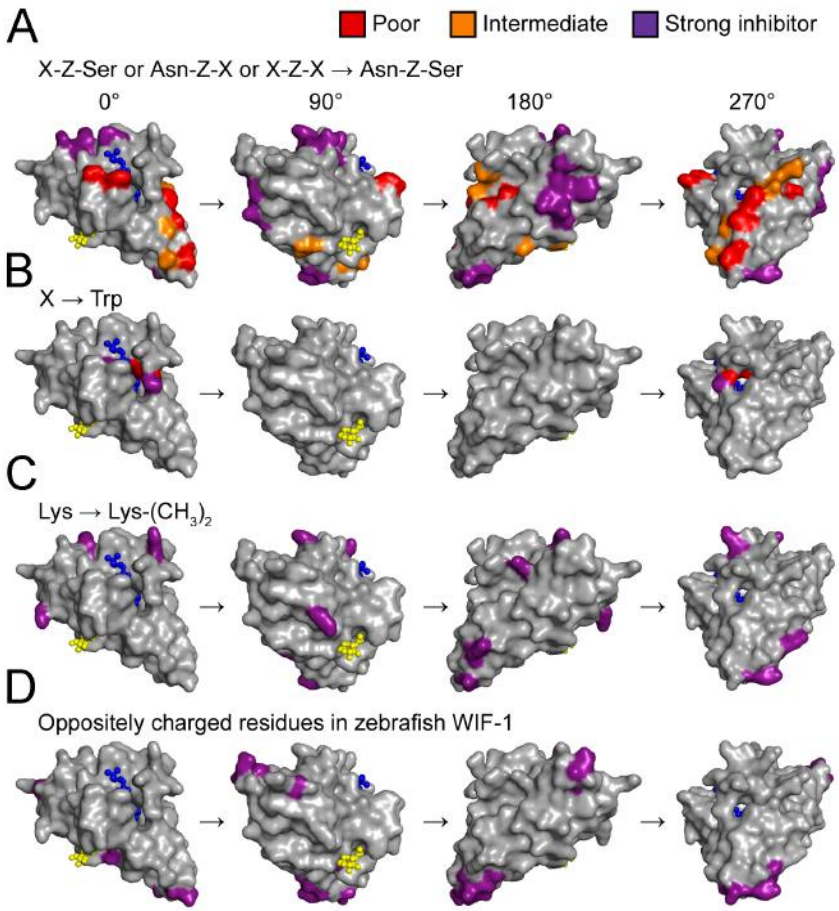
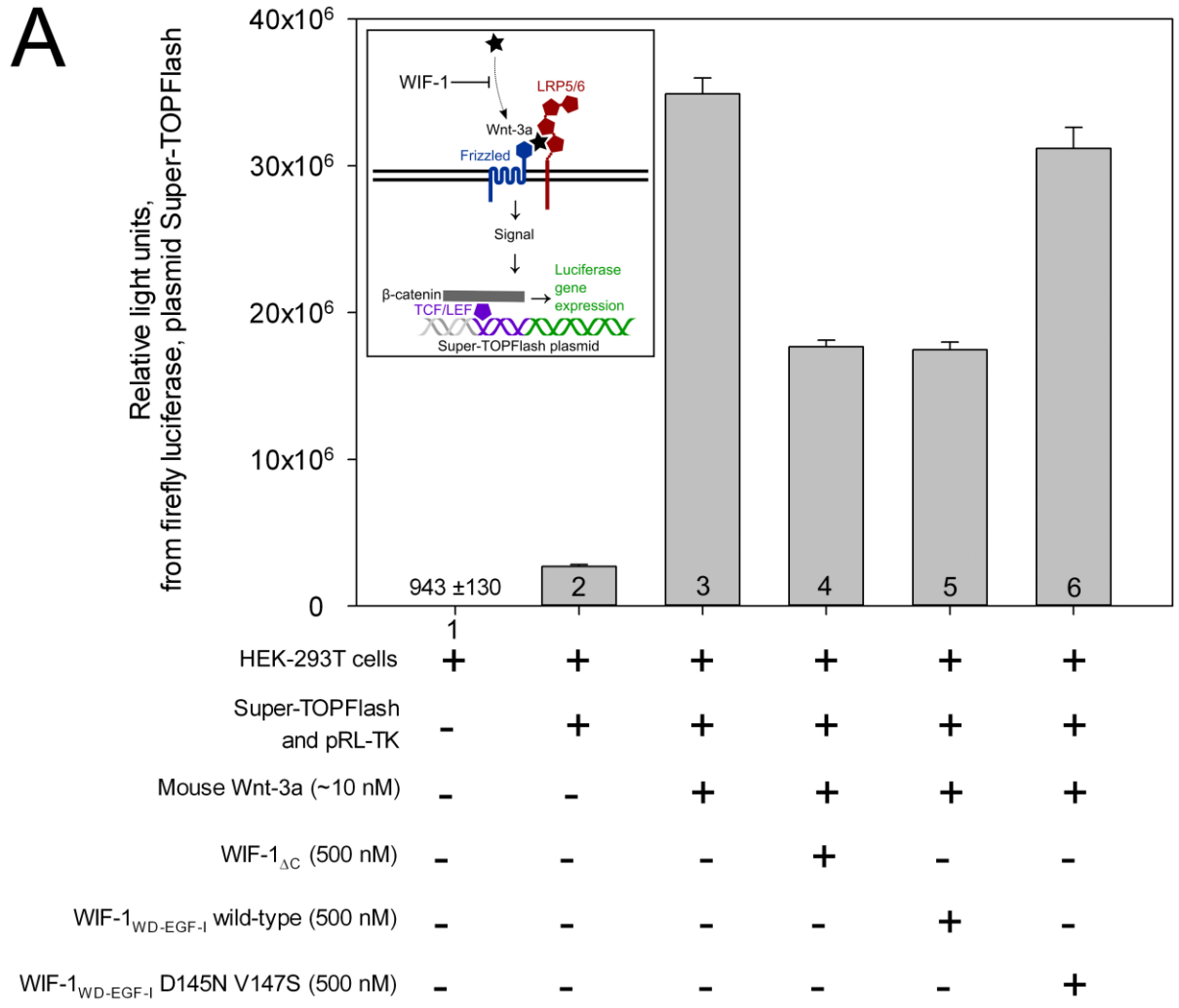


Figure 4.7. Residues tested in the cellular assay for human Wnt-3a signalling inhibition mapped on the WD surface. **A**, the locations of introduced glycosylation sites on the WD surface. Orientation of the WD is the same in the 0° view as in **Figure 3.7**, **C** (0° view). The wild-type GlcNAc moiety on Asn88 is shown as yellow ball-and-sticks. Residues for which the putative glycosylation site generating mutations resulted in poor inhibition of human Wnt-3a signalling are coloured in red (Phe174, Phe51, Glu53, Glu151, Asp153, Asp145, and Val147), residues that were mutated and resulted in intermediate (Thr176, Gln178, Ile172, Thr118, and Asn120) and strong (Met157, Phe65, Leu141, Lys143, Ser59, Gly61, Gly161, Leu124, Glu60, Lys62, Phe100, and Glu102) inhibition of Wnt-3a signalling are coloured in orange and purple, respectively. Colour coding corresponds to **Figure 4.6**, **A** and **B**. **B**, the locations of mutations which introduced tryptophans: Met77 (red), Ile55 (purple), and Pro78 (purple). **C**, the locations of lysines (purple). **D**, the locations of human WIF-1 residues for which the equivalent residues in zebrafish WIF-1 differ in electrostatic charge (purple, human WIF-1 residue and number/corresponding zebrafish WIF-1 residue): Ser86/His, Gly95/Asp, Gly112/Asp, Ala115/Asp, Leu141/Arg, Lys143/Asp, Asn158/Asp, Gln179/Arg, and Glu160/Gly. The combined information from **A** to **D** is presented in **Figure 4.6**, **B**. It indicates a potential WD surface that is (orange and red) and is not (purple) recognized by human Wnt-3a.

To further validate the potential Wnt-3a binding site, we have tested WIF-$1_{WD\text{-}EGF\text{-}I}$ Asp145Asn/Val147Ser in the Wnt/β-catenin cellular assay. The double mutation Asp145Asn/Val147Ser reduced the inhibitory effect of WIF-1 in the context of the WIF-$1_{\Delta C}$ construct (**Figure 4.6**, **A**) but was not evaluated in the context of the WIF-1 construct lacking EGF-like domains. The WIF-$1_{WD\text{-}EGF\text{-}I}$ Asp145Asn/Val147Ser construct lacks four C-terminal EGF-like domains that contribute to the inhibitory function of WIF-1.

Wnt/β-catenin signalling pathway was readily activated using mouse Wnt-3a purified in-house (this experiment also serves as a control experiment to show that mouse Wnt-3a purified in-house and used for the SPR binding assays (see **Section 4.7**) is biologically functional; the amino acid sequence identity between human and mouse Wnt-3a is 96 %) (final concentration ~10 nM; Wnt signalling level taken as 100 %; **Figure 4.8**, **A**). The mouse Wnt-3a signalling level was reduced to 51 and 50 % by the wild-type WIF-$1_{\Delta C}$ and WIF-$1_{WD\text{-}EGF\text{-}I}$ constructs (final concentration 500 nM), respectively. Consistent with the results obtained from the cellular assay using commercially available human Wnt-3a and WIF-$1_{\Delta C}$ Asp145Asn/Val147Ser, WIF-$1_{WD\text{-}EGF\text{-}I}$ Asp145Asn/Val147Ser inhibited mouse Wnt-3a/β-

catenin signalling poorly (Wnt signalling level 89 %). These data are consistent with previous human Wnt-3a/β-catenin signalling and SPR binding assay-derived results, and implicate a region proximal to the lipid head group in Wnt-3a binding (the distance between Val147 and the lipid head group in WIF-$1_{\Delta C}$ structure is 12 Å).


A
Relative light units, from firefly luciferase, plasmid Super-TOPFlash
40x10^6
30x10^6
20x10^6
10x10^6
0
WIF-1
LRP5/6
Wnt-3a
Frizzled
Signal
β-catenin
TCF/LEF
Luciferase gene expression
Super-TOPFlash plasmid
943 ±130
1
2
3
4
5
6
HEK-293T cells
Super-TOPFlash and pRL-TK
Mouse Wnt-3a (~10 nM)
WIF-1ΔC (500 nM)
WIF-1WD-EGF-I wild-type (500 nM)
WIF-1WD-EGF-I D145N V147S (500 nM)

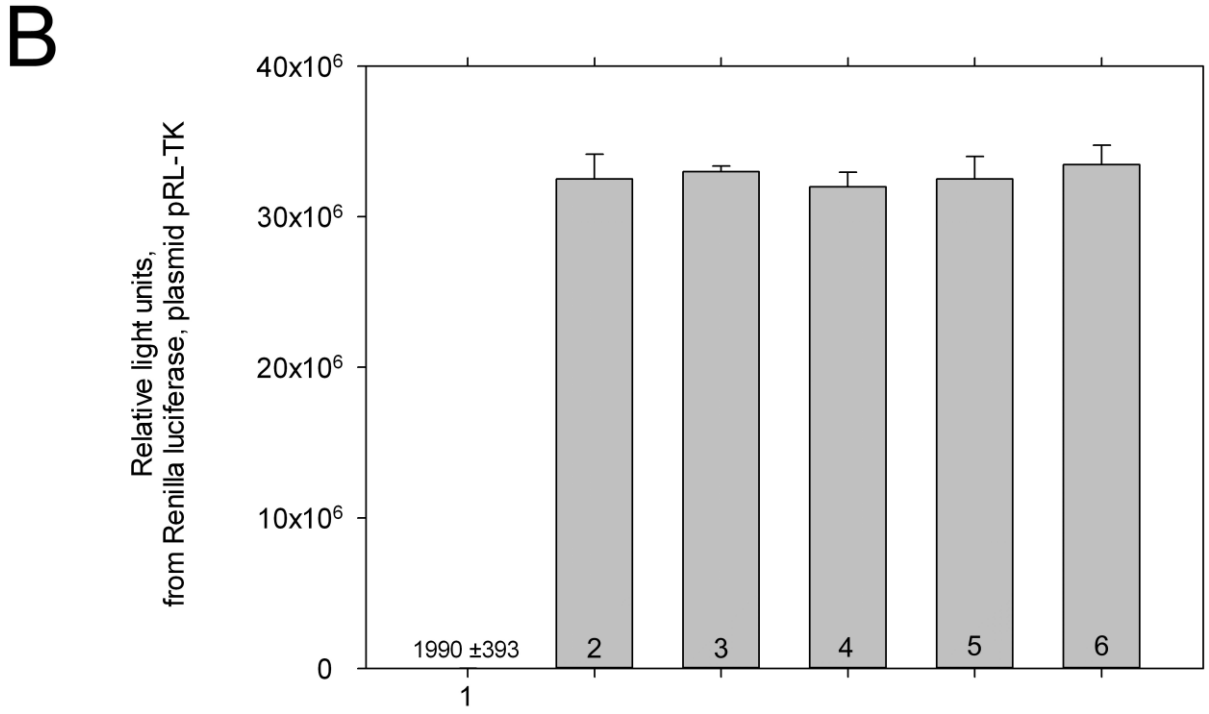

B
Relative light units, from Renilla luciferase, plasmid pRL-TK
40x10^6
30x10^6
20x10^6
10x10^6
0
1990 ±393
1
2
3
4
5
6

Figure 4.8. The inhibition of mouse Wnt-3a/β-catenin signalling by the wild-type WIF-$1_{\Delta C}$, WIF-$1_{WD\text{-}EGF\text{-}I}$ and WIF-$1_{WD\text{-}EGF\text{-}I}$ Asp145Asn/Val147Ser construct. HEK-293T cells were transfected with a plasmid Super-TOPFlash encoding the firefly luciferase gene upstream of a minimal TA promoter and 11 Wnt-3a-responsive TCF/LEF transcription factor binding sites and a plasmid pRL-TK encoding *Renilla* luciferase. Wnt-3a signalling was activated by adding mouse Wnt-3a (final conc. ~10 nM) to the transfected cells (**A**, compare histogram columns 2 and 3, with (+) and without (-) added Wnt-3a, respectively). Wnt-3a signalling was significantly reduced in the presence of wild-type WIF-$1_{\Delta C}$, WIF-$1_{WD\text{-}EGF\text{-}I}$ (A, histogram columns 4 and 5, respectively; final conc. 500 nM), but not mutant WIF-$1_{WD\text{-}EGF\text{-}I}$ (A, histogram column 6; final conc. 500 nM). **B**, pRL-TK plasmid is not Wnt-3a responsive and encodes a *Renilla* luciferase. Enzymatic activity of *Renilla* luciferase was monitored independently of the firefly luciferase activity. A similar level of enzymatic activity of the *Renilla* luciferase in the transfected cells (**B**, histogram columns 2 to 6; column 1 corresponds to untransfected HEK-293T cells) indicates a similar level of transfection efficiency (both for pRL-TK and Super-TOPFlash plasmids) and serves as a quality control experiment. Experiments were performed in triplicate, the height of the error bar is equal to the standard deviation. Further details are provided in **Section 2.11.2**.

4.9 WIF-$1_{WD\text{-}EGF\text{-}I}$ Met77Trp crystal structure highlights a potential Wnt-3a-binding site

Within the putative human Wnt-3a-binding site, the major impact on function of the WIF-$1_{\Delta C}$ Met77Trp mutant highlights the region in the proximity of the lipid head group (**Figure 4.7**, **B**). The side chain of Met77 is partially buried at the entrance to the lipid-binding pocket and provides a hydrophobic environment for the DPPC acyl chains extending from the lipid glycerol backbone (**Figure 4.9**, **A**). To assess changes introduced by the Met77Trp mutation I determined the crystal structure of WIF-$1_{WD\text{-}EGF\text{-}I}$ Met77Trp at 2.22 Å resolution (**Table 4.2** and **Section 2.3**). Superposition of the wild-type and Met77Trp WIF-$1_{WD\text{-}EGF\text{-}I}$ structures revealed no major long-range perturbations (RMSD of 0.22 Å for 170 equivalent Cα atoms) and the acyl chains of the DPPC lipid remain buried within the pocket of the WD. However, unlike the side chain of Met77, which makes no direct interactions with the lipid in the wild-type WIF-$1_{WD\text{-}EGF\text{-}I}$ structure, the tryptophan side chain in WIF-$1_{WD\text{-}EGF\text{-}I}$ Met77Trp forms a hydrogen bond with a carbonyl oxygen of the palmitoyl tail group (distance between the corresponding N and O atoms is 3.1 Å) and oxygen atom of the phosphatidylcholine head group (distance between the corresponding N and O atoms is 4 Å) whilst the indole ring provides stacking interactions which further restrict the orientation of the lipid head group (**Figure 4.9**, **B**–**C**). I assessed whether lipid exchange was still possible in the presence of these hydrogen bonds and hydrophobic interactions. I found that 16:0 Liss Rhod PE could still exchange the natively bound PCs (**Figure 4.10**). Thus the notable reduction in the ability of the WIF-$1_{\Delta C}$ Met77Trp mutant to inhibit Wnt-3a signalling suggests that Met77 and the protein surface proximal to the lipid head group is involved in the Wnt-3a-binding site. The involvement of the lipid may be indirect. I was unable to inhibit wild-type WIF-$1_{WD\text{-}EGF\text{-}I}$ binding to mouse Wnt-3a in SPR experiments using a 40-fold excess of the lipid head group

analogue phosphocholine (**Table 4.1**), suggesting that the contribution of the head group to Wnt-3a binding is limited.

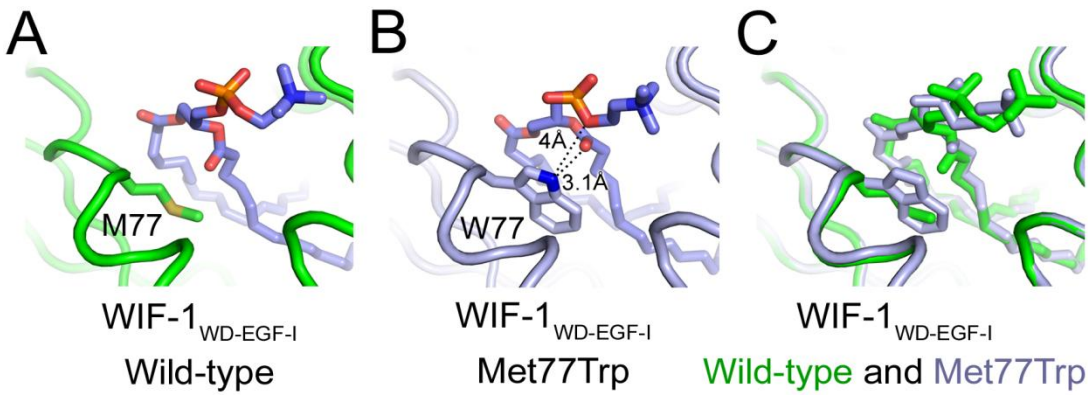


Figure 4.9. Comparison of the wild-type (**A**) and Met77Trp (**B**) WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structures, that are superimposed onto each other in **C**. DPPC is shown as sticks. Nitrogen, oxygen, phosphorus and sulfur atoms are coloured in blue, red, orange and yellow in **A** and **B**, respectively.

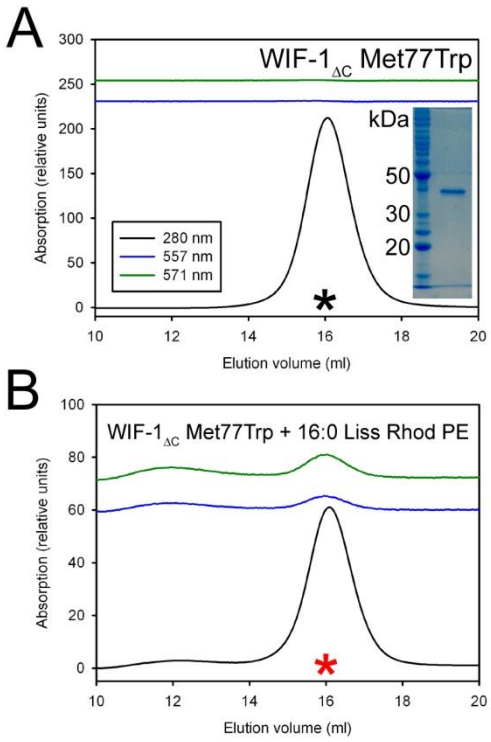


Figure 4.10. Size-exclusion chromatography of the purified WIF-$1_{\Delta C}$ Met77Trp incubated: without fluorescent molecules (**A**) and with 16:0 Liss Rhod PE (**B**). Separation by size exclusion was performed using a Superdex 200 10/300 column (GE Healthcare Life Sciences) at a flow speed of 0.4 ml/min in PBS buffer at room temperature. Protein and

fluorescent molecule elution was monitored by absorbance at 280, 557 and 571 nm (black, blue and green curves, respectively) simultaneously. Binding between WIF-$1_{\Delta C}$ Met77Trp and 16:0 Liss Rhod PE was confirmed by fluorescence (excitation at 535 nm, emission at 595 nm) intensity measurements: intensity of WIF-$1_{\Delta C}$ Met77Trp peak fractions containing 16:0 Liss Rhod PE (indicated with a red asterisk in **B**) was 25-fold (100.0 %) higher than WIF-$1_{\Delta C}$ Met77Trp alone (4.0 %; indicated with a black asterisk in **A**). An inset in **A** shows a reducing SDS-polyacrylamide (15 % w/v) gel of the purified WIF-$1_{\Delta C}$ Met77Trp used in the experiment. Molecular weight (MW) marker: BenchMark Protein Ladder, Invitrogen. The gel was stained with SimplyBlue SafeStain, Invitrogen. Further details of the lipid-exchange experiment protocol are provided in **Section 2.9**.

Table 4.2. Crystallographic data collection and refinement statistics of the $WIF\text{-}1_{WD\text{-}EGF\text{-}I}$ Met77Trp.

Construct	**$WIF\text{-}1_{WD\text{-}EGF\text{-}I}$ Met77Trp**
Data collection	
X-ray source	Diamond, I24
Resolution range (Å)	29.32–2.22 (2.28–2.22)
Space group	C222
Cell dimensions (Å)	a = 50.8 b = 134.2 c = 60.2
Solvent content (%) (molecules per asymmetric unit)	50 (1)
Wavelength (Å)	0.978
R_{merge} (%)	9.4 (57.3)
Total number of observations	66673 (4073)
Total number of unique observations	10433 (683)
Mean ((I)/σ(I))	12.7 (2.3)
Completeness	99.4 (92.0)
Redundancy	6.4 (6.0)
Refinement	
Resolution range (Å)	29.32–2.22 (2.45–2.22)
Number of reflections	10431 (2490)
R_{work} (%)	18.6 (25.0)
R_{free} (%)	22.9 (29.5)
Number of atoms (protein/DPPC/GlcNAc, fucose/sucrose octasulfate/water)	1598 (1412/50/14/0/111)
Wilson B factor ($Å^2$)	35
Average B factor (protein/DPPC/GlcNAc, fucose/sucrose octasulfate/water) ($Å^2$)	51 (50/79/48/–/46)
Ramachandran statistics	
RMSD bonds (Å)	0.015
RMSD angles (degrees)	2.2
Favoured (%)	95.4
Number of outliers	0

4.10 Why does *Drosophila* WIF-1 (Shifted) not bind to Wnt proteins?

Drosophila WIF-1 (also known as Shifted) does not interact with Wnt proteins but is required for the stabilisation and diffusion of lipid-modified Hedgehog morphogens (Glise et al., 2005; Gorfinkiel et al., 2005). Both human WIF-1 and *Drosophila* Shifted share the same domain architecture, comprised of the N-terminal WIF domain (WD, **Figure 4.11**) and five C-terminal EGF-like domains, plus substantial amino acid sequence similarity (26 % sequence identity). Moreover, analysis of the WD sequences (**Figures 3.11** and **3.12**, **A**) shows that the WD of Wnt receptor RYK is similar to the WD of Shifted. Thus it is surprising that the WD of Shifted does not bind to Wnt proteins.

To start to explain this property of Shifted, we simplistically assume that (1) all Wnts interact with all RYK and WIF-1 WDs in a similar way and bind to the same corresponding WD surface, (2) the WD fold is conserved across species. I further consider possibility that for the vertebrate WIF-1 WD amino acid residues involved in Wnt-3a binding (**Sections 4.7** and **4.8**) corresponding similar residues should be present in the Wnt receptor RYK WD, but not Shifted WD; Shifted should have residues different compared to RYK and WIF-1 WDs at the surface which corresponds to the Wnt-3a-binding region. I focus my attention on eight residues identified as important for Wnt-3a recognition and the WD surface encompassing these residues (**Sections 4.7** and **4.8**). The eight human WIF-1 residues (**Figure 3.12**, **B**) and corresponding residues in RYK and Shifted are briefly discussed below.

1. Phe51. Most of the vertebrate WIF-1 proteins have phenylalanine at the corresponding position. The only exception is chicken WIF-1 (which is likely to be a Wnt-binding protein similar to other vertebrate WIF-1 proteins) which lacks any residue at the corresponding position, suggesting that Phe51 is not essential for Wnt binding. Red deer WIF-1, vertebrate RYK, and mosquito Shifted have leucines while

Drosophila Shifted has a tyrosine at the corresponding position. The similar hydrophobic nature of the corresponding residues across vertebrate WIF-1, RYK and Shifted protein families suggests that this residue is not used by Shifted to prevent Wnt binding to the WD.

2. Glu53. Most of the vertebrate WIF-1 proteins and *Drosophila* Shifted have glutamate at the corresponding position. Together with the data for Phe51, which is located proximal to Glu53, it suggests that the region encompassing Phe51–Glu53 is not used by Shifted to prevent Wnt binding to the WD.
3. Met77. *Drosophila* Shifted has leucine while most RYKs have asparagine at the corresponding position which would complicate assignment of the functional significance for this region. Interestingly, the adjacent Pro78 is highly conserved among vertebrate WIF-1 family members but any corresponding residue is absent in *Drosophila* Shifted. Thus it is interesting to speculate how deletion of Pro78 could affect the WD fold and its function in Shifted. This deletion could potentially alter the relative position of the α2 helix (residues 70–77, **Figure 3.12**, **B**) and bring it closer to the predicted Wnt-3a-binding site on the WD. Thus it is tempting to speculate that Shifted may prevent Wnt binding by having an α helix which obstructs ligand binding. The argument against this speculation is the presence of the corresponding proline in other insect Shifted-like proteins (that are not biochemically characterized but, in analogy with *Drosophila* Shifted, are assumed to bind to Hedgehog but not to Wnt proteins).
4. Asp145. Aspartate is conserved among vertebrate WIF-1 and RYK family members while Shifted and two other insect Shifted-like proteins have serine at the corresponding position. Together with our Wnt/β-catenin cellular assay-derived data, this suggests that aspartate is essential for Wnt binding and *Drosophila* Shifted is no

longer able to bind Wnts because it has serine (but not the conserved aspartate) at the corresponding position. However, *C.elegans* LIN-18 WD (which is similar to RYK WD, **Figure 4.11** and **4.12**, **A**) has a serine at the corresponding position, thus it is not possible to conclude that a serine in Shifted is the sole residue that prevents Wnt binding to the WD.

5. Val147. Valine is conserved across vertebrate WIF-1 family members but is also present in Silk moth Shifted-like protein, suggesting that the hydrophobic side chain of the residue is compatible both with Wnt and Hedgehog binding to the WD. In addition, members of the RYK family of Wnt receptors have glutamate at the corresponding position, suggesting that evolutionary conservation of this residue is not essential to maintain Wnt-binding ability of the WD or/and that WDs present in RYK and WIF-1 proteins bind to different Wnt family members.
6. Glu151. Similarly, to the Val146, Glu151 is conserved across the vertebrate WIF-1 family members but is present as glutamate in Silk moth Shifted-like protein too, suggesting that the acidic side chain of the residue is compatible both with Wnt and Hedgehog binding to the WD. In addition, members of the RYK family of Wnt receptors have leucines at the corresponding position, suggesting that the evolutionary conservation of this residue is not essential to maintain the Wnt-binding ability of the WD or/and that WDs present in RYK and WIF-1 proteins bind to different Wnt family members.
7. Asp153. Aspartate is not conserved across the vertebrate WIF-1 family members but is conserved as glycine in three Shifted-like proteins, suggesting that glycine might be important for Shifted-Hedgehog interactions.
8. Phe174. Phenylalanine is conserved across the vertebrate WIF-1 family members, but is present as arginine or lysine in Shifted and RYK family members. Thus it is not

clear whether the residue at this position is more important to maintain Wnt or Hedgehog binding properties. Mutation Phe174Lys had a minor effect on WIF-$1_{\Delta C}$ function (**Figure 4.6**, **A**).

In addition, the WD of Shifted (in contrast to the WD of WIF-1 or RYK) contains an insertion of six amino acids (AlaLeuTyrPheProGln) which may interfere with Wnt binding (**Figures 3.14** and **4.11**). To summarise, the currently available data on *Drosophila* Shifted-Hedgehog interactions is limited to two published studies (Glise et al., 2005; Gorfinkiel et al., 2005) while WIF-1– and RYK–Wnt interactions also remain elusive because both Wnt and WIF-1/RYK proteins are difficult to express and purify in sufficient quantities for biochemical studies. Thus it is difficult to explain how proteins with the evolutionarily conserved domain architecture, WIF-1 in vertebrates and Shifted in insects, can specifically regulate two evolutionarily highly divergent, Wnt and Hedgehog, signalling pathways.

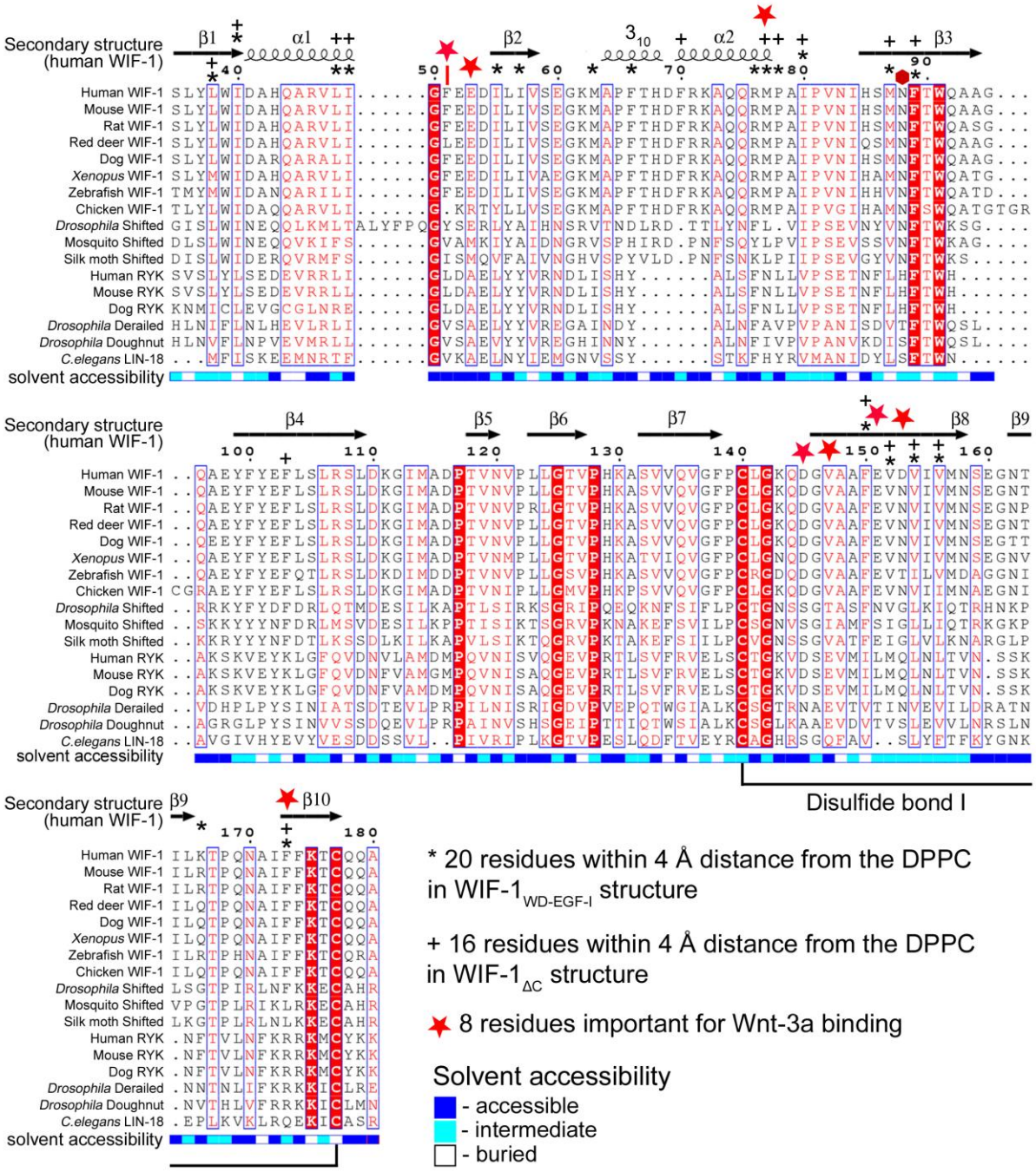


Figure 4.11. Amino acid sequence alignment of the WIF domains. Sequences, alignment protocol and labels are described in **Section 3.62** and **Figure 3.8**, **A**. Eight amino acid residues important for Wnt-3a binding and described in **Section 4.9** are marked with red stars.

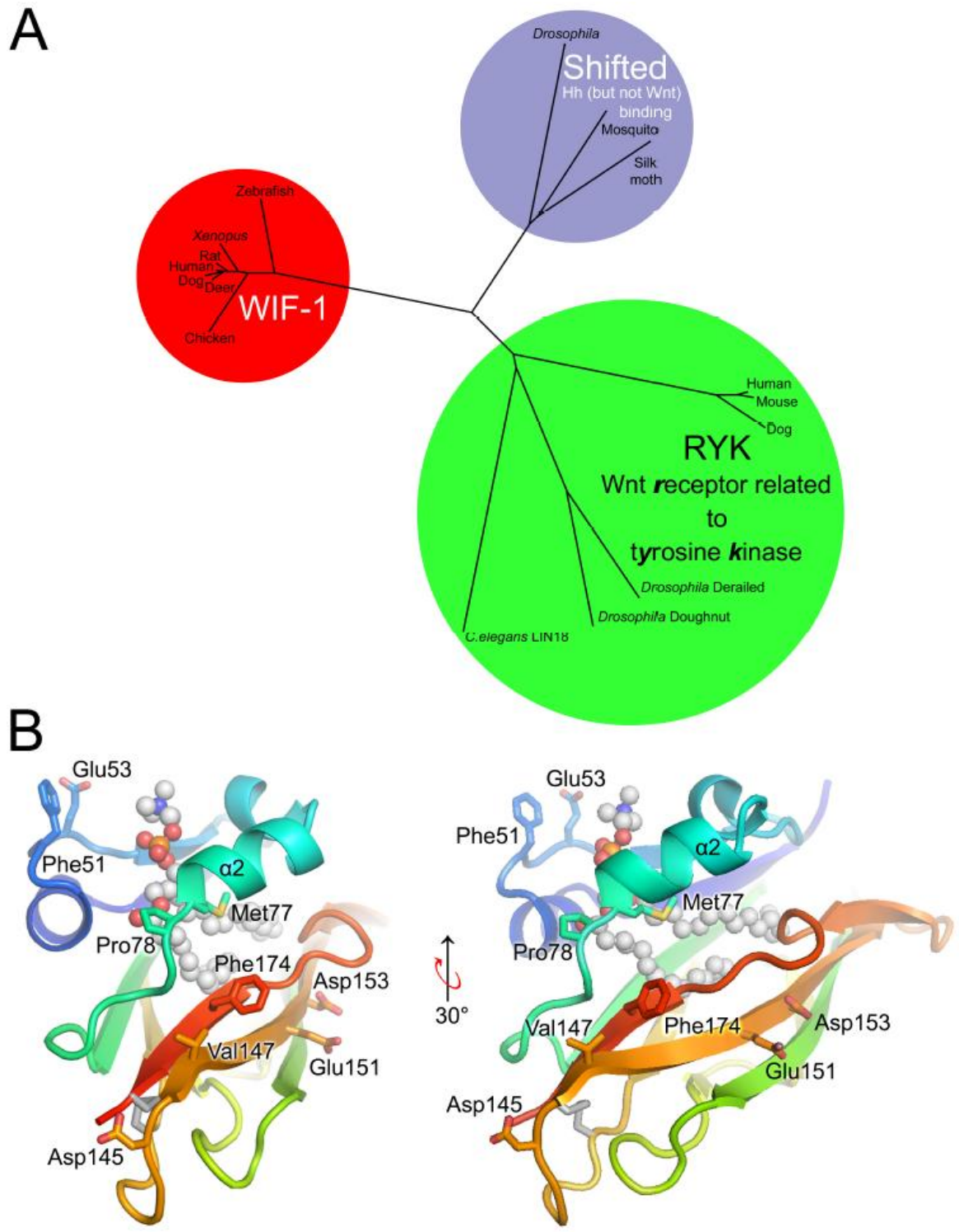


Figure 4.12. **A**, phylogenetic tree of WIF domains. The tree was constructed using the ClustalW (Chenna et al., 2003; http://www.ebi.ac.uk/Tools/msa/clustalw2/), PHYLIP (PHYLogeny Inference Package; Felsenstein, 2004) and Phylodendron 0.8d (by Don Gilbert, http://iubio.bio.indiana.edu/treeapp/treeprint-form.html) programs based on the amino acid sequences presented in **Figure 4.11**. **B**, ribbon diagram of WIF-1_{WD} in two views that differ by a 30° rotation about a vertical axis. WD is colour coded as a rainbow with N- and C-termini shown in blue and red, respectively. Disulfide bond (grey) is shown as sticks. DPPC is shown in spheres (carbons – grey; nitrogen – blue; oxygens – red; phosphorus – orange).

Side chains of the residues important for Wnt-3a binding are labelled and shown as sticks.

4.11 Future perspectives

Biophysical and cellular assays, combined with site-directed mutagenesis, locate a Wnt-3a-binding surface on the WIF domain, proximal to the lipid head group. Results also implicate the five epidermal growth factor (EGF)-like domains in Wnt-3a binding. To gain better understanding of WIF-1–Wnt interactions, future work could be focused upon trying to obtain structures of, first, the WIF-1 construct containing both WIF and five EGF-like domains, and, second, WIF-1 in complex with a Wnt family member. The first challenge will be addressed in **Chapter 5**. The major obstacle to the second challenge, the WIF-1–Wnt complex structure solution, is the lack of pure Wnt proteins in sufficient quantities for structural studies. In addition, although not addressed in this thesis, characterization of other WIF domain containing proteins, such as RYK and *Drosophila* WIF-1 (Shifted) is of high value. Understanding of RYK–Wnt and Shifted–Hedgehog interactions could provide insights into regulation of two signalling pathways, Wnt and Hedgehog, essential for the development and maintenance of multicellular organisms.

Chapter 5 – Contribution of the EGF-like domains to WIF-1 function

5.1 Summary

In **Chapter 4**, I found that both five EGF-like domains and a WIF domain (WD) of human Wnt inhibitory factor 1 (WIF-1) directly bind to Wnt-3a and contribute to Wnt-3a/β-catenin signalling inhibition. However, the structural and functional analysis of the WD (WIF-1_{WD}) and WD-EGF-like domain I (WIF-$1_{WD\text{-}EGF\text{-}I}$) constructs I described in **Chapter 3** provides limited understanding of how the five EGF-like domains contribute to signalling inhibition. To start to answer this question I set out to determine the crystal structure of the WIF-$1_{\Delta C}$ construct containing both WD and EGF-like domains I to V. The WIF-$1_{\Delta C}$ crystal structure reveals EGF-like domains I–V wrapped-back to interface with the WD at EGF-like domain III. The wrapped-back conformation is consistent with both the WD and EGF regions contributing to an extended Wnt-3a-binding surface and thus inhibitory WIF-1 function. In addition, I identified a heparan sulfate (HS)-binding site on the EGF-like domains II–V, consistent with evolutionarily conserved, positively charged residues present in EGF-like domain IV. This combination of HS- and Wnt-binding properties suggests a modular mechanism for WIF-1 localisation on the cell surface, and signalling inhibition, within Wnt morphogen gradients.

5.2 Introduction

In **Chapter 4**, I showed that both an N-terminal WD and C-terminal EGF-like domains I–V of human WIF-1 contribute to Wnt-3a signalling inhibition (**Section 4.6**), and directly bind to Wnt-3a in a surface plasmon resonance-based binding assay (**Section 4.7**). The contribution of the EGF-like domains to inhibitory WIF-1 function extends the previous observation that WD is sufficient for Wnt signalling inhibition (Hsieh et al., 1999). The described contribution

of the EGF-like domains to WIF-1 function (**Sections 4.6–4.7**) is consistent with the necessity of the EGF-like domains for full WIF-1 activity in *Xenopus* embryo-based assays (Hsieh et al., 1999). However, my structural and functional analysis of the WD (WIF-1_{WD}) and WD-EGF-like domain I (WIF-$1_{WD\text{-}EGF\text{-}I}$) constructs, described in **Chapter 3**, provides only a limited understanding of how EGF-like domains contribute to Wnt signalling inhibition. To understand the role of the EGF-like domains better, I performed a structural and functional analysis of the WIF-$1_{\Delta C}$ construct containing both WD and EGF-like domains I–V as described in **Chapter 5**.

The additional interest in the function of the human WIF-1 EGF-like domains stems from two studies on *Drosophila* WIF-1, called Shifted (Gorfinkiel et al., 2005; Glise et al., 2005). Shifted is the ortholog of the human WIF-1 and does not interact with the *Drosophila* Wnt protein, called Wingless. In contrast to human WIF-1, Shifted is required for Hedgehog morphogen diffusion and signalling. Moreover, the EGF-like domains (but not the WD) of Shifted are suggested to act as a protein (in this case, Hedgehog) recognition unit (Gorfinkiel et al., 2005). Structural and biochemical studies of Shifted EGF-like domains–Hedgehog interactions are absent. Thus analysis of the human WIF-1 EGF-like domains could potentially provide the first structural insights into Shifted–Hedgehog interactions too.

Studies of the Shifted-shaped Hedgehog diffusion and signalling also highlight the importance of morphogen gradients (Gorfinkiel et al., 2005; Glise et al., 2005). Morphogens, such as Wnts and Hedgehogs, form concentration gradients to specify different cell fates in a concentration-dependent manner (reviewed by Tabata and Takei, 2004). How morphogen gradients are established and maintained is a key question in developmental biology. Covalent lipid modifications provide a mechanism by which Wnt and Hedgehog morphogens can be anchored to cell membranes for short-range activity or co-localise with lipoprotein particles for long-range action (Panáková et al., 2005). More generally, a restricted diffusion

model fits experimental observations and suggests that diffusing morphogens interact with their receptors, antagonists and extracellular matrix components such as HSPGs (reviewed by Yan and Lin, 2009). HSPGs consist of a protein core to which heparan sulfate (HS) glycosaminoglycan (GAG) chains are covalently linked. HSPGs bind to Wnts and Hedgehogs, and modulate their diffusion, prevent dilution in the extracellular space and abnormal signalling in the distantly located cells. The molecular mechanisms of Wnt signalling inhibition within lipoprotein particle- and HSPG-shaped morphogen gradients are poorly understood because post-translationally modified Wnts (notably, Wnts are lipid-modified; Willert et al., 2003; Takada et al., 2006) and their antagonists are difficult to purify in sufficient quantities for biochemical studies. However, Shifted has been shown to co-localise with HSPGs and Hedgehogs. This raises a question whether human WIF-1 is a HS-binding protein and has a potential to co-localise with Wnts and HSPGs. This question is addressed in **Chapter 5**.

5.3 Aims

The main aim of **Chapter 5** is to investigate the crystal structure of human WIF-$1_{\Delta C}$ to provide insights into the combined contribution of the WD and EGF-like domains to WIF-1 function. The second aim is to investigate the heparan sulfate binding properties of WIF-1 and provide insights into heparan sulfate proteoglycan-shaped Wnt gradient control.

5.4 Crystallisation and structure determination of WIF-$1_{\Delta C}$

To investigate further the mechanism of Wnt signalling inhibition, the human WIF-$1_{\Delta C}$ (residues Ser35–Ile346) construct was transiently expressed in mammalian HEK-293S cells, purified and crystallised as described in **Sections 2.2.2** and **4.5.1**. As reductive lysine methylation was essential to obtain crystals of WIF-$1_{WD\text{-}EGF\text{-}I}$ that diffract X-rays beyond 2 Å resolution (**Section 3.5.2**), the same strategy was initially applied during WIF-$1_{\Delta C}$

crystallisation. However, in contrast to WIF-$1_{WD\text{-}EGF\text{-}I}$, crystals of the methylated WIF-$1_{\Delta C}$ (**Figure 5.1**, **A**) have never diffracted X-rays beyond 4 Å resolution. Analysis of the methylated WIF-$1_{\Delta C}$ low resolution structures suggested that methylated lysines do not make major crystal contacts as observed in WIF-$1_{WD\text{-}EGF\text{-}I}$ crystals (**Figure 3.6**, **C**). To improve the diffraction quality of WIF-$1_{\Delta C}$ crystals, the lysine methylation step was omitted. Protein was crystallised in the presence of sucrose octasulfate (SOS). SOS mimics the structure of heparan sulfate (HS) which is a known ligand for *Drosophila* WIF-1 (Shifted), and thus potentially may bind to the structurally similar human WIF-1, and contribute to the diffraction quality of crystals. WIF-$1_{\Delta C}$ crystals in the presence of SOS appeared after 5 days and grew over 1 month (**Figure 5.1**, **B**). A diffraction data set was collected from a single crystal (**Figure 5.1**, **B**) cryoprotected using crystallisation solution supplemented with 25 % v/v glycerol (as described in **Section 2.4**) at 100 K on beamline I03 at the Diamond Light Source synchrotron (**Figure 5.1**, **C**). Statistical descriptors of the WIF-$1_{\Delta C}$ X-ray diffraction data set are presented in **Table 5.1**.

The crystal structure of WIF-$1_{\Delta C}$ was phased by molecular replacement using the WIF-1_{WD} crystal structure with DPPC omitted. Phases were improved by solvent flattening and histogram matching using the program Parrot (Zhang et al., 1997). The resulting electron density map was sufficiently clear to place models of the first three EGF-like domains (**Figure 5.2**). The structure of WIF-$1_{\Delta C}$ in complex with DPPC and SOS was refined to 3.95 Å resolution as described in **Section 2.6.1**. WIF-$1_{\Delta C}$ crystals contain one molecule per asymmetric unit and 82 % solvent (**Figure 5.3**). The refined WIF-$1_{\Delta C}$ structure has good values for crystallographic validation parameters compared to other low resolution structures (among the top 9 % best refined structures of 3 – 9999 Å resolution range as evaluated with the program MolProbity (Davis et al., 2007)) (**Table 5.1**).

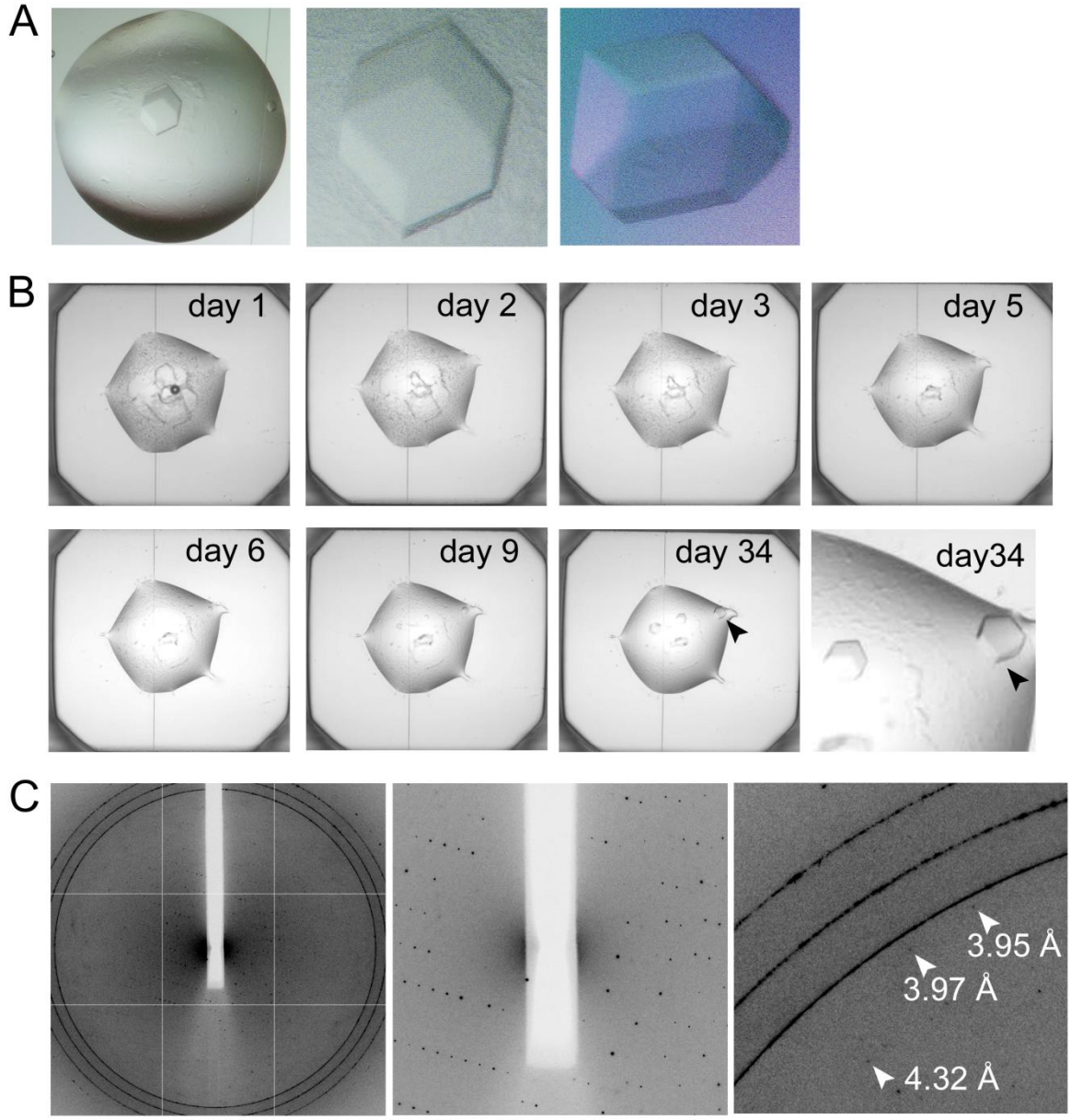


Figure 5.1. Crystallisation of WIF-$1_{\Delta C}$. **A**, crystal of the deglycosylated, methylated WIF-$1_{\Delta C}$ grown in a sitting-drop containing 200 nl of the methylated WIF-$1_{\Delta C}$ (7.45 mg/ml) in 300 mM NaCl, 20 mM Tris-HCl pH 8.0 plus 200 nl of 0.7 M sodium succinate pH 5.51 (left) at 21 °C. A zoom-in view of the same crystal in the drop (middle) and in a cryoprotectant solution (0.7 M sodium succinate pH 5.51, 30 % v/v glycerol; polarized light was used to illuminate the crystal) (right). **B**, growth of the deglycosylated (not methylated) crystals of WIF-$1_{\Delta C}$ monitored over time. The drop contains 100 nl WIF-$1_{\Delta C}$ (6.7 mg/ml in 300 mM NaCl, 20 mM Tris-HCl pH 8.0, 70 mg/ml sucrose octasulfate sodium salt) plus 100 nl 0.75 M sodium succinate pH 5.7 at 21 °C. **C**, crystals shown in **B** were frozen in 0.75 M sodium succinate pH 5.7, 25 % v/v glycerol and diffraction data were collected as described in **Section 2.4**. Although the crystal freezing was suboptimal (as indicated by three prominent regions of X-ray diffraction from ice-like material at 3.44, 3.66 and 3.90 Å resolution; visible on the left and right images), X-ray reflections were relatively well separated and sharp (best visible in the middle and right images). High resolution (up to 3.95 Å) reflections are indicated with arrows on the right image. White rectangular regions (left and middle images) correspond to the beamstop which is impenetrable to X-rays and was used to protect the X-ray detector. The dimensions of the crystal were ~100 × 100 × 100 μm. The crystal was exposed over a 1° oscillation range. The edge of the detector corresponds to a resolution of 3.76 Å. Crystallographic data collection and refinement statistics for the WIF-$1_{\Delta C}$ are shown in **Table 5.1**.

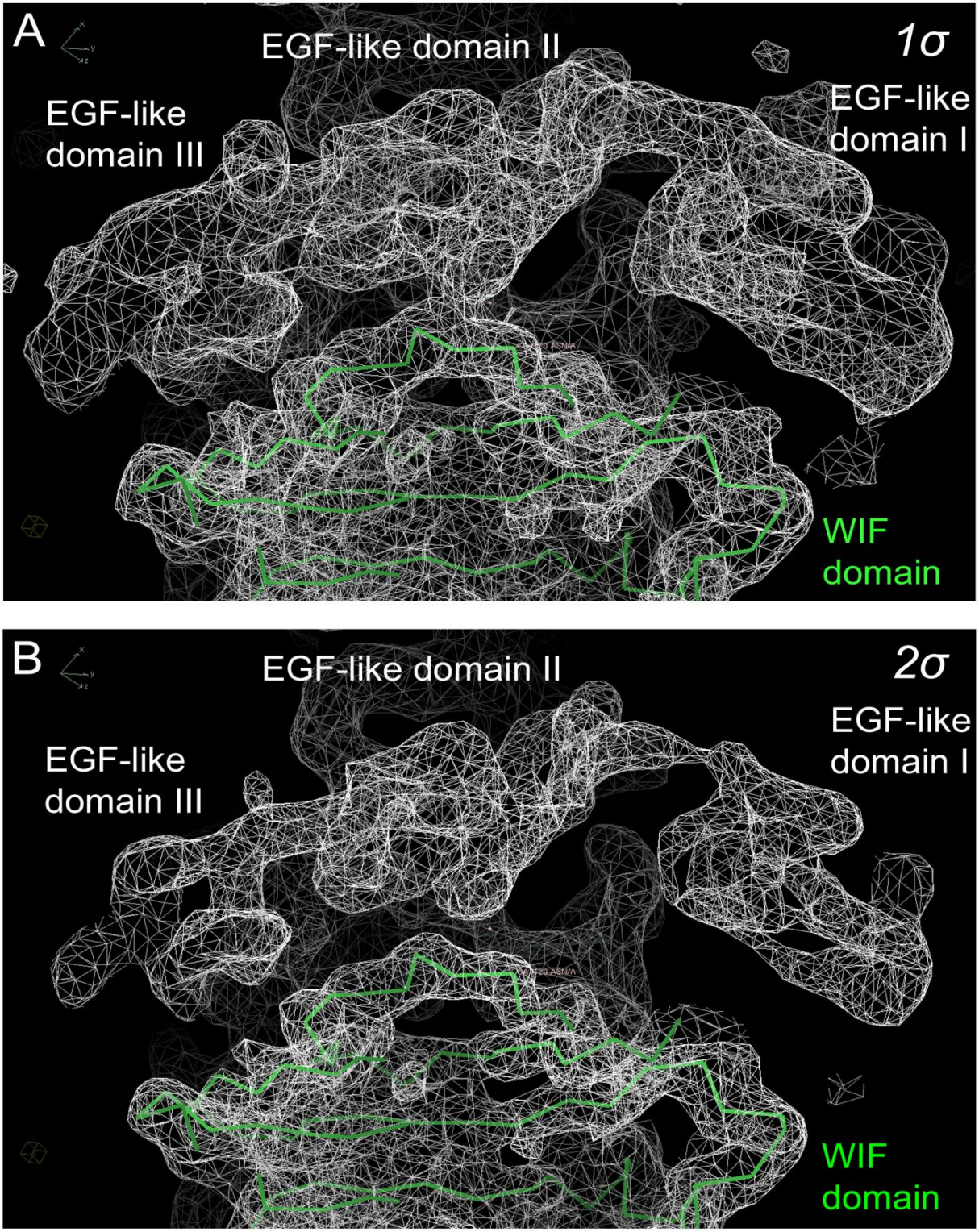


Figure 5.2. Initial WIF-$1_{\Delta C}$ $2F_o$-F_c electron density map. The initial WIF-$1_{\Delta C}$ $2F_o$-F_c electron density map was calculated after molecular replacement using the high resolution crystallographic structure of the WIF domain as a search model and subsequent phase improvement by solvent flattening (crystals of WIF-$1_{\Delta C}$ contain 82 % solvent) and histogram matching with the program Parrot (Zhang et al., 1997). Electron density is contoured at 1 σ (**A**) and 2 σ (**B**) in program Coot (Emsley et al., 2010). The Cα backbone of the WIF domain is shown in green. Electron density corresponding to the EGF-like domains I to III is apparent and labelled.

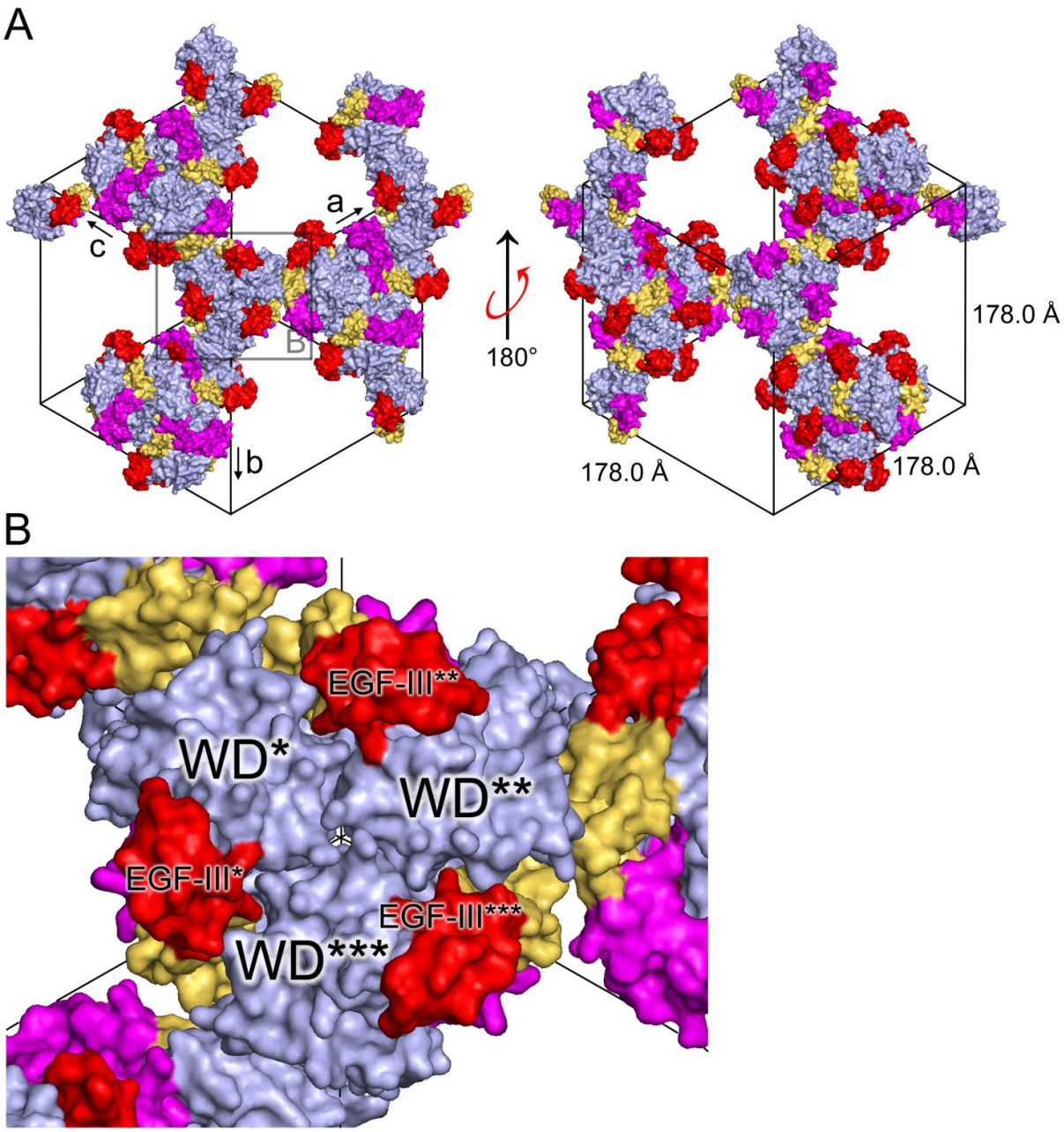


Figure 5.3. Crystal packing of WIF-$1_{\Delta C}$ crystals in space group $I2_13$. **A**, surfaces of the WIF domain (WD) and EGF-like domains I to III are coloured in light blue, magenta, yellow and red, respectively. Unit cell dimensions are indicated on the right and correspond to **Table 5.1**. The three unit cell constants (a, b and c) are indicated. Two views of the crystal packing were obtained by 180° rotation of the unit cell around the vertical axis (lattice vector **b**). **B**, a close-up view of three WIF-$1_{\Delta C}$ molecules arranged around the unit cell origin. A WD and EGF-like domain III that belong to the same molecule are indicated with the same number of asterisks (*, ** or ***). EGF-like domains IV–V extend into the solvent channel and are not presented. DPPC, glycans and sucrose octasulfate are not shown.

Table 5.1. Crystallographic data collection and refinement statistics of WIF-$1_{\Delta C}$.

Construct	**WIF-$1_{\Delta C}$**
Data collection	
X-ray source	Diamond, I03
Resolution range (Å)	40.00–3.95 (4.09–3.95)
Space group	$I2_13$
Cell dimensions (Å)	a = b = c = 178.0
Solvent content (%) (molecules per asymmetric unit)	82 (1)
Wavelength (Å)	0.976
R_{merge} (%)	10.0 (87.4)
Total number of observations	92940 (8938)
Total number of unique observations	8403 (820)
Mean ((I)/σ(I))	27.2 (2.8)
Completeness	99.9 (99.8)
Redundancy	11.1 (10.9)
Refinement	
Resolution range (Å)	39.80–3.95 (4.52–3.95)
Number of reflections	8048 (2525)
R_{work} (%)	29.7 (31.7)
R_{free} (%)	33.1 (32.0)
Number of atoms (protein/DPPC/GlcNAc, fucose/sucrose octasulfate/water)	1969 (1826/50/38/55/0)
Wilson B factor ($Å^2$)	141
Average B factor (protein/DPPC/GlcNAc, fucose/sucrose octasulfate/water) ($Å^2$)	231 (221/208/303/540/–)
RMSD bonds (Å)	0.014
RMSD angles (degrees)	2.1
Ramachandran statistics	
Favoured (%)	93.7
Number of outliers	0

5.5 The WIF-$1_{\Delta C}$ structure reveals the wrapped-back orientation of the EGF-like domains

The WIF-$1_{\Delta C}$ structure reveals a novel molecular arrangement consisting of a lipid-binding WD loosely wrapped by EGF-like domains I–III and ending with EGF-like domains IV–V extending away from the WD into the solvent (**Figures 5.3** and **5.4**). The electron density corresponding to EGF-like domains IV–V in the WIF-$1_{\Delta C}$ crystal structure is fragmentary, suggesting conformational flexibility of these domains. Thus EGF-like domains IV–V were omitted from the final WIF-$1_{\Delta C}$ refined coordinates.

The WD retains structure determined at higher resolution for the isolated domain of WIF-1_{WD} (root mean square deviation (RMSD) of 0.53 Å for 143 equivalent Cα atoms). DPPC again provides a suitable model for the additional electron density observed in the WD core. The extended conformation of the DPPC acyl chains in the WD lipid-binding cavity is conserved between all four of the WIF-1 structures (WIF-1_{WD}, WIF-$1_{WD\text{-}EGF\text{-}I}$, WIF-$1_{WD\text{-}EGF\text{-}I}$ Met77Trp and WIF-$1_{\Delta C}$) as described in this thesis. In contrast, the position of the DPPC head group in WIF-$1_{\Delta C}$ differs markedly from that observed in the WIF-1_{WD}, WIF-$1_{WD\text{-}EGF\text{-}I}$ and WIF-$1_{WD\text{-}EGF\text{-}I}$ Met77Trp crystal structures. The lipid head group is orientated towards the C-terminal region of the WD and the linker to EGF-like domain I in the WIF-$1_{\Delta C}$ structure (**Figure 5.4**, **A**–**C**). This DPPC head group orientation is orthogonal compared to the one observed in crystals of the shorter WIF-1 constructs such as WIF-$1_{WD\text{-}EGF\text{-}I}$ (**Figure 5.4**, **C**). This re-orientation seems to occur because of differences in arrangement of neighbouring molecules within the various crystal lattices. This suggests that the orientation of the lipid head group can alter in response to external forces.

The second notable difference between the structures of WIF-$1_{\Delta C}$ and WIF-$1_{WD\text{-}EGF\text{-}I}$ is the relative orientation of EGF-like domain I to the WD. The EGF-like domain I is rotated by approximately 30° towards the WD in the WIF-$1_{\Delta C}$ structure. With this orientation EGF-like

domains I–V wrap back along the WD so that residues Phe257 and Ile263 of the EGF-like domain III make hydrophobic interactions with WD residues Phe100, Leu124, and Met157 (**Figure 5.4**, **D**; also marked with an asterisk (*) on the 180° view in **Figure 4.6**, **B**). These three variants of WIF-$1_{\Delta C}$ with mutations introducing Asn-linked glycosylation sites in the WD (Phe100Asn/Glu102Ser, Leu124Asn, and Met157Asn) (**Figure 4.3**) were designed to disrupt the WD–EGF-like domain III hydrophobic interface. Three WIF-$1_{\Delta C}$ variants did not show any impairment of WIF-1 inhibitory function in the cellular Wnt signalling assay (**Figure 4.6**, **A–B**). These results argue against the direct involvement of the hydrophobic WD region in Wnt-3a recognition. The interface between WD and EGF-like domain III is not extensive; the total buried surface area between the domains is 260 Å as calculated with a program PISA (Krissinel and Henrick, 2007). However, this interface does serve to define a preferred relative orientation between the WD and EGF-like domains. This juxtaposition of WD and EGF-like domains is consistent with both regions contributing to an extended Wnt-3a-binding surface. Thus the structure of WIF-$1_{\Delta C}$, in combination with the results of SPR binding studies (**Section 4.7**) and the cellular Wnt signalling assay (**Section 4.8**), suggests a modular mechanism for Wnt binding combining sites on the WD and EGF-like domains.

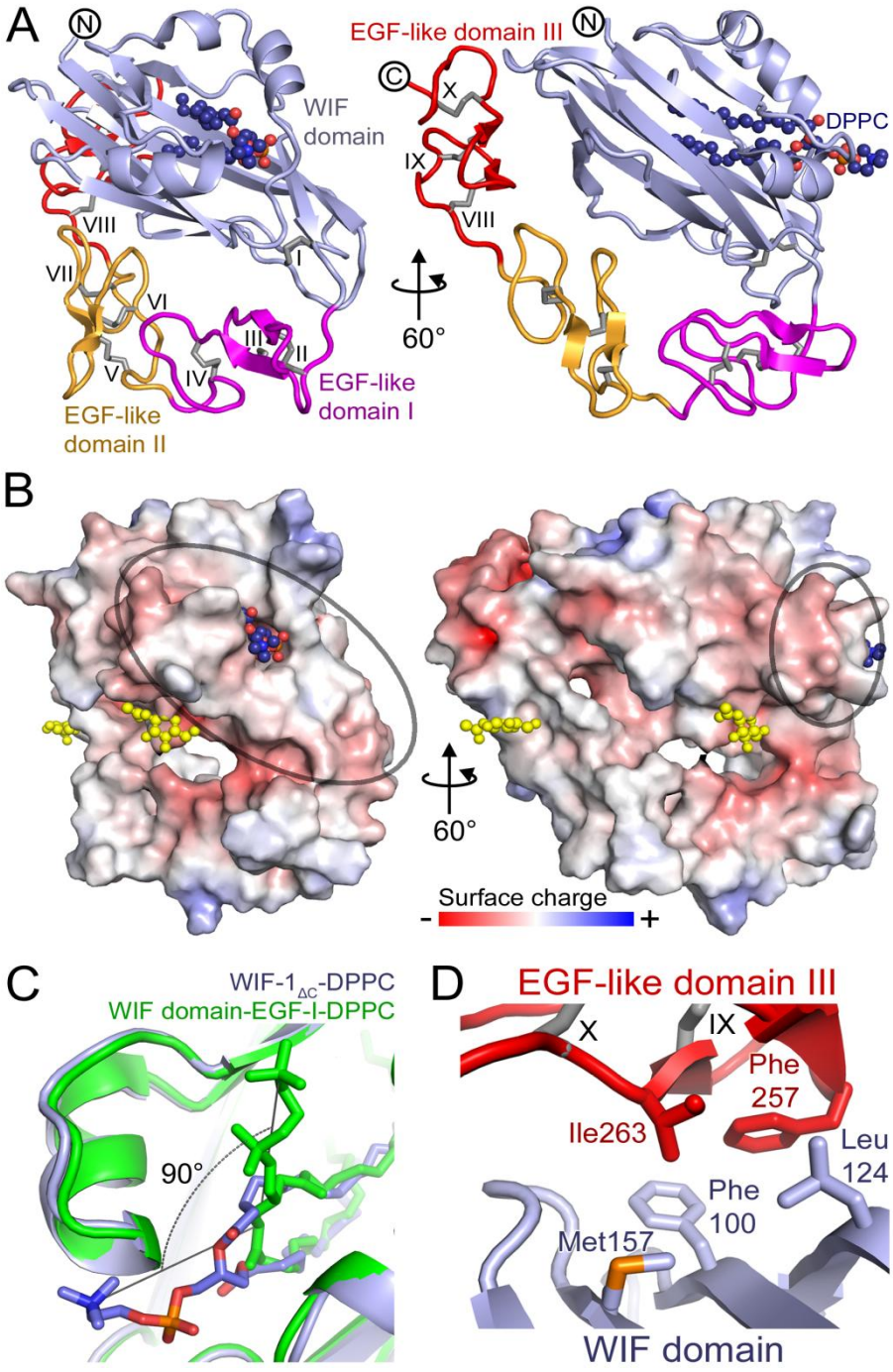


Figure 5.4. The structure of WIF-$1_{\Delta C}$. **A**, ribbon diagram of WIF-$1_{\Delta C}$. The orientation of the WD is the same as in **Figure 3.7**, **C**, 0° view. The WIF domain and EGF-like domains I to III are coloured in light blue, magenta, yellow, and red, respectively. Ten disulfide bridges are shown as grey sticks and marked with Roman numerals. DPPC is shown as spheres. **B**, electrostatic properties of WIF-$1_{\Delta C}$. The protein is shown as solvent-accessible surface coloured by electrostatic potential contoured at ± 8 kT/e (red, acidic; blue, basic). GlcNAc moieties on Asn88 (WD) and Asn245 (EGF III) are shown as yellow spheres. A potential Wnt-3a-binding surface is circled. **C**, orthogonal relative orientation of the DPPC head group between the superimposed structures of WIF-$1_{\Delta C}$ (light blue; DPPC is coloured as in **Figure 3.7**, **D**) and WIF-$1_{WD\text{-}EGF\text{-}I}$ (green). **D**, hydrophobic contact formed between the WD hydrophobic patch (marked with an asterisk (*) in **Figure 4.6**, **B**, 180° view) and EGF III. Interacting residues are shown as sticks. **Figure 5.4** is an adapted version of Figure 4 from

Malinauskas et al., 2011.

5.6 WIF-1 binds to glycosaminoglycans via the EGF-like domains II–V

Genetic studies on *Drosophila* Shifted, the ortholog of human WIF-1, suggest that the EGF-like domains I–V mediate interactions with Hedgehog morphogens and heparan sulfate proteoglycans (HSPGs) (Gorfinkiel et al., 2005; Glise et al., 2005). These interactions were required for Shifted–Hedgehog co-localisation on the cell surface and morphogen signalling. HSPGs provide a mechanism for localizing macromolecular assemblies, such as lipoprotein particles associated with Wnts and Hedgehogs, near the target cell surface (Panáková et al., 2005; Neumann et al., 2009). For example, the extracellular Wnt concentration and signalling is reduced in *Drosophila* cells with genes encoding defective EXT glycosyltransferases, enzymes required for heparan sulfate (HS) biosynthesis (Takei et al., 2004). Mutations in the human ortholog EXT1 cause hereditary diseases characterised by short stature and benign bone tumours (Ahn et al., 1995). Moreover, injection of the heparin cleaving enzyme, heparinase, into wild-type *Drosophila* embryos causes the degradation of the heparin-like glycosaminoglycans (GAGs) and a *wingless*-like phenotype (Binari et al., 1997).

Based on the studies described above which loosely link the EGF-like domains of Shifted, HSPGs and morphogen signalling in *Drosophila*, I hypothesized that, in addition to contributing to Wnt binding, the EGF-like domains of human WIF-1 may play a further, distinctive role in Wnt signalling inhibition. The presence of eight evolutionarily conserved lysines and arginines in the sequence of EGF-like domain IV (**Figure 5.5**, **A**), when mapped onto the homology modelled structure of the domain (**Figure 5.5**, **B**), generates a highly positively charged surface (**Figure 5.5**, **C**). This suggests that human WIF-1 may bind to the highly-sulfated, negatively charged GAGs (**Figure 5.5**, **D**–**E**), typically present on the eukaryotic cell surface and in the extracellular matrix, via EGF-like domain IV. There are six glypicans (HSPGs, GPI-anchored to the cell membrane) and four syndecans (HSPGs

containing a single-pass transmembrane helix, type I transmembrane proteins) in mammals (Häcker et al., 2005). I hypothesised that WIF-1 may potentially bind via its positively charged EGF-like domain IV to negatively charged GAG chains covalently linked to these ten mammalian HSPGs.

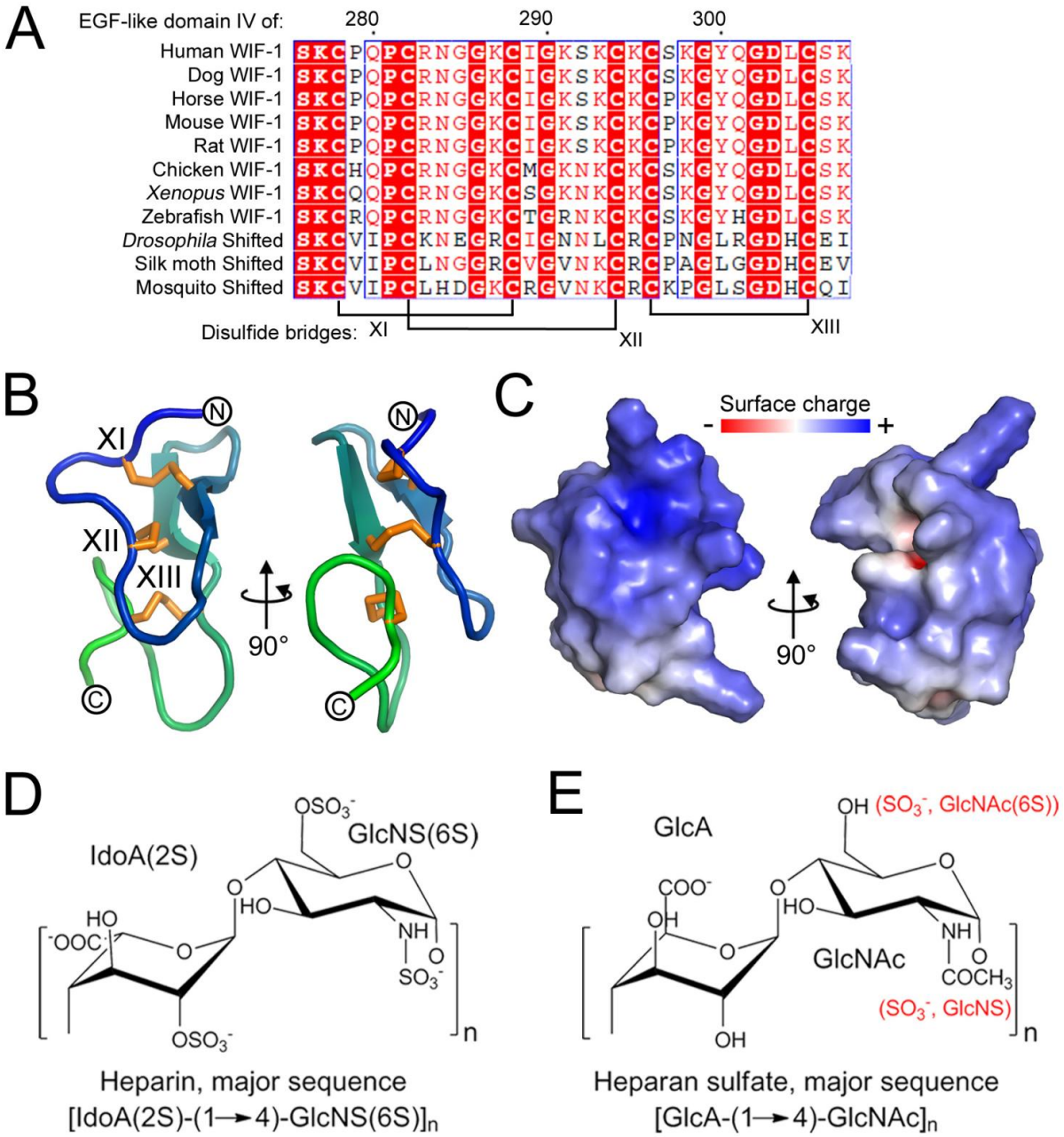


Figure 5.5. Charge complementarity between human WIF-1 EGF-like domain IV (positively charged) and glycosaminoglycans (negatively charged). **A**, amino acid sequence alignment of the WIF-1 EGF-like domains IV for mammals, chicken, *Xenopus* and zebrafish plus of the equivalent domain from sequences for the insect protein Shifted. Evolutionarily conserved residues are shown in a red background. Three disulfide bridges are indicated and marked with Roman numerals. Numbering corresponds to the full length human WIF-1 (including secretion signal), SwissProt identification accession number Q9Y5W5. **B**, an homology model of the human WIF-1 EGF-like domain IV is shown as a ribbon diagram. Disulfide bridges XI–XIII are shown as orange sticks. **C**, electrostatic properties of WIF-1 EGF-like domain IV model. The domain is shown as solvent-accessible surface coloured by electrostatic potential contoured at ± 8 kT/e (red, acidic; blue, basic). Chemical structures of major repeating disaccharide units of heparin (**D**) and heparan sulfate (**E**). **Figure 5.5** is an adapted version of Figure 4 and Supplementary Figure 8 from Malinauskas et al., 2011.

Physical interactions between WIF-$1_{\Delta C}$, WIF-$1_{WD\text{-}EGF\text{-}I}$ and four different GAGs were studied using the SPR binding assay. GAGs were biotinylated on their reducing terminus (as described in **Section 2.7.3**) and immobilized on the streptavidin-coated SPR sensor surface. Binding to purified WIF-1 constructs was measured. The tightest binding was observed between the most sulfated GAG, heparin (2.8 sulfate groups per disaccharide on average; Zhang et al., 2007), and WIF-$1_{\Delta C}$ (K_d 2.7 ± 1.5 μM, **Figure 5.6**, **A**). Weaker binding was detected between HS (1.2 sulfate groups per disaccharide; Zhang et al., 2007) and WIF-$1_{\Delta C}$ (K_d 8.0 ± 2.3 μM, **Figure 5.6**, **B**). WIF-$1_{\Delta C}$ did not bind to chondroitin sulfate and chondroitin 4-sulfate, the two least sulfated GAGs (Zhang et al., 2007). The correlation between higher GAG sulfation level and tighter binding to WIF-$1_{\Delta C}$ suggests the importance of electrostatic interactions between negatively charged GAG sulfate groups and positively charged EGF-like domain IV residues. Indeed, no binding was observed between WIF-$1_{WD\text{-}EGF\text{-}I}$ and any of the four GAGs, suggesting that GAG–WIF-$1_{\Delta C}$ interactions are mediated by EGF-like domains II–V. We used WIF-$1_{WD\text{-}EGFs\ I\text{–}III}$ to further dissect the contribution of the EGF-like domains in HS and heparin binding. SPR sensorgrams corresponding to GAG–WIF-$1_{WD\text{-}EGFs\ I\text{–}III}$ interactions showed qualitatively different binding characteristics compared to these measured for GAG–WIF-$1_{\Delta C}$ and could not be fitted to a simple 1:1 binding model to give a reliable estimate of K_d (**Figure 5.6**, **C**–**D**).

In agreement with the SPR binding assay, purified WIF-$1_{EGFs\ I\text{–}V}$ and WIF-$1_{\Delta C}$ (but not WIF-$1_{WD\text{-}EGF\text{-}I}$) constructs were able to bind to a heparin column under low-ionic strength conditions (10 mM Tris-HCl pH 8.0, 10 mM NaCl buffer; **Figure 5.6**, **E**). This binding was disrupted under higher-ionic strength conditions (250 mM and 300 mM NaCl, for WIF-$1_{\Delta C}$ and WIF-$1_{EGFs\ I\text{–}V}$ constructs, respectively). Observation of this binding mode is consistent with electrostatic forces playing a significant role in GAG–WIF-1 binding. The conserved cluster of lysines and arginines within the EGF-like domain IV seems to provide a suitable

surface for GAG recognition.

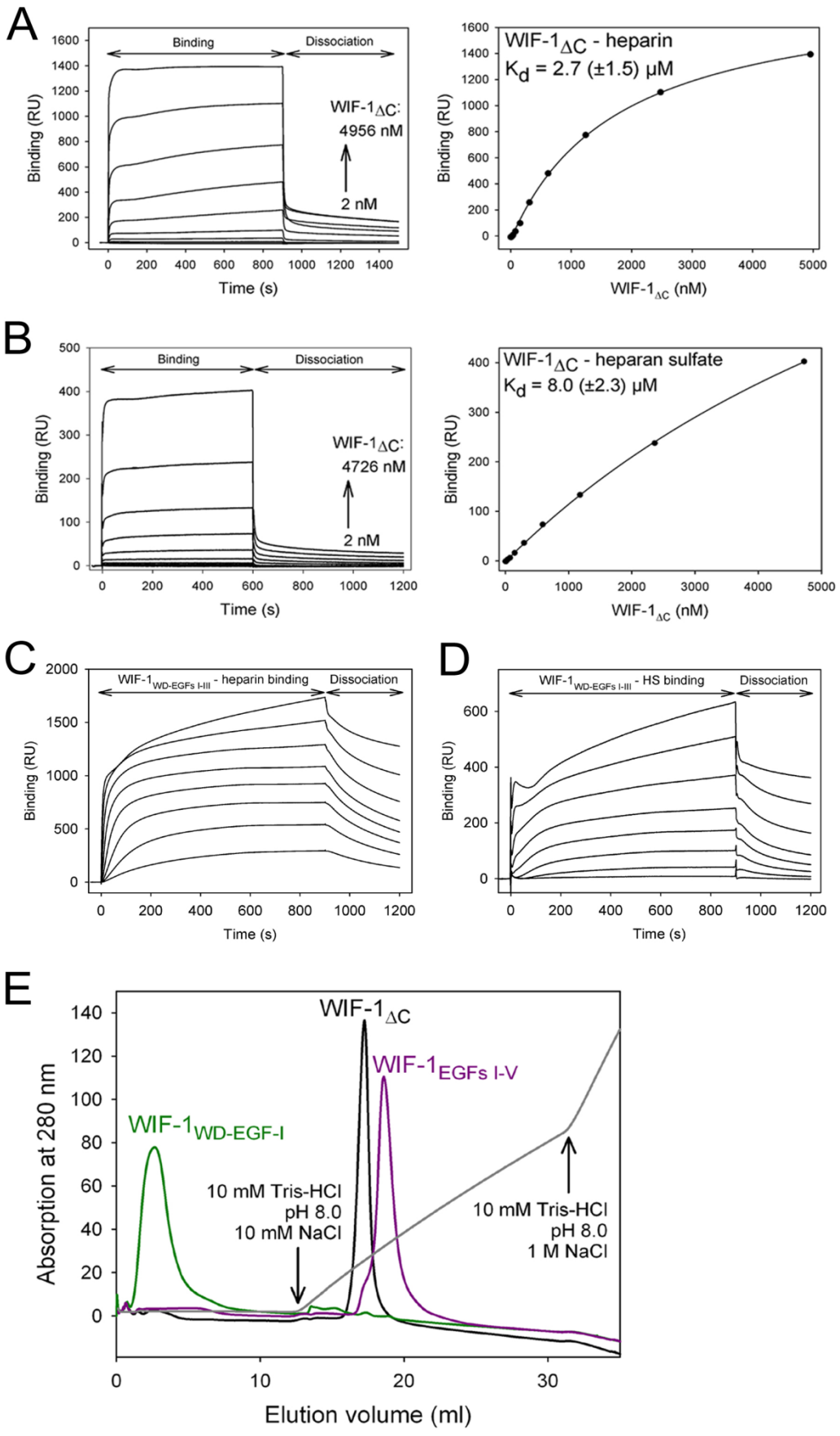

A
Binding
Dissociation
WIF-1ΔC:
4956 nM
2 nM
Binding (RU)
Time (s)
WIF-1ΔC - heparin
Kd = 2.7 (±1.5) μM
WIF-1ΔC (nM)
B
WIF-1ΔC:
4726 nM
WIF-1ΔC - heparan sulfate
Kd = 8.0 (±2.3) μM
C
WIF-1WD-EGFs I-III - heparin binding
D
WIF-1WD-EGFs I-III - HS binding
E
WIF-1ΔC
WIF-1EGFs I-V
WIF-1WD-EGF-I
10 mM Tris-HCl
pH 8.0
10 mM NaCl
10 mM Tris-HCl
pH 8.0
1 M NaCl
Absorption at 280 nm
Elution volume (ml)

Figure 5.6. Heparin- and heparan sulfate-binding properties of human WIF-1. **A**, surface plasmon resonance (SPR) binding sensorgrams of WIF-$1_{\Delta C}$ to heparin and representative plot of the binding response (response units (RU)) derived from the sensorgram as a function of WIF-$1_{\Delta C}$ concentration. **B**, SPR binding sensorgrams of WIF-$1_{\Delta C}$ to heparan sulfate and representative plot of the binding response (response units, RUs) derived from the sensorgram as a function of WIF-$1_{\Delta C}$ concentration. **C** and **D**, SPR binding sensorgrams of WIF-$1_{WD\text{-}EGFs\ I–III}$ to GAGs. SPR binding sensorgrams in which a concentration series (15.0 nM–1.92 μM, two-fold dilutions) of WIF-$1_{WD\text{-}EGFs\ I–III}$ were injected over immobilized heparin (**C**) or HS (**D**). These SPR sensorgrams could not be fitted to conventional 1:1 equilibrium or kinetic interaction models using BIAevaluation (Biacore Life Sciences) software, possibly indicating some contribution to binding from disruption of the terminal EGF-like domain III structure, consistent with the observed sensitivity to the exact C-terminal truncation point (**Figure 4.2**, **D**). **E**, binding of WIF-1 constructs to a heparin column. Three constructs of WIF-1 (WIF-$1_{WD\text{-}EGF\text{-}I}$ (green), WIF-$1_{EGFs\ I\text{-}V}$ (purple), and WIF-$1_{\Delta C}$ (black)) were dialysed against 10 mM Tris-HCl pH 8.0, 10 mM NaCl and loaded onto a pre-equilibrated 5 ml heparin column (HiTrap Heparin HP, GE Healthcare Life Sciences). Protein elution was followed by absorption at 280 nm. The elution gradient is represented by the grey curve. Arrows indicate the gradient range of NaCl (10–1000 mM). The WIF-$1_{WD\text{-}EGF\text{-}I}$ construct did not bind to the column and was eluted with the binding buffer (elution volume 1–5 ml). The WIF-$1_{EGFs\ I–V}$ and WIF-$1_{\Delta C}$ constructs were eluted with higher NaCl concentration (300 mM and 250 mM, respectively) indicating binding to heparin.

5.7 Endocytosis of WIF-1 constructs

Based on the observed WIF-1–GAG interactions (**Section 5.6**), we hypothesized that endogenously expressed HSPGs may be sufficient to localize WIF-1 on the cell surface. To test this hypothesis, we used purified WIF-1 constructs and three cell lines: human HEK-293T, Chinese hamster ovary (CHO) K1 (wild-type) and CHO-PgsA-745 cells. CHO-PgsA-745 cells are a CHO mutant cell line deficient in xylosyltransferase (UDP-D-xylose:serine-1,3-D-xylosyltransferase) (Esko et al., 1985). This cell line (American Type Culture Collection (ATCC) CRL-2242) was derived from CHO-K1 cells (ATCC CCL-61) treated with the mutagen ethylmethanesulfonate and screened for mutants defective in proteoglycan synthesis. The defect in xylosyltransferase, the enzyme responsible for the first sugar transfer in glycosaminoglycan (GAG) synthesis, means that CHO-PgsA-745 cells do not produce GAGs and fully glycosylated heparan sulfate proteoglycans (HSPGs).

The three cell lines were cultured separately and incubated with purified WIF-1 constructs. Following incubation, cells were washed with PBS buffer (to remove excess of the supplemental WIF-1 constructs), lysed and probed with anti-His antibodies during Western blot analysis (as described in **Section 2.12**) to check whether any of the WIF-1 constructs bind tightly to the cell surface (supposedly via HSPGs) and thus are present in the cell lysates. Assuming my hypothesis is correct, I would expect the following results. First, I would expect to detect binding between WIF-1 constructs, containing the HSPG-binding EGF-like domains II–V, and the HEK-293T and CHO-K1 cells, which presumably endogenously express HSPGs. As a consequence, these WIF-1 constructs should be present in HEK-293T and CHO-K1 cell lysates. Second, I would expect that WIF-1 constructs, lacking EGF-like domains II–V, do not bind to the cell surface and would be absent in HEK-293T and CHO-K1 cell lysates. Third, I would expect that none of the WIF-1 constructs bind to CHO-PgsA-745 cells that are not able to synthesize GAGs and thus glycosylated HSPGs.

However, I found that this hypothesis was not supported by my experimental results. First, WIF-1 constructs containing GAG-binding EGF-like domains II–V were not retained on the cell surface and were not detected in HEK-293T cell lysates (**Figure 5.7**, **A**), suggesting that either HEK-293T cells do not express fully glycosylated and sulfated HSPGs or that interactions between WIF-1 and HSPGs are not sufficiently strong to localize protein on the cell surface. Second, I found that WIF-$1_{\Delta C}$ construct, containing both WD and EGF-like domains, was detected in the cell lysates of CHO-K1 and CHO-PgsA-745 cell lysates, suggesting that either CHO cells do not express fully glycosylated and sulfated HSPGs in sufficient quantities to localise WIF-1 on the surface or that HSPGs may play a minor role in WIF-1 localisation (**Figure 5.7**, **B**). Interestingly, note that WIF-$1_{\Delta C}$ was endocytosed by CHO but not HEK-293T cells. Third, I found that WIF-1_{WD}, which lacks the EGF-like domains, was endocytosed by HEK-293T cells, suggesting that the overall domain architecture of WIF-1 may play a role in WIF-1 localisation and endocytosis (**Figure 5.7**, **A**). In addition, I tested whether overexpression of *Drosophila* HSPG Dally (Nakato et al., 1995; Tsuda et al., 1999) in HEK-293T cells could contribute to WIF-$1_{\Delta C}$ localisation on the cell surface. I detected no binding between WIF-$1_{\Delta C}$ and HEK-293T cells overexpressing Dally (data not shown).

To summarise, my preliminary cell-based experiments suggest that the domain architecture of WIF-1 may play a role in its localisation and endocytosis. However, further cell-based experiments, coupled to HSPG expression and sulfation level control, are required to understand the role of WIF-1–HSPG interactions and their implications for Wnt signalling inhibition.

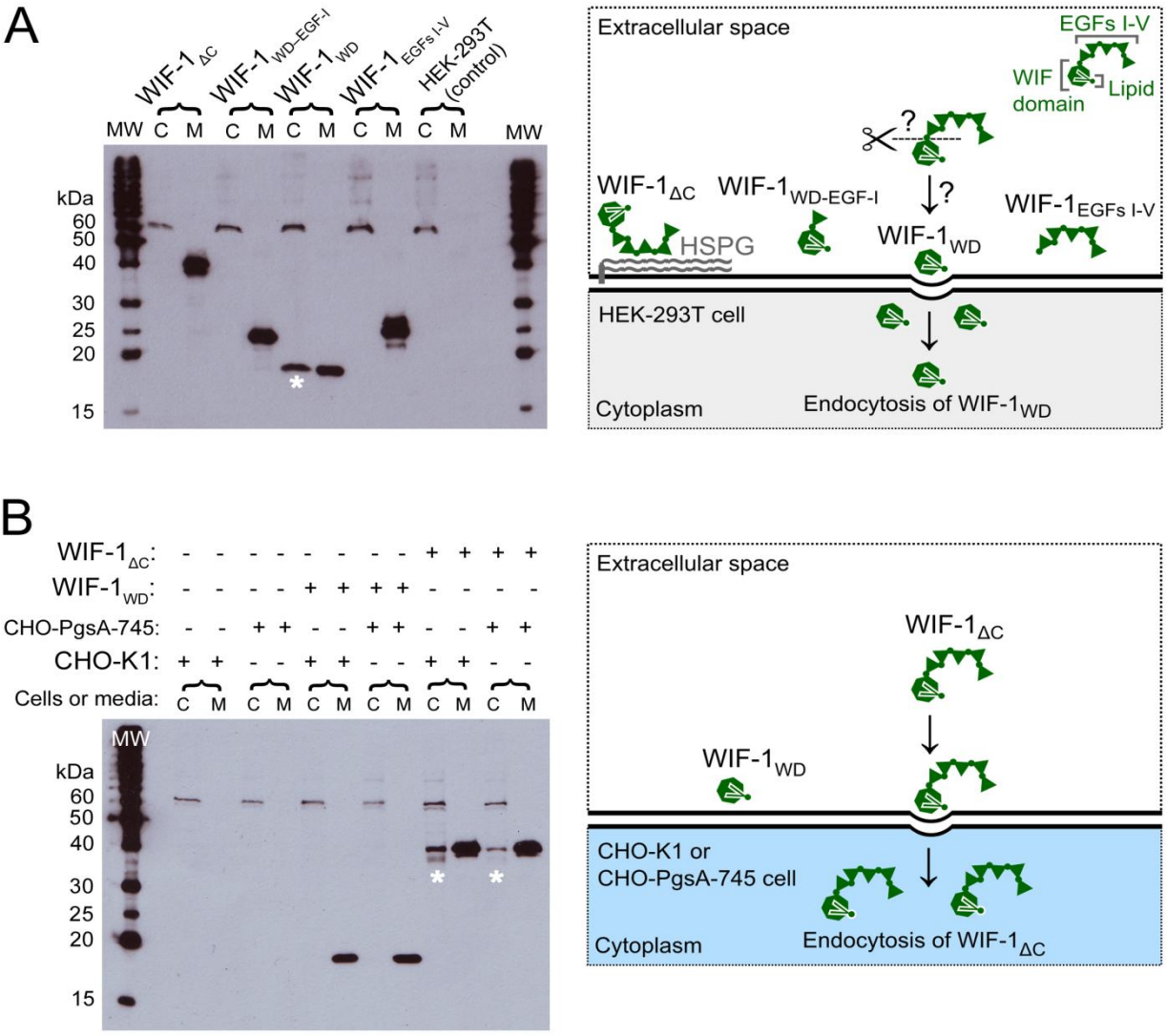


Figure 5.7. Endocytosis of human WIF-1 constructs. **A**, HEK-293T cells were seeded in 6-well plates and allowed to recover overnight in DMEM supplemented with L-Gln, non-essential amino acids and 10 % v/v FCS in a humidified 37 °C incubator with 5 % CO_2. Then the conditioned media was changed and supplemented with four purified human WIF-1 constructs (0.1 mg/well, i.e. 0.1 mg/2 ml media): WIF-$1_{\Delta C}$, WIF-$1_{WD\text{-}EGF\text{-}I}$, WIF-$1_{WD}$ and WIF-$1_{EGFs\ I–V}$. Cells were incubated for 6 hours and then washed three times with PBS buffer. The presence of the His_6-tagged WIF-1 constructs in the cells was checked by Western blot analysis as described in **Section 2.2.1**. Western blot lanes containing cell lysates and conditioned media are labelled with C and M, respectively. I have also checked whether WIF-1_{WD} was associated with the cell surface (rather than endocytosed) using anti-His antibodies linked to fluorescent molecules and light microscopy. I found no association between WIF-1_{WD} and the cell surface (data not shown), suggesting that WIF-1_{WD} was endocytosed (indicated with an asterisk *). Whether WIF-1 is proteolytically cleaved (indicated with scissors) in the extracellular space *in vivo* is not known. **B**, similarly to **A**, I have checked whether two WIF-1 constructs (WIF-$1_{\Delta C}$ and WIF-1_{WD}) are endocytosed by Chinese hamster ovary (CHO) K1 (wild-type) and CHO-PgsA-745 (GAG deficient) cells. Interestingly, in contrast to HEK-293T cells, WIF-1_{WD} was not (whereas WIF-$1_{\Delta C}$ was) endocytosed (indicated with an asterisk *) by CHO cells, suggesting that the endocytosis of WIF-1 constructs differs in distinct cell types.

5.8 Conclusions and future perspectives

Interactions between the evolutionarily conserved EGF-like domains of WIF-1 and HS provide a mechanism for vertebrate WIF-1 and *Drosophila* Shifted localisation within Wnt and Hedgehog morphogen gradients, respectively. Several Wnt signalling antagonists bind to heparin, an HS mimic. These include Sclerostin (Veverka et al., 2009), Dkk-1 (Fedi et al., 1999), and sFRP-1 (Finch et al., 1997). Similarly to WIF-1, sFRP-1 binds to Wnts and heparin via its N- and C-terminal domains, respectively (Bafico et al., 1999 and Üren et al., 2000, respectively). My structural and functional analyses on human WIF-1 illustrate how the interplay between Wnt- and HS-binding sites in Wnt signalling antagonists may link signal inhibition and morphogen gradients (**Figure 5.8**). Future functional and structural studies could potentially focus on the assembly of Wnt antagonist–HSPG–Wnt complexes to provide further insights into morphogen gradient control.

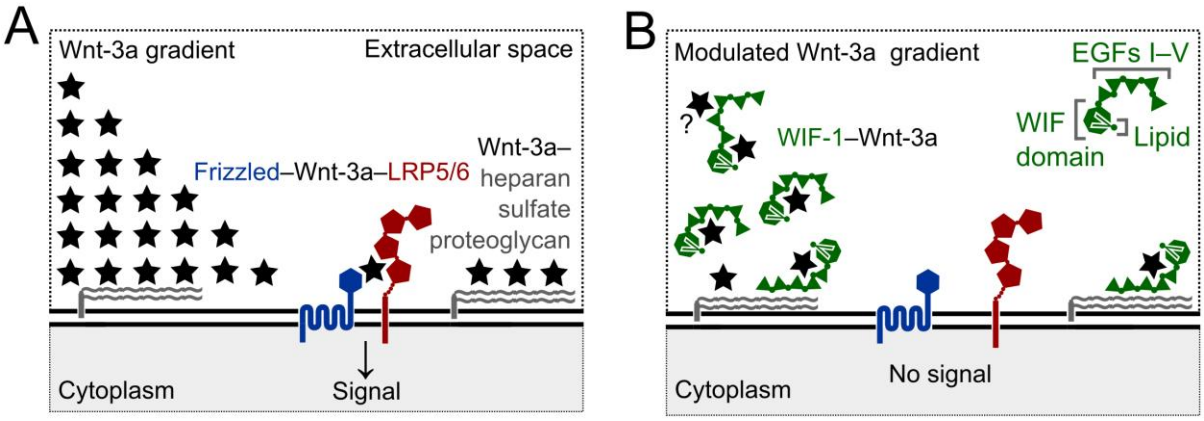


Figure 5.8. Modular mechanism of Wnt signalling inhibition by WIF-1. **A**, the Wnt/β-catenin pathway is activated at the cell surface when secreted Wnt-3a protein (drawn as a black star) forms a ternary complex with a seven-pass transmembrane receptor Frizzled (blue) and a single-pass transmembrane co-receptor LRP5/6 (dark red). This triggers downstream signal transduction and transcription of Wnt target genes. Heparan sulfate proteoglycans (HSPGs, grey) present on the cell surface have covalently linked heparan sulfate chains, bind to Wnt proteins and modulate their extracellular gradients. **B**, the combination of heparan sulfate- and Wnt-binding properties may provide a modular mechanism to control WIF-1 (green) localisation, and hence Wnt signalling inhibition, within the Wnt concentration gradient. **Figure 5.8** is an adapted version of Figure 5 from Malinauskas et al., 2011.

Chapter 6 – The cysteine-rich domain of the human secreted Frizzled-related protein 1

6.1 Summary

Secreted Frizzled-related proteins (sFRPs) are modulators of the Wnt/β-catenin signalling pathway. Members of the sFRP family consist of an N-terminal cysteine-rich domain (CRD), which is similar to the extracellular domains of members of the Wnt receptor family Frizzled, and a C-terminal netrin-like domain. sFRPs inhibit Wnt signalling by binding to Wnts and/or Frizzled receptors. However, sFRPs may also enhance Wnt signalling possibly, by acting as transporters of Wnts in the extracellular space. To start to dissect the contrasting mechanisms of action of sFRPs, we determined crystal structures of the human sFRP-1 CRD from four different crystal forms. Structural analyses of the sFRP-1 CRD crystal structures reveal a hydrophobic CRD homodimer interface and a positively charged CRD surface region which is located proximal to a putative Wnt-3-binding site. We show that the sFRP-1 CRD binds to mouse Wnt-3a and two negatively charged glycosaminoglycans, heparin and heparan sulfate. We further show that the sFRP-1 CRD binds weaker to Wnt-3a in the presence of heparin. This suggests that Wnt-3a and heparin may bind to the CRD via two partially overlapping sites and thus may compete with each other for CRD binding. The crystal structures, presented in **Chapter 6**, provide a structural basis for further studies of the CRDs, and the interplay between heparan sulfate proteoglycans and the CRDs in Wnt signal modulation.

6.2 Introduction

Seven-pass transmembrane proteins of the Frizzled family (Bhanot et al., 1996; Wang et al., 1996), together with co-receptors LRP5/6 (Wehrli et al., 2000; Tamai et al., 2000; Pinson et al., 2000), bind to Wnt morphogens and transduce signal across the cell membrane. It has

been shown that *Drosophila* cells transfected with the Frizzled-2 gene respond to added Wingless protein by increasing the level of Armadillo protein (*Drosophila* homolog of β-catenin) (Bhanot et al., 1996). It has been later shown that the purified extracellular domains of mouse Frizzled-8 and human LRP6 form a ternary complex mediated by human Wnt-3a (Bourhis et al., 2010). The cysteine-rich domain (CRD), present at the N-terminus of Frizzled receptors and members of the sFRP family, binds to Wnt proteins (Bhanot et al., 1996; Dann et al., 2001; **Figure 6.1**).

Following the characterization of Frizzled proteins as Wnt receptors, several secreted, CRD-containing, Frizzled-related proteins have been identified. A sFRP called FrzB was identified in protein fractions with chondrogenic (promoting development of the connective tissue) activity *in vivo* (Hoang et al., 1996). A *Xenopus* homologue of FrzB, expressed by the Spemann organiser (an organiser for gastrulation in amphibians) during early gastrulation, has been shown to co-immunoprecipitate with *Xenopus* Wnt-8, and mouse Wnt-1 as well as to inhibit Wnt signalling *in vivo* (Wang et al., 1997; Leyns et al., 1997). Concurrently, additional members of the sFRP family were identified via amino acid sequence homology with Frizzled receptors (Rattner et al., 1997) or as heparin-binding molecules (Finch et al., 1997). Surprisingly, a sFRP-1 derivative lacking the CRD retains ability to bind Wingless (Üren et al., 2000). Interestingly, sFRP-1 promotes Wingless signalling at low, but inhibits at high concentration in *Drosophila* cell-based assays (Üren et al., 2000). Moreover, sFRPs have been shown to interfere with BMP signalling (Yabe et al., 2003), act as axon-guidance cues (Rodriguez et al., 2005) and bind to the *r*eceptor *a*ctivator of *n*uclear factor *k*appa-B *l*igand (RANKL) (Häusler et al., 2004; Chuman et al., 2004). These diverse experimental observations remain poorly understood and prompted me to study the human sFRP-1 in detail.

The sFRP family members contain approximately 300 amino acid residues, harbouring an N-

terminal secretion signal peptide, followed by the CRD and a netrin-like domain (**Appendix 3**). As well as comprising the extracellular region of the transmembrane Frizzled receptors, the CRDs are also present in ROR family receptor tyrosine kinases, muscle-specific kinase (MuSK), collagen XVIII, Hedgehog signalling pathway activator Smoothened, and carboxypeptidase Z (**Figure 6.1**). sFRPs could be classified into two groups based on the sequence similarity of the Wnt-binding CRD: (1) sFRP-3 and -4, (2) sFRPs-1, -2 and -5. Classification based on the netrin-like domains yields the identical two sFRP groups.

Crystal structures of the mouse sFRP-3 and Frizzled-8 CRDs revealed a predominantly α helical fold with all 10 cysteines forming disulfide bonds (Dann et al., 2001). The sFRP-3 and Frizzled-8 CRDs form structurally homologous dimers with a highly complementary dimer interface. However, both purified CRDs are monomeric in aqueous solution at 100 μM protein concentration, suggesting a relatively weak affinity between two CRDs. Site-directed mutagenesis studies suggest a region of 950 $Å^2$ on the sFRP-3 and Frizzled-8 CRD surface as a putative *Xenopus* Wnt-8 binding site (Dann et al., 2001). This putative Wnt-8-binding surface partially overlaps with the interface of the CRD dimer observed in the crystals and suggests a mechanism in which CRD–CRD complex formation may interfere with Wnt binding (Dann et al., 2001). In contrast to the CRDs of Frizzled-8 and sFRP-3, the CRDs of MuSK form an asymmetric homodimer whose interface is also predominantly hydrophobic (Stiegler et al., 2009).

Given the speculation that the sFRP-1 contains a CRD which is more similar to non-Frizzled proteins (Semenov and He, 2000) and may potentially differ from already known CRD structures (Dann et al., 2001), we performed structural and functional analyses of this domain as presented in **Chapter 6**. Our crystal structures of the sFRP-1 CRDs from four different crystal forms (in total containing 19 CRD molecules in four asymmetric units) reveal a positively charged surface region of the domain, consistent with previously observed sFRP-1

interactions with a negatively charged heparin polysaccharide (Finch et al., 1997). This suggests that the sFRP-1 CRD could potentially interact with heparan sulfate proteoglycans (HSPGs), proteins typically present on the cell surface and implicated in the regulation of Wnt signalling (discussed in **Section 5.2**). Surface plasmon resonance studies of the sFRP-1 CRD binding to heparin, heparan sulfate, HSPGs and Wnt-3a are also presented in **Chapter 6**, and implications for the regulation of Wnt signalling are discussed.

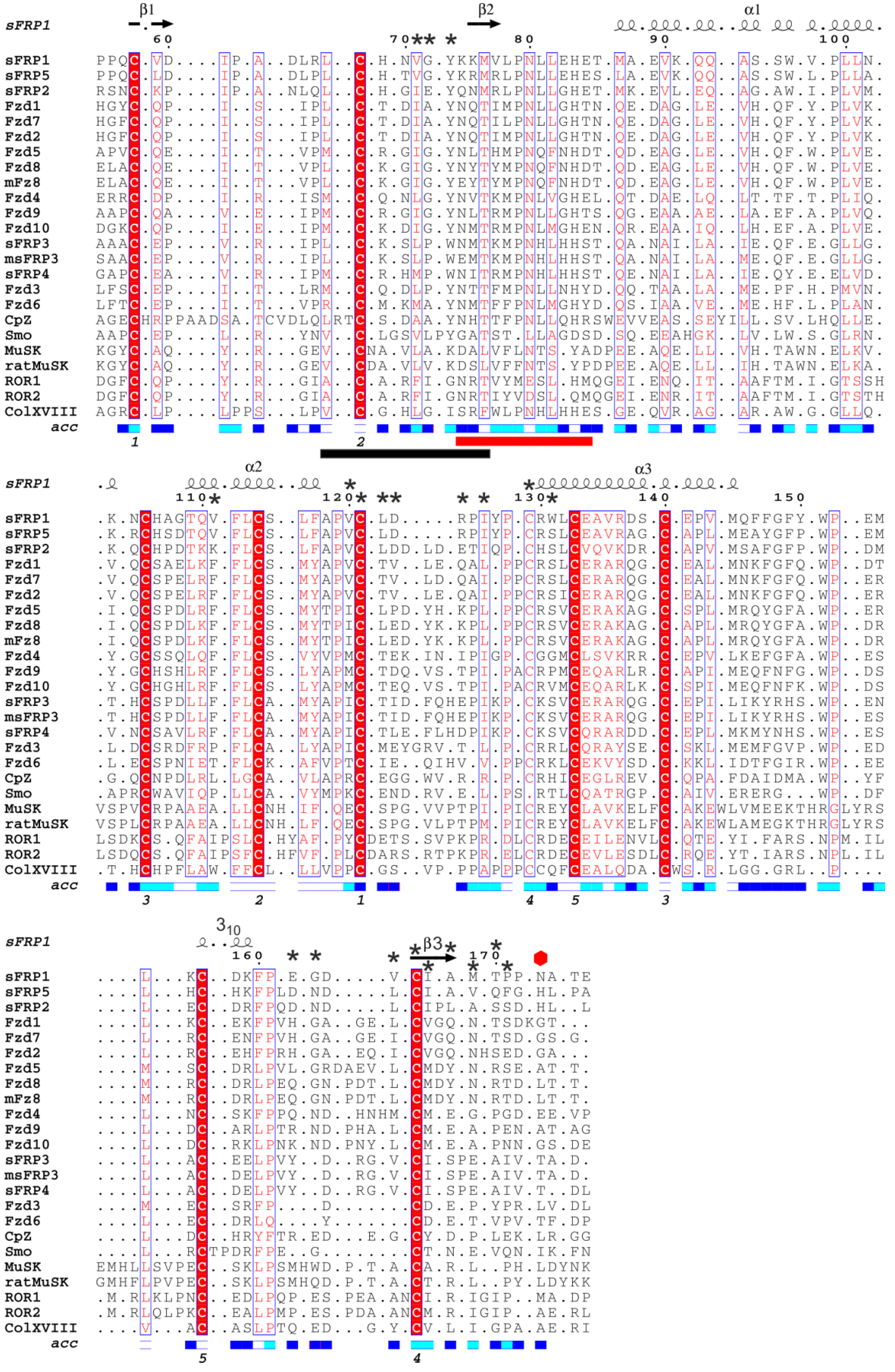

sFRP1
β1
β2
α1
sFRP1
sFRP5
sFRP2
Fzd1
Fzd7
Fzd2
Fzd5
Fzd8
mFz8
Fzd4
Fzd9
Fzd10
sFRP3
msFRP3
sFRP4
Fzd3
Fzd6
CpZ
Smo
MuSK
ratMuSK
ROR1
ROR2
ColXVIII
acc
α2
α3
β3
3₁₀

Figure 6.1. Amino acid sequence alignment of all human Frizzled (abbreviated as Fzd), secreted Frizzled-related protein (sFRP) cysteine-rich/Wnt-binding domains (CRD) with selected CRDs. Alignment of human sFRP-1 to -5 (UniProtKB/Swiss-Prot identification accession numbers Q8N474, Q96HF1, Q92765, Q6FHJ7 and Q5T4F7, respectively), Frizzled-1 to -10 receptors (Q9UP38, Q14332, Q9NPG1, Q9ULV1, Q13467, O60353, O75084, Q9H461, O00144 and Q9ULW2, respectively), mouse Frizzled-8 (abbreviated as mFz8, Q61091; PDB ID code 1IJY; Dann et al., 2001), mouse sFRP-3 (msFRP3, P97401; PDB ID code 1IJX; Dann et al., 2001), human carboxypeptidase Z (Q66K79), human Smoothened homolog (also known as SMO, protein Gx, Q99835), human muscle-specific tyrosine-protein kinase receptor (MuSK, O15146), rat MuSK (Q62838; PDB ID code 3HKL; Stiegler et al., 2009), human tyrosine-protein kinase transmembrane receptors ROR1 and ROR2 (Q01973 and Q01974) and human collagen α-1 (XVIII) chain (P39060). Numbering corresponds to the full length human sFRP-1. Evolutionarily conserved residues are shown in a red background. The sequences were aligned using MultAlin (Corpet, 1988; http://multalin.toulouse.inra.fr/multalin/) and formatted with ESPript (Gouet et al., 1999; http://espript.ibcp.fr/ESPript/ESPript/). Secondary structure elements of the sFRP-1 CRD crystal structure were assigned with ESPript. Disulfide bonds 1 to 5 and the solvent accessibilities (calculated with ESPript) of the sFRP-1 residues are presented below the alignments (blue, accessible; cyan, intermediate; white, buried). An Asn-linked N-acetylglucosamine is marked by a filled red hexagon above glycosylated Asn173. Amino acid residues implicated in *Drosophila* Frizzled-2 CRD– and mouse Frizzled-6 CRD–*Xenopus* Wnt-8 binding are highlighted with a black bar below the sequence alignments (Dann et al., 2001). Amino acid residues implicated in the sFRP-1 antagonism of Wnt-3 signalling are highlighted with a red bar below the sequence alignments (Bhat et al., 2007). Residues whose side chains are buried in the dimer interface are marked with asterisks (*).

6.3 Aims

The main aim is to provide structural insights into the Wnt-3a–sFRP-1 CRD complex assembly and the role of glycosaminoglycans in these interactions, potentially leading to a better understanding of sFRP- and HSPG-modulated Wnt signal transduction.

6.4 Crystallisation and structure determination of the sFRP-1 CRD

To investigate the mechanism of Wnt signalling modulation by sFRP-1, constructs of the human sFRP-1 cysteine-rich domain (CRD) were cloned into the pHLsec expression vector (Aricescu et al., 2006) as described in **Section 2.1**. A number of sFRP-1 constructs were not secreted by HEK-293T cells. These include the full length sFRP-1, amino acid residues Ser32–Lys314, and the sFRP-1 CRD construct Ser32–Ala174, lacking Thr175, an essential Thr for the glycosylation of the preceding Asn173 (data not shown). The sFRP-1 CRD constructs Ser32–Glu176 and Tyr52–Glu176 were secreted by HEK-293T and -293S cells and purified (**Figure 6.2** and **Section 2.2.2**). Prior to crystallisation, the sFRP-1 CRD was enzymatically deglycosylated (Aricescu et al., 2006) and concentrated by ultrafiltration in SEC buffer (typically, 150–300 mM NaCl, 10–20 mM Tris-HCl pH 8.0) to 4.1–5.6 mg/ml (**Table 6.1**). Sitting drop vapour diffusion crystallisation experiments were set up as described in **Section 2.3**. The sFRP-1 CRD crystallised in several different sulfate-containing conditions (**Table 6.1**).

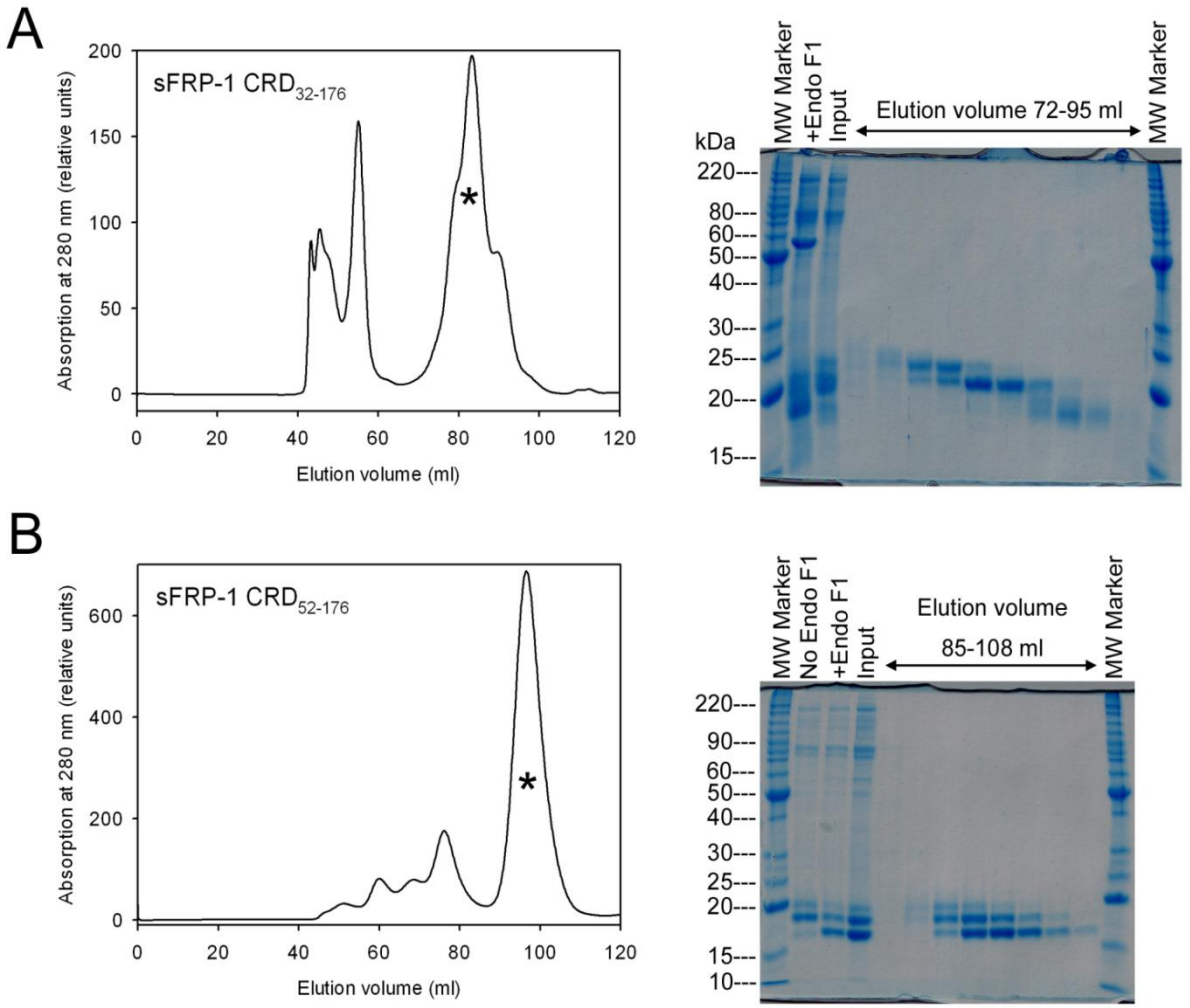
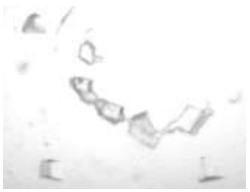
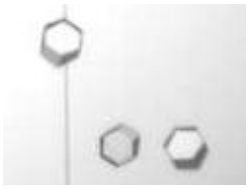


Figure 6.2. Size-exclusion chromatography of the human sFRP-1 cysteine-rich domain constructs: sFRP-1 $CRD_{32–176}$ (**A**) and sFRP-1 $CRD_{52–176}$ (**B**). Separation by size exclusion chromatography was performed using a HiLoad 16/60 Superdex 75 column (GE Healthcare Life Sciences) at a flow speed of 0.8 ml/min in 20 mM Tris-HCl pH 8.0, 300 mM NaCl. Protein elution was monitored by absorbance at 280 nm (black curves). The sFRP-1 CRD is present in the peak marked with an asterisk (*). The sFRP-1 CRD peak fractions were analyzed on reducing SDS-polyacrylamide (15 % w/v) gels (shown on the right side). For crystallisation trials, sFRP-1 CRD was deglycosylated with endoglycosidase F1 (abbreviated as "Endo F1" on the right). Molecular weight (MW) marker: BenchMark Protein Ladder, Invitrogen. Gels were stained with SimplyBlue SafeStain, Invitrogen.

Table 6.1. Crystallisation of the human sFRP-1 cysteine-rich domain. Crystal numbering (I to III) corresponds to **Table 6.2**.

No.	Construct boundaries	Crystal number	Conc. (mg/ml)	Crystallisation conditions at 21 °C	Crystal size (µm), in the longest dimension	Cryoprotectant	Crystal
1.	Ser32–Glu176		4.1	100 nl 4.1 mg/ml protein in 300 mM NaCl, 20 mM Tris-HCl pH 8.0 plus 100 nl 200 mM NaCl, 100 mM cacodylate pH 6.5, 2 M $(NH_4)_2SO_4$	60	30 % glycerol	
2.	Tyr52–Glu176	Crystal I	5.6	100 nl 4.1 mg/ml protein in 300 mM NaCl, 20 mM Tris-HCl pH 8.0 plus 100 nl 200 mM NaCl, 100 mM cacodylate pH 6.5, 2 M $(NH_4)_2SO_4$, 2.5 % 1-propanol	80	30 % glycerol	
3.	Tyr52–Glu176	Crystal II	5.2	100 nl 5.2 mg/ml protein in 150 mM NaCl, 10 mM Tris-HCl pH 8.0, 10 mg/ml sucrose octasulfate plus 100 nl 2 M $(NH_4)_2SO_4$, ~2–5 % 1-propanol	160	Perfluoropolyether oil PFO-X125/03 (Lancaster Synthesis)	
4.	Tyr52–Glu176	Crystal III	5.2	100 nl 5.2 mg/ml protein in 150 mM NaCl, 10 mM Tris-HCl pH 8.0, 10 mg/ml sucrose octasulfate plus 100 nl 10 mM $MgSO_4$, 1.8 M Li_2SO_4, 50 mM cacodylate pH 6.0	50	25 % glycerol	

Diffraction data sets were collected from four single crystals cryoprotected using crystallisation solution supplemented with cryoprotectants (**Table 6.1**) at 100 K at the Diamond Light Source and ESRF synchrotrons (**Figure 6.3**). Statistical descriptors of the X-ray diffraction data sets are presented in **Table 6.2**. The crystal structures were phased by molecular replacement using a homology model of the sFRP-1 CRD (based on the mouse sFRP-3 CRD crystal structure (Dann et al., 2001; PDB identification code 1IJX)) and program Phaser (McCoy et al., 2007). Structures were refined using iterative cycles of manual rebuilding in program Coot (Emsley et al., 2010) and automatic refinement in autoBUSTER (Blanc et al., 2004). As the sFRP-1 CRD crystals diffracted to a moderate resolution (3.1–3.8 Å) but contain multiple molecules in the asymmetric unit (3–6), non-crystallographic symmetry (NCS) restraints were used during structure refinement. The refined sFRP-1 CRD atomic coordinates have good values for crystallographic validation parameters compared to other low resolution structures (among the top 3 % best refined structures of 3 – 9999 Å resolution range as evaluated with the program MolProbity (Davis et al., 2007)) (**Table 6.2**).

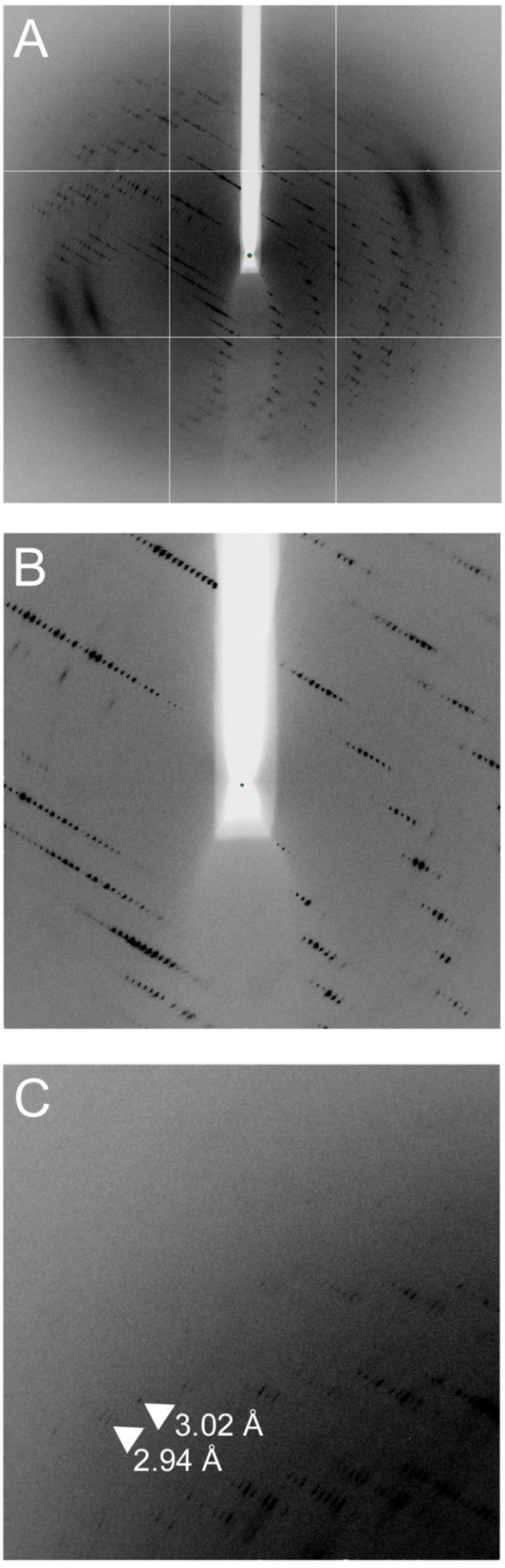


Figure 6.3. The sFRP-1 $CRD_{52–176}$ crystal shown in **Table 6.1** (number II) was frozen in the crystallisation solution (200 mM NaCl, 100 mM cacodylate pH 6.5, 2 M $(NH_4)_2SO_4$, 2.5 % 1-propanol) supplemented with glycerol (30 % final concentration) and diffraction data were collected as described in **Section 2.4**. Although the crystal diffraction quality was suboptimal (as indicated by elongated shape of the X-ray reflections (**A**–**C**) and their relatively poor separation (visible in **B**)) and the three dimensional unit cell was relatively large in one dimension (c=307.02 Å), the diffraction data were successfully processed and used for structure determination. High resolution (up to 2.94 Å) reflections are indicated with white arrows in **C**. White rectangular regions in **A** and **B** correspond to the beamstop which is impenetrable to X-rays and was used to protect the X-ray detector. The crystal was exposed over a 1° oscillation range. The edge of the detector corresponds to a resolution of ~2.9 Å. Crystallographic data collection and refinement statistics for the sFRP-1 CRDs are shown in **Table 6.2**.

Table 6.2. Crystallographic data collection and refinement statistics for the sFRP-1 CRD crystals. Root-mean-square deviation (RMSD) values of pairwise comparison of the 19 sFRP-1 CRD structures are presented in **Table 6.3**.

Construct	**sFRP-1 $CRD_{32–176}$**	**sFRP-1 $CRD_{52–176}$ Crystal I**	**sFRP-1 $CRD_{52–176}$ Crystal II**	**sFRP-1 $CRD_{52–176}$ Crystal III**
Data collection				
X-ray source	Diamond Light Source, I03	Diamond Light Source, I04	ESRF, ID23-1	ESRF, ID23-1
Resolution range (Å)	39.66–3.60 (3.80–3.60)	57.85–3.10 (3.27–3.10)	47.74–3.70 (3.90–3.70)	61.53–3.80 (4.01–3.80)
Space group	$P4_32_12$	$P6_322$	P3	$P6_1$
Cell dimensions	a=81.13 Å, b=81.13 Å, c=189.11 Å, α=90°, β=90°, γ=90°	a=72.12 Å, b=72.12 Å, c=307.02 Å, α=90°, β=90°, γ=120°	a=72.31 Å, b=72.31 Å, c=147.57 Å, α=90°, β=90°, γ=120°	a=142.09 Å, b=142.09 Å, c=118.03 Å, α=90°, β=90°, γ=120°
Molecules per asymmetric unit	4	3	6	6
Dimers in the asymmetric unit	0	1	2	1
Wavelength (Å)	0.9763	0.9763	0.9801	0.9801
R_{merge} (%)	24.9 (76.6)	11.9 (86.2)	16.1 (78.9)	22.3 (90.2)
Total number of observations	33368 (5012)	54069 (8551)	24726 (3576)	73227 (10832)
Mean ((I)/σ(I))	5.6 (1.8)	8.3 (1.8)	4.1 (1.4)	6.7 (2.0)
Completeness	98.9 (99.8)	98.1 (99.1)	99.7 (99.2)	100.0 (100.0)
Multiplicity	4.3 (4.5)	5.9 (6.5)	2.7 (2.6)	5.4 (5.5)
Refinement				
Resolution range (Å)	39.66–3.60 (4.03–3.60)	57.85–3.10 (3.47–3.10)	47.74–3.70 (4.14–3.70)	61.53–3.80 (4.10–3.80)
Number of reflections	7735 (2163)	9124 (2527)	9174 (2584)	13452 (2778)
R_{work} (%)	28.7 (27.6)	29.60 (30.20)	27.03 (27.84)	25.00 (26.53)
R_{free} (%)	30.6 (26.9)	31.62 (29.57)	29.43 (30.00)	25.40 (29.98)
Number of atoms; number of non-protein atoms	3692; 0	2755; 0	5524; 0	5539; 14
Wilson B factor ($Å^2$)	48	97	93	68
Average B factor ($Å^2$)	138	123	172	89
RMSD bonds (Å)	0.008	0.008	0.007	0.007
RMSD angles (degrees)	0.91	0.93	0.97	0.98
Ramachandran statistics				
Favoured (%)	95.18	93.24	92.67	95.44
Number of outliers (total residues per AU)	0 (456)	3 (340)	1 (682)	1 (680)

6.5 The sFRP-1 CRD forms homodimers in the crystals

Here I report the crystal structures of the human sFRP-1 CRD from four different crystal forms, 19 molecules in total, at 3.1–3.8 Å resolution (**Table 6.2**). The crystal structures reveal that the sFRP-1 CRD adopts predominantly an α helical fold previously observed for the CRDs of mouse Frizzled-8 (RMSD of 1.4–1.6 Å for 103 aligned residues; sFRP-1 $CRD_{52–176}$, crystal II, chain F was used for alignment with the program PDBeFold; Krissinel and Henrick, 2004), secreted Frizzled-related protein 3 (sFRP-3) (RMSD of 1.3–1.5 Å for 100 aligned residues) and muscle-specific receptor tyrosine kinase (MuSK) (RMSD of 2.9 Å for 90 aligned residues), highlighting evolutionary conservation of the domain architecture (**Figure 6.4**). All ten cysteines form five disulfide bonds (Cys58–Cys121, Cys68–Cys114, Cys105–Cys140, Cys129–Cys166 and Cys133–Cys157) that stabilize the domain architecture. The sFRP-1 CRD fold consists of tightly packed secondary structure elements: two short N-terminal antiparallel β strands 1 and 2 (residues Cys58–Asp60 and Lys75–Val77, respectively), three α helices (9 to 17 residues each), one short 3_{10} helix (residues Cys157–Lys159) plus a C-terminal β3 which contacts β3 from a neighbouring CRD molecule (**Figures 6.1** and **6.4**) in all four different crystal forms (**Table 6.2**).

The sFRP-1 CRDs are arranged as homodimers within all four crystal forms (**Figure 6.5**). Eight of them form four non-crystallographic dimers, the remaining eleven form similarly packed crystallographic dimers arranged along the two-fold symmetry axes (**Table 6.3**). ~1000 $Å^2$ solvent-accessible area (14 % of the total CRD solvent-accessible surface area) is buried at the complementary homodimer interface. The extensive interface area (typically observed in the biologically important protein–protein interactions; Jones and Thornton, 1996; Moreira et al., 2007) suggests that CRDs could potentially form dimers *in vitro* and *in vivo*. Indeed, it has been shown that CRDs of sFRPs and Frizzled receptors may form

homodimers and heterodimers. It has been shown that the full length sFRP-3 (also known as FRP), or the isolated CRD of sFRP-3, binds to human Frizzled-6 (Bafico et al., 1999). In addition, sFRP-3 can form homo complexes and its CRD is sufficient to mediate these interactions (Bafico et al., 1999). It has also been shown that dimerisation of *Xenopus* Frizzled-3 is sufficient to activate the canonical Wnt/β-catenin pathway and that the CRD of Frizzled-3 is sufficient for dimerisation (Carron et al., 2003). Moreover, sFRP-1 regulates the growth of retinal ganglion cell axons through the Frizzled-2 receptor, and the sFRP-1 activity is mediated by its CRD (Rodriguez et al., 2005). However, in my hands the sFRP-1 CRD is monomeric in aqueous solution at 2.3 mg/ml (147 μM) as measured by multi-angle laser light scattering and analytical ultracentrifugation analyses (**Figure 6.5**, **C** and **D**). This suggests that the sFRP-1 CRD–CRD binding affinity is low (147 μM < K_d < ~50 mM, concentration of the CRDs in the crystals) and could be potentially enhanced by unknown molecules *in vivo*.

About 20 amino acid residues whose side chains are exposed at the CRD surface are located at the dimer interface (**Figures 6.1** and **6.5**). Most of these residues are contributed from three amino acid sequence regions of the CRD: Val71–Tyr73, Val120–Trp131 and Glu162–Pro171 (**Figure 6.1**). Evolutionary conservation of these three regions is very limited among the CRDs, suggesting that dimerisation might be a unique property of a subset of Frizzled and sFRP CRDs (**Figure 6.1**). E.g. sFRP-1 Leu122, which is buried at the dimer interface (**Figure 6.5**, **A**), is conserved between sFRP-1, -2, -5 and Frizzled-5, -8, suggesting a possible homo or hetero dimerisation between these proteins. In contrast, a corresponding residue in Smoothened and carboxypeptidase Z is glutamate, which excludes a possibility of an identical hydrophobic surface-mediated dimerisation of these proteins.

To summarise, the observed human sFRP-1 CRD homodimers in the crystals, together with previously observed mouse sFRP-3 and Frizzled 8 CRD homodimers (Dann et al., 2001),

provide a structural basis for further structure-guided, site-directed mutagenesis studies to elucidate the role of the CRD dimerisation in Wnt signalling pathway.

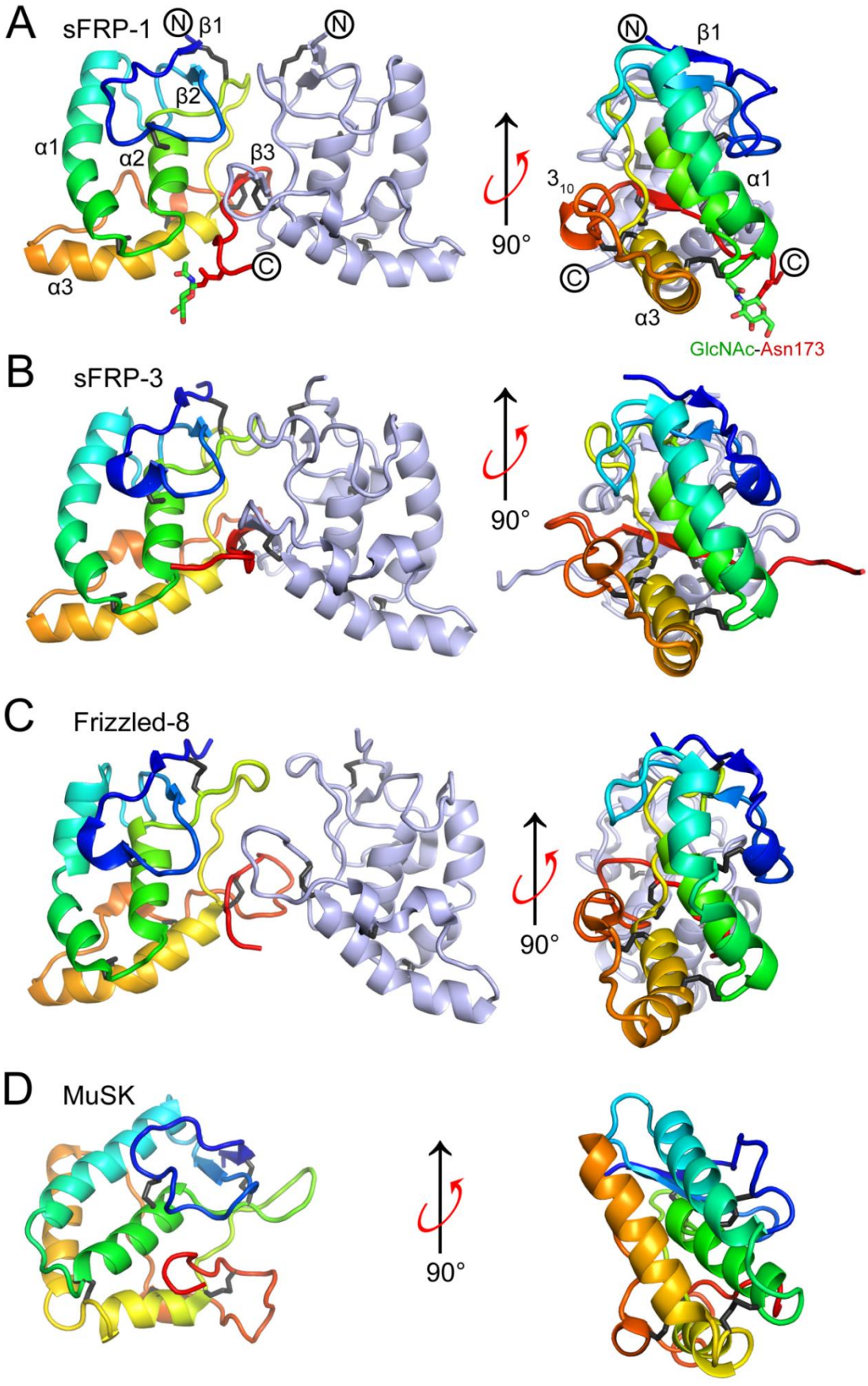
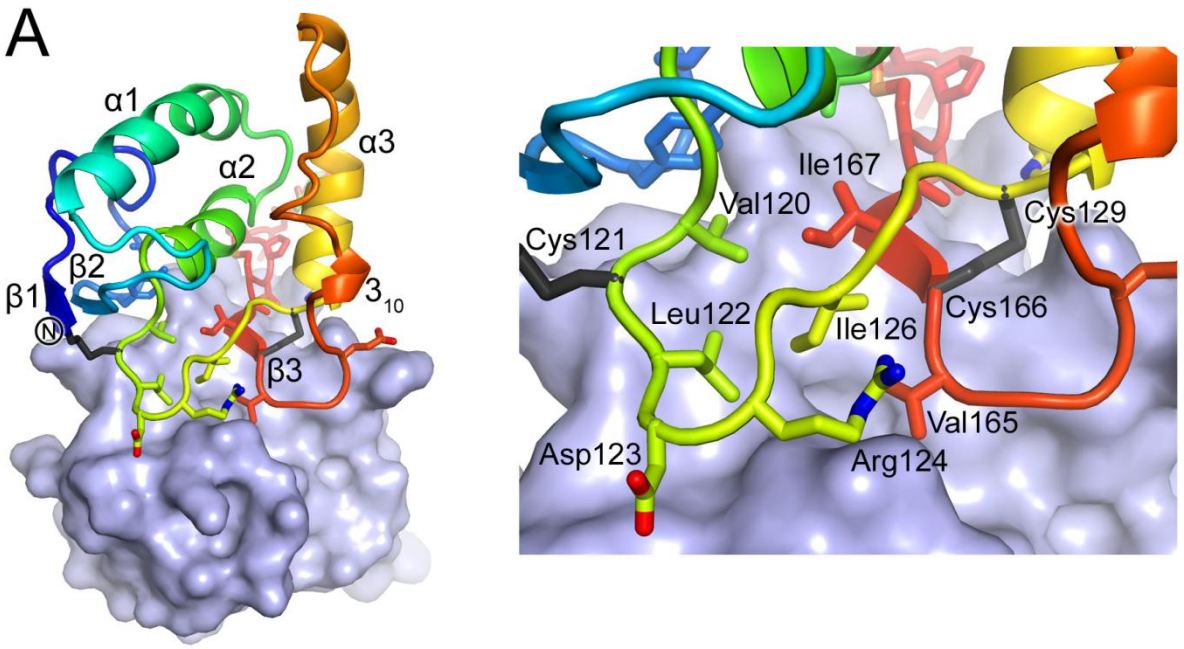
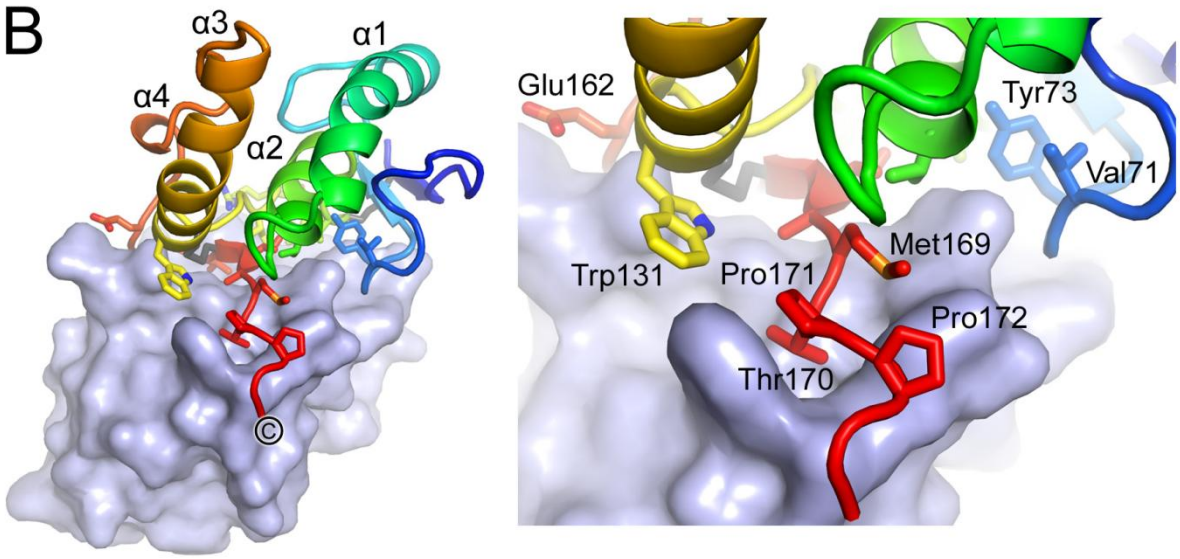

A
sFRP-1
N
β1
β2
α1
α2
β3
α3
C
90°
3₁₀
GlcNAc-Asn173
B
sFRP-3
C
Frizzled-8
D
MuSK

Figure 6.4. Comparison of the dimeric human sFRP-1 $CRD_{52–176}$ crystal III structure (**A**) (chains A and B, **Tables 6.1**–**6.3**) with the mouse sFRP-1 CRD dimer (**B**) (PDB ID 1IJX, chains A and C), mouse Frizzled-8 CRD dimer (**C**) (PDB ID 1IJY, chains A and B), and mouse muscle-specific tyrosine kinase (MuSK) CRD monomer (**D**) (PDB ID 3HKL, chain A). One CRD of the dimer is colour coded as a rainbow with N- and C-termini shown in blue and red, respectively; the second CRD is coloured in light blue/grey. Disulfide bonds are shown as grey sticks. Asn173 (red) linked N-acetyl-D-glucosamine is shown as sticks (C: green; O: red; N: blue). Two views of each structure that differ by a 90° rotation about a vertical axis are presented. α helices and β strands are numbered in **A**, and correspond to **Figure 6.1**.


A
α1
α2
α3
β1
β2
β3
3₁₀
N
Ile167
Val120
Cys129
Cys121
Leu122
Ile126
Cys166
Val165
Asp123
Arg124


B
α3
α1
α4
α2
C
Glu162
Tyr73
Val71
Met169
Trp131
Pro171
Pro172
Thr170

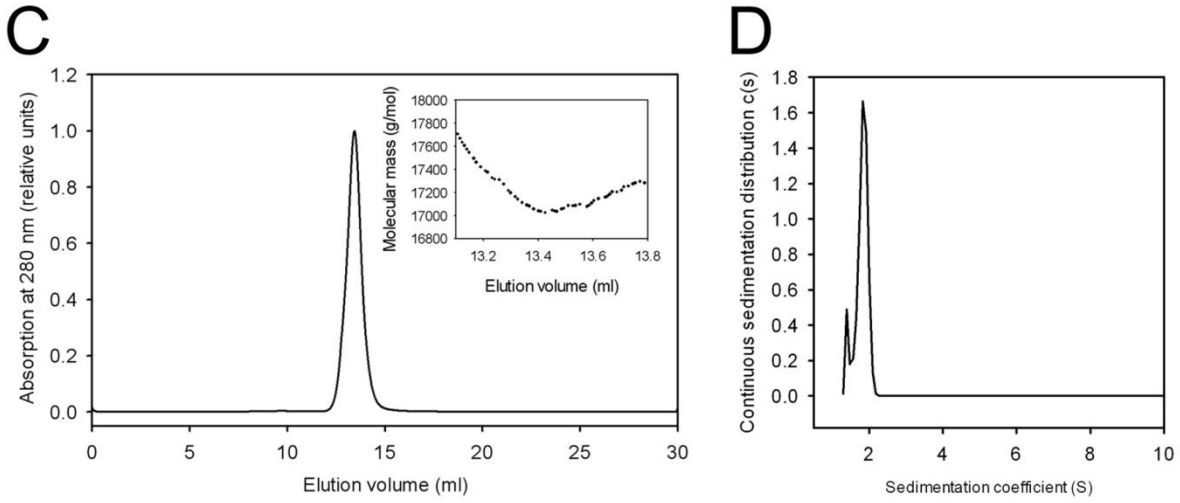

C
Absorption at 280 nm (relative units)
Elution volume (ml)
Molecular mass (g/mol)
Elution volume (ml)
D
Continuous sedimentation distribution c(s)
Sedimentation coefficient (S)

Figure 6.5. Amino acid residues forming the sFRP-1 CRD–CRD homodimer interaction surface. The same molecules as in **Figure 6.4**, **A** are shown. Secondary structure elements and amino acid residues are indicated, and correspond to the amino acid sequence alignment of CRDs in **Figure 6.1**. One CRD of the dimer is colour coded as a rainbow with N- and C-termini shown in blue (best visible in **A**) and red (best visible in **B**), respectively. The solvent-protein contact surface of the second CRD is coloured in light blue/grey. Disulfide bonds are shown as grey sticks. Asn173-linked GlcNAc is not shown for simplicity. **C**, multi-angle laser light scattering (MALLS) revealed that the sFRP-1 $CRD_{52–176}$ is monomeric in solution at 2.3 mg/ml (147 μM). An inset shows a distribution of the calculated molar mass, which is in a good agreement with SDS-PAGE analysis (**Figure 6.2**, **B**), of the CRD. Glycosylation of the protein was not taken into account during MALLS data analysis. **D**, sedimentation coefficient distribution of sFRP-1 $CRD_{52–176}$ determined by analytical ultracentrifugation (AUC) at 0.6 mg/ml. No dimers of the sFRP-1 $CRD_{52–176}$ were detected during MALLS (**C**), AUC (**D**) and SEC (**Figure 6.2**) experiments.

6.6 Structural characteristics of the sFRP-1 CRD

Analyses of nineteen sFRP-1 CRD molecules (**Figure 6.6**, **A**) from four different crystal forms (**Tables 6.2** and **6.3**) revealed the following four notable features of the structure.

First, electron density corresponding to the position of N-terminal residues Glu52–Pro56 was fragmentary, consistent with their potential conformational flexibility, thus these residues were omitted from the final refined sFRP-1 CRD atomic coordinates. The first N-terminal residue with a distinctive electron density was Cys58 forming a disulfide bond with Cys121, consistent with a contribution of five disulfide bonds to the restriction of the conformational heterogeneity of the CRD (**Table 6.3**).

Second, the C-terminal residues Pro171–glycosylated Asn173 adopt different conformations depending on the crystal packing and do not form distinctive hydrophobic/hydrophilic interactions with any regions of the CRD, suggesting that their conformation can change depending on the minor external forces (**Figure 6.6**, **A**). In the full length human sFRP-1, the CRD is followed by a predicted linker of ~10 amino acid residues (residues Glu176–Val185; residue Thr175 is defined here as the C-terminal residue of the CRD as it is required for the glycosylation of Asn173; the sFRP-1 CRD construct lacking this glycosylation site was not secreted by HEK-293T cells (data not shown), suggesting its possible role in sFRP-1 folding and/or trafficking inside the cell) and a netrin-like domain. Multiple conformations of the C-terminal residues and a distinctive linker between the two domains suggest multiple possibilities for how the CRD and the netrin-like domain could be arranged in relation to each other. Whether the two domains are closely packed, remains unknown. However, it has been shown that the netrin-like domain contributes to Wnt signalling inhibition (Üren et al., 2000; Lopez-Rios et al., 2008), suggesting a possibility where both the CRD and the netrin-like domain are closely juxtaposed and form an extended Wnt-binding surface. Whether that is true, and whether the CRD and the netrin-like domain interact with each other remains to

be tested experimentally.

Third, a highly hydrophobic region, encompassing residues Phe146-PheGlyPhe-Tyr150, is exposed on the CRD surface (**Figure 6.6**, **A**). It is located between helices α3 and 3_{10}, approximately 20–30 Å distant from the C-terminal residue Thr175, suggesting its spatial accessibility to a netrin-like domain, linked to the CRD via the amino acid linker of ten residues. Surface-exposed hydrophobic residues often form protein domain–domain interaction surfaces (Tsai et al., 1997; Moreira et al., 2007). This suggests that Phe146-PheGlyPhe-Tyr150 could potentially serve as a binding site for the netrin-like domain of sFRP-1 or other sFRP-1-binding proteins such as RANKL (Häusler et al., 2004; Chuman et al., 2004). This hydrophobic region may potentially be involved in Wnt binding too. A putative Wnt-3-binding site, encompassing residues Lys74–Glu85 (Bhat et al., 2007), is also located within 20–30 Å distance (**Figure 6.6**, **A**, 90° view). A 20–30 Å length region (separating Lys74–Glu85 from Phe146–Tyr150) could be potentially covered by one Wnt molecule (~40–45 kDa). This surface-exposed hydrophobic region is therefore a candidate site for further site-directed mutagenesis studies of sFRP-1 CRD interactions with numerous ligands (e.g. Wnts and RANKL).

The fourth notable feature is conformational invariability of a putative Wnt-3-binding-site, and the CRD core, stabilised by five disulfide bonds, in general (**Figure 6.6**, **B** and **Table 6.3**). The putative binding site is located on the CRD surface bordered by the N-terminus, start of the helix α1 and helix 3_{10} (**Figure 6.6**, **B**; Bhat et al., 2007). It encompasses only one, relatively short secondary structure element, β2 (residues Lys75–Val77). The site does not contain any cysteine residues (KKMVLPNLLEHE) and thus lacks direct conformational stabilisation by disulfide bonds. However, it shows low conformational variability among all 19 crystal structures (**Figure 6.6**). This could be explained by two structural features of the CRD. First, the peptide bond between Lys75 and Met76 is shielded from direct solvent

accessibility by the first disulfide bond of the domain (Cys58–Cys121). This sterical stabilisation of the selected residues might be a canonical feature of the CRDs as the corresponding cysteines are conserved across all CRDs analysed (**Figure 6.1**). Second, out of 12 residues involved in Wnt-3 recognition, only 3 of them have surface-exposed side chains (Lys75, Glu83 and Glu85), 2 have deeply buried side chains (Met76 and Asn80) while other residues are partially buried (**Figure 6.6**, **B**). This integration of partially- and fully-buried hydrophobic residues results in the putative Wnt-3-binding site being an only moderately hydrophobic, tightly packed structural unit.

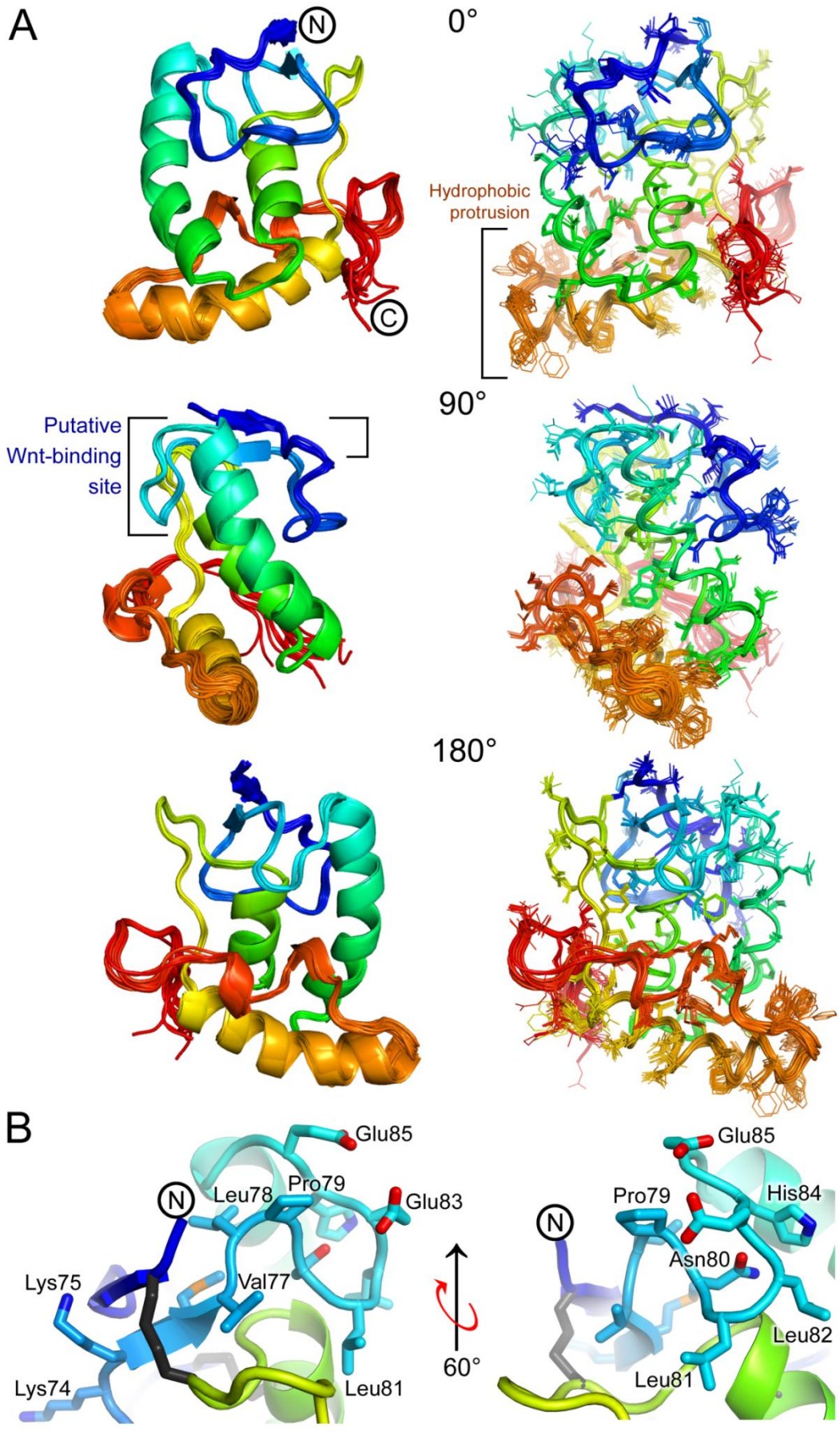

A
N
C
0°
Hydrophobic
protrusion
90°
Putative
Wnt-binding
site
180°
B
Glu85
N
Leu78
Pro79
Glu83
Lys75
Val77
Lys74
Leu81
60°
Glu85
Pro79
His84
N
Asn80
Leu82
Leu81

Figure 6.6. **A**, superposition of the 19 sFRP-1 CRD conformers from four crystal structures (**Tables 6.2** and **6.3**). The CRDs are colour coded as a rainbow with N- and C-termini shown in blue and red, respectively. A cartoon representation with secondary structure elements is shown on the left; backbone representation with side chains shown as lines is shown on the right. Three different views (from top to bottom) that differ by a 90° rotation about a vertical axis are presented. Orientation of the superposed CRDs in the 0° view is the same as in **Figure 6.4**, **A**. Putative Wnt-binding region (between residues Tyr73 and Thr86; Bhat et al., 2007) is indicated in the 90° view and coloured in blue to cyan transition. A surface exposed hydrophobic region (residues Phe146-Phe-Gly-Phe-Tyr150) is indicated in the 0° view and coloured in orange. Disulfide bonds and GlcNAc are not shown for simplicity. **B**, a putative Wnt-3-binding site (Bhat et al., 2007). Two different views that differ by a 60° rotation about a vertical axis are presented. Side chains of Lys74–Glu85 are shown as sticks (N, blue; O, red; S, orange). Disulfide bond between Cys58 and Cys121 is shown as dark grey sticks.

Table 6.3. Root-mean-square deviation (RMSD) values (given in Å) of pairwise comparison of the 19 sFRP-1 CRD structures (protein backbone Cα atoms only) (Krissinel and Henrick, 2004; http://www.ebi.ac.uk/msd-srv/ssm/). Crystal numbering and construct boundaries (e.g. 32–176) correspond to **Table 6.1**. CRD molecule chains from each single crystal form are indicated with letters A to F; CRDs forming non-crystallographic dimers are highlighted in green (e.g. sFRP-1 $CRD_{52–176}$, crystal I, chains A and B); the CRD pairs with the most divergent conformations (RMSD > 0.7 Å) are highlighted in yellow.

	32–176, A	32–176, B	32–176, C	32–176, D	52–176, crystal I, A	52–176, crystal I, B	52–176, crystal I, C	52–176, crystal II, A	52–176, crystal II, B	52–176, crystal II, C	52–176, crystal II, D	52–176, crystal II, E	52–176, crystal II, F	52–176, crystal III, A	52–176, crystal III, B	52–176, crystal III, C	52–176, crystal III, D	52–176, crystal III, E	52–176, crystal III, F
32–176, A	0	0.17	0.18	0.21	0.5	0.45	0.48	0.58	0.62	0.72	0.6	0.64	0.74	0.46	0.39	0.47	0.39	0.47	0.38
32–176, B	0.17	0	0.11	0.13	0.48	0.47	0.45	0.56	0.6	0.7	0.62	0.65	0.72	0.4	0.34	0.46	0.38	0.49	0.32
32–176, C	0.18	0.11	0	0.13	0.46	0.45	0.43	0.54	0.57	0.67	0.63	0.67	0.7	0.4	0.34	0.44	0.39	0.45	0.35
32–176, D	0.21	0.13	0.13	0	0.46	0.46	0.42	0.54	0.58	0.69	0.65	0.68	0.71	0.39	0.34	0.45	0.4	0.49	0.32
52–176, crystal I, A	0.5	0.48	0.46	0.46	0	0.24	0.18	0.44	0.45	0.48	0.57	0.61	0.49	0.5	0.53	0.51	0.51	0.56	0.48
52–176, crystal I, B	0.45	0.47	0.45	0.46	0.24	0	0.23	0.48	0.49	0.56	0.52	0.59	0.56	0.55	0.54	0.55	0.52	0.59	0.53
52–176, crystal I, C	0.48	0.45	0.43	0.42	0.18	0.23	0	0.41	0.42	0.49	0.57	0.61	0.5	0.49	0.5	0.51	0.49	0.56	0.47
52–176, crystal II, A	0.58	0.56	0.54	0.54	0.44	0.48	0.41	0	0.13	0.47	0.39	0.42	0.28	0.58	0.59	0.5	0.53	0.55	0.55
52–176, crystal II, B	0.62	0.6	0.57	0.58	0.45	0.49	0.42	0.13	0	0.25	0.44	0.49	0.28	0.61	0.62	0.53	0.57	0.60	0.59
52–176, crystal II, C	0.72	0.7	0.67	0.69	0.48	0.56	0.49	0.47	0.25	0	0.47	0.49	0.12	0.67	0.70	0.55	0.63	0.60	0.65
52–176, crystal II, D	0.6	0.62	0.63	0.65	0.57	0.52	0.57	0.39	0.44	0.47	0	0.2	0.48	0.69	0.68	0.59	0.58	0.64	0.63
52–176, crystal II, E	0.64	0.65	0.67	0.68	0.61	0.59	0.61	0.42	0.49	0.49	0.2	0	0.51	0.73	0.72	0.62	0.62	0.66	0.63
52–176, crystal II, F	0.74	0.72	0.7	0.71	0.49	0.56	0.5	0.28	0.28	0.12	0.48	0.51	0	0.68	0.71	0.57	0.64	0.62	0.67
52–176, crystal III, A	0.46	0.4	0.4	0.39	0.5	0.55	0.49	0.58	0.61	0.67	0.69	0.73	0.68	0	0.21	0.43	0.43	0.55	0.33
52–176, crystal III, B	0.39	0.34	0.34	0.34	0.53	0.54	0.5	0.59	0.62	0.70	0.68	0.72	0.71	0.21	0	0.44	0.4	0.5	0.32
52–176, crystal III, C	0.47	0.46	0.44	0.45	0.51	0.55	0.51	0.5	0.53	0.55	0.59	0.62	0.57	0.43	0.44	0	0.35	0.31	0.38
52–176, crystal III, D	0.39	0.38	0.39	0.4	0.51	0.52	0.49	0.53	0.57	0.63	0.58	0.62	0.64	0.43	0.4	0.35	0	0.39	0.31
52–176, crystal III, E	0.47	0.49	0.45	0.49	0.56	0.59	0.56	0.55	0.60	0.60	0.64	0.66	0.62	0.55	0.5	0.31	0.39	0	0.42
52–176, crystal III, F	0.38	0.32	0.35	0.32	0.48	0.53	0.47	0.55	0.59	0.65	0.63	0.63	0.67	0.33	0.32	0.38	0.31	0.42	0

6.7 The sFRP-1 CRD interactions with Wnt-3a and glycosaminoglycans

The sFRP-1 CRD crystal structures reveal multiple, surface-exposed, positively charged amino acid residues that form an extended, positively charged, but fragmentary patch on the domain surface (**Figure 6.7**, **A** and **B**). The patch is proximal to a putative Wnt-binding site and extends across one side of the CRD (**Figure 6.7**, **B**). It is composed of Arg66, His69, Lys74, Lys75 and Lys103. This sFRP-1 CRD surface property directed me to test whether the CRD binds to heparin and heparan sulfate (HS). Heparin and heparan sulfate are highly sulfated (thus negatively charged), heterogeneous glycosaminoglycans (GAGs) (discussed in **Section 5.2**). They are typically attached to the core of heparan sulfate proteoglycans (HSPGs) present on the cell surface. HSPGs bind to morphogens (or lipoprotein particles that co-localise with morphogens; Panáková et al., 2005) directly via HSPG core–morphogen or GAG–morphogen interactions (reviewed by Yan and Lin, 2009). These interactions increase the concentration of morphogens at the cell surface and may promote signalling. For example, *Drosophila* HSPG, called Dally-like protein, promotes Wingless signalling when present at low concentration (Gallet et al., 2008) but inhibits signalling when present at high concentration (Yan et al., 2009). This biphasic function of Dally-like protein suggests that, when present at low concentration, it prevents Wingless loss into the distant extracellular space and localises it next to the receptor Frizzled-2; however, when present at high concentration, Dally-like protein competes with Frizzled-2 for Wingless binding (Yan et al., 2009). I tested whether the human sFRP-1 CRD binds to heparin, HS, human HSPG Glypican-3 (GPC3), mouse Wnt-3a, and whether heparin and sFRP-1 CRD competes for Wnt-3a binding.

Physical interactions between the sFRP-1 $CRD_{52–176}$ and two different GAGs, heparin and HS, were studied using an SPR binding assay. GAGs were biotinylated on their reducing terminus (as described in **Section 2.7.3**) and immobilised on the streptavidin-coated SPR

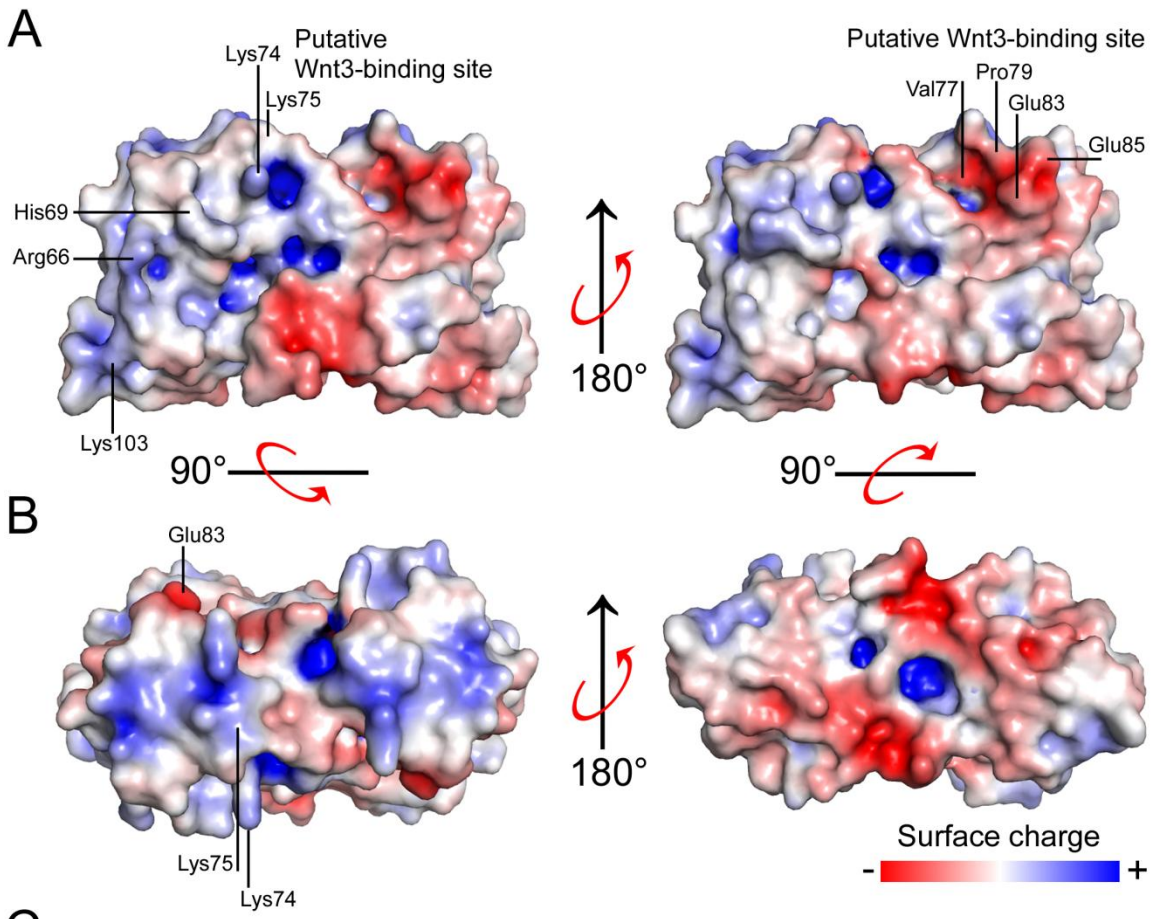
sensor surface. Binding to purified sFRP-1 $CRD_{52–176}$ was measured. The tightest binding was observed between the most sulfated GAG, heparin (2.8 sulfate groups per disaccharide on average; Zhang et al., 2007), and sFRP-1 $CRD_{52–176}$ (K_d 6.2 μM) (**Figure 6.7, C**). A two-fold weaker binding was detected between less sulfated GAG, HS, (1.2 sulfate groups per disaccharide; Zhang et al., 2007) and sFRP-1 $CRD_{52–176}$ (K_d 13.1 μM). The correlation between higher GAG sulfation level and tighter binding to sFRP-1 $CRD_{52–176}$ suggests the importance of electrostatic interactions between the negatively charged GAG sulfate groups and the positively charged sFRP-1 $CRD_{52–176}$ residues. To assess the contribution of the positively charged primary amine groups of lysines to GAG binding, interactions between reductively methylated (Walter et al., 2006) sFRP-1 $CRD_{52–176}$ and GAGs were measured. The binding affinity between chemically modified sFRP-1 $CRD_{52–176}$ (compared to a non-modified CRD) and GAGs was approximately 10–20 times lower as indicated by higher K_d values (**Figure 6.7**, **D**), suggesting an involvement of the ε-amino groups of lysine residues in GAG recognition.

I further tested whether the sFRP-1 $CRD_{52–176}$ binds to human GPC3. Three GPC3 constructs, encompassing putative GAG attachment sites were expressed in HEK-293T cells (**Section 2.7.5**), biotinylated and immobilised on the streptavidin-coated SPR sensor surface. Binding to a purified mouse Sonic Hedgehog (a kind gift from Dr. Christian Siebold; protein construct ShhN is described by Bishop et al., 2008; mouse Sonic Hedgehog was reported to bind to GPC3 with high-affinity, K_d 32 nM; Capurro et al., 2008) and sFRP-1 $CRD_{52–176}$ was measured. In contrast to the study by Capurro et al., 2008, a very weak binding was observed between all three GPC3 constructs and Sonic Hedgehog or sFRP-1 $CRD_{52–176}$ (**Figure 6.8**, **A** and **B**). The binding response suggests that only a fraction (less than 50 %, assuming 1:1 binding stoichiometry) of immobilised GPC3 binds to Sonic Hedgehog and sFRP-1 $CRD_{52–176}$. The reasons underlying this weak binding remain to be investigated. One possible cause

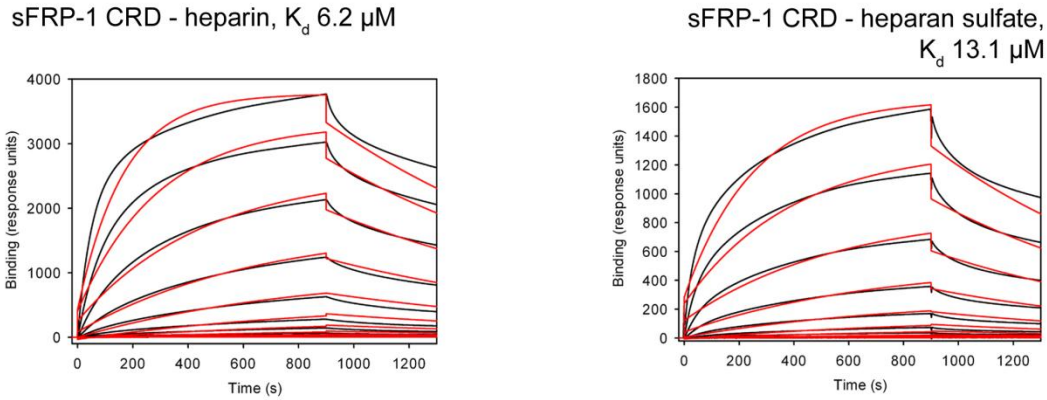
of weak binding is a low GAG sulfation level for the GPC3 expressed by HEK-293T cells. Further quantitative analysis (possibly by mass spectrometry) to assess the sulfation level of the GAGs attached to GPC3 would be informative.

The potential of GAGs to compete with the sFRP-1 $CRD_{52–176}$ for Wnt-3a binding was investigated using an SPR binding assay. Purified mouse Wnt-3a (**Sections 2.2.3** and **4.4**) was immobilised on the SPR sensor surface via primary amines and binding was measured for sFRP-1 $CRD_{52–176}$. The calculated K_d (11 μM) for Wnt-3a–sFRP-1 $CRD_{52–176}$ interactions is consistent with previously reported K_ds of 11.2 and 85.9 nM calculated for commercially available mouse Wnt-3a – full length human sFRP-1 interactions (both proteins are sold in a lyophilised state by R&D Systems), which assumed the presence of at least two binding sites for Wnt-3a on sFRP-1 (Wawrzak et al., 2007). The lower binding affinity between sFRP-1 CRD–Wnt-3a (reported here) compared to the full length sFRP-1–Wnt-3a (Wawrzak et al., 2007) suggests that the netrin-like domain of sFRP-1 contributes to Wnt-3a binding. Indeed, it has been reported that the netrin-like domain of sFRP-1 binds to Wnt-8 and antagonises the canonical Wnt signalling pathway (Lopez-Rios et al., 2008).

I further tested whether heparin affects binding between sFRP-1 $CRD_{52–176}$ and Wnt-3a. The binding affinity between the two proteins was weaker in the presence of 1 mM heparin, as indicated by a higher K_d value of 45.2 μM (K_d in the absence of heparin was 11 μM) (**Figures 6.8**, **C** and **D**). This is consistent with a previous observation that Lys74 and Lys75 are important for Wnt-3 binding (Bhat et al., 2007); these two lysines may also potentially bind to heparin and HS, consistent with the observation that Wnt-3a and GAGs compete with each other for the sFRP-1 CRD binding. This suggests that glycosaminoglycans, such as heparin and heparan sulfate (and thus HSPGs), may potentially promote Wnt signalling by preventing Wnt interactions with antagonists (such as sFRP-1) or inhibit signalling by weakening Wnt binding to the CRDs of Frizzled receptors (**Figure 6.9**).


A
Lys74
Putative
Wnt3-binding site
Lys75
His69
Arg66
Lys103
90°
180°
Putative Wnt3-binding site
Val77
Pro79
Glu83
Glu85
90°
B
Glu83
180°
Lys75
Lys74
Surface charge
-
+

C


sFRP-1 CRD - heparin, K$_d$ 6.2 µM
Binding (response units)
Time (s)
sFRP-1 CRD - heparan sulfate,
K$_d$ 13.1 µM
Binding (response units)
Time (s)

D

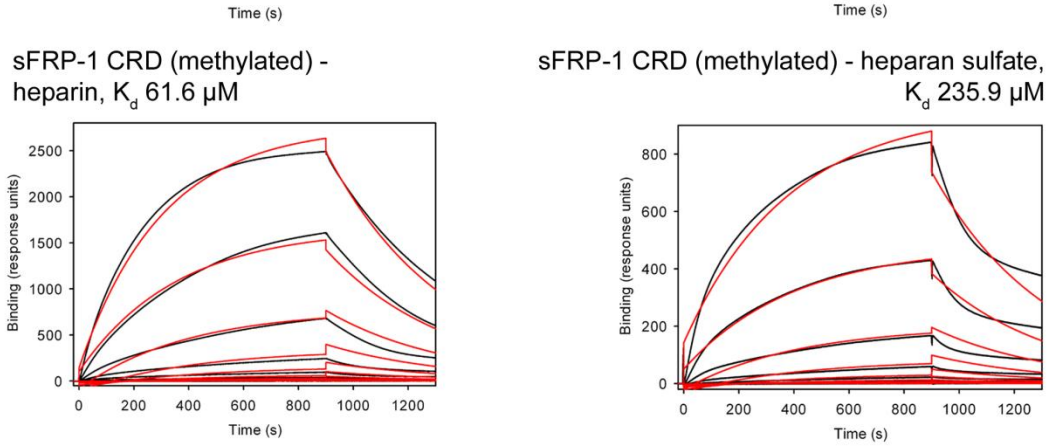

sFRP-1 CRD (methylated) -
heparin, K$_d$ 61.6 µM
Binding (response units)
Time (s)
sFRP-1 CRD (methylated) - heparan sulfate,
K$_d$ 235.9 µM
Binding (response units)
Time (s)

Figure 6.7. Electrostatic properties of sFRP-1 $CRD_{52–176}$ crystal III structure (chains A and B). The CRD homodimer is oriented as in **Figure 6.4**, **A** and is shown as solvent-accessible surface coloured by electrostatic potential contoured at ± 8 kT/e (red, negative (acidic); blue, positive (basic)). **A**, front and back views that differ by a 180° rotation about a vertical axis are presented on the left and right, respectively. **B**, top and bottom views that differ by a 180° rotation about a vertical axis are presented on the left and right, respectively. Residues that interfere with the sFRP-1 antagonism of Wnt-3 signalling (Lys74–Glu85) (Bhat et al., 2007) and form an extended, fragmentary, positively charged surface (Arg66, His69, Lys74, Lys75 and Lys103) are indicated. Note that Lys75 potentially binds GAGs, but is also implicated in Wnt-3 binding (Bhat et al., 2007). **C**, surface plasmon resonance (SPR) binding sensorgrams of sFRP-1 $CRD_{52–176}$ to heparin (left; ~800 RUs of biotinylated heparin was immobilised on the streptavidin-coated sensor chip surface) and heparan sulfate (right; ~700 RUs of biotinylated heparin was immobilised on the streptavidin-coated sensor chip surface). SPR binding sensorgrams (shown as black curves) in which a concentration series (17.0 nM–34.8 μM, two-fold dilutions) of sFRP-1 $CRD_{52–176}$ was injected over immobilized glycosaminoglycans (**C**). **D**, SPR binding sensorgrams of the methylated sFRP-1 $CRD_{52–176}$ to heparin (left) and heparan sulfate (right). The same SPR sensor chip as in **C** was used. SPR binding sensorgrams (shown as black curves) in which a concentration series (13.4 nM–27.4 μM, two-fold dilutions) of sFRP-1 $CRD_{52–176}$ was injected over immobilized glycosaminoglycans (**D**). Binding responses (black curves) were fit to a simple 1:1 kinetic interaction model (red curves) using BIAevaluation (Biacore Life Sciences) software; dissociation constants (K_ds) calculated from a single experiment are shown above the plots.

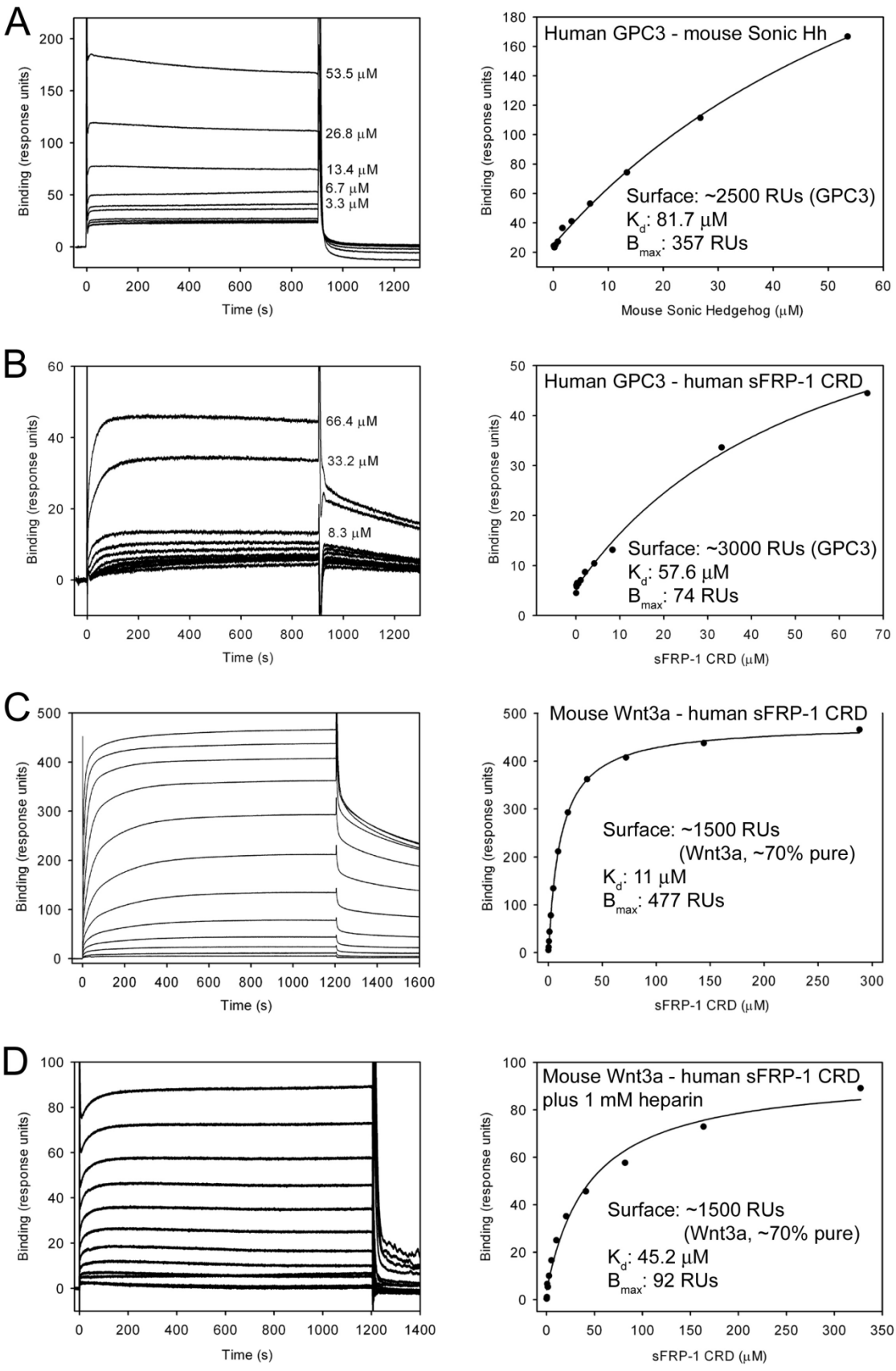

A
Binding (response units)
Time (s)
53.5 μM
26.8 μM
13.4 μM
6.7 μM
3.3 μM
Human GPC3 - mouse Sonic Hh
Surface: ~2500 RUs (GPC3)
K_d: 81.7 μM
B_max: 357 RUs
Mouse Sonic Hedgehog (μM)
B
66.4 μM
33.2 μM
8.3 μM
Human GPC3 - human sFRP-1 CRD
Surface: ~3000 RUs (GPC3)
K_d: 57.6 μM
B_max: 74 RUs
sFRP-1 CRD (μM)
C
Mouse Wnt3a - human sFRP-1 CRD
Surface: ~1500 RUs
(Wnt3a, ~70% pure)
K_d: 11 μM
B_max: 477 RUs
D
Mouse Wnt3a - human sFRP-1 CRD
plus 1 mM heparin
Surface: ~1500 RUs
(Wnt3a, ~70% pure)
K_d: 45.2 μM
B_max: 92 RUs

Figure 6.8. A very weak binding of mouse Sonic Hedgehog (**A**, a positive control experiment) and human sFRP-1 $CRD_{52–176}$ (**B**) to biotinylated human heparan sulfate proteoglycan GPC3 (residues Pro31–Ile514) immobilised on the streptavidin-coated sensor chip surface. Plots of the binding response derived from the sensorgrams (shown on the left) as a function of Sonic Hedgehog (**A**) and sFRP-1 $CRD_{52–176}$ (**B**) concentration are shown on the right. Binding of human sFRP-1 $CRD_{52–176}$ to mouse Wnt-3a (covalently linked to the sensor chip surface via primary amines) in the absence (**C**) or presence (**D**) of 1 mM heparin. The amount of immobilised molecules on the sensor chip surface (expressed in response units, RUs), K_ds (calculated for 1:1 equilibrium interaction model using BIAevaluation (Biacore Life Sciences) software from a single experiment) and estimated B_{max} (maximum binding response at saturating analyte (e.g. sFRP-1 CRD) concentration, calculated using BIAevaluation software) are shown on the right, under the binding plot curves.

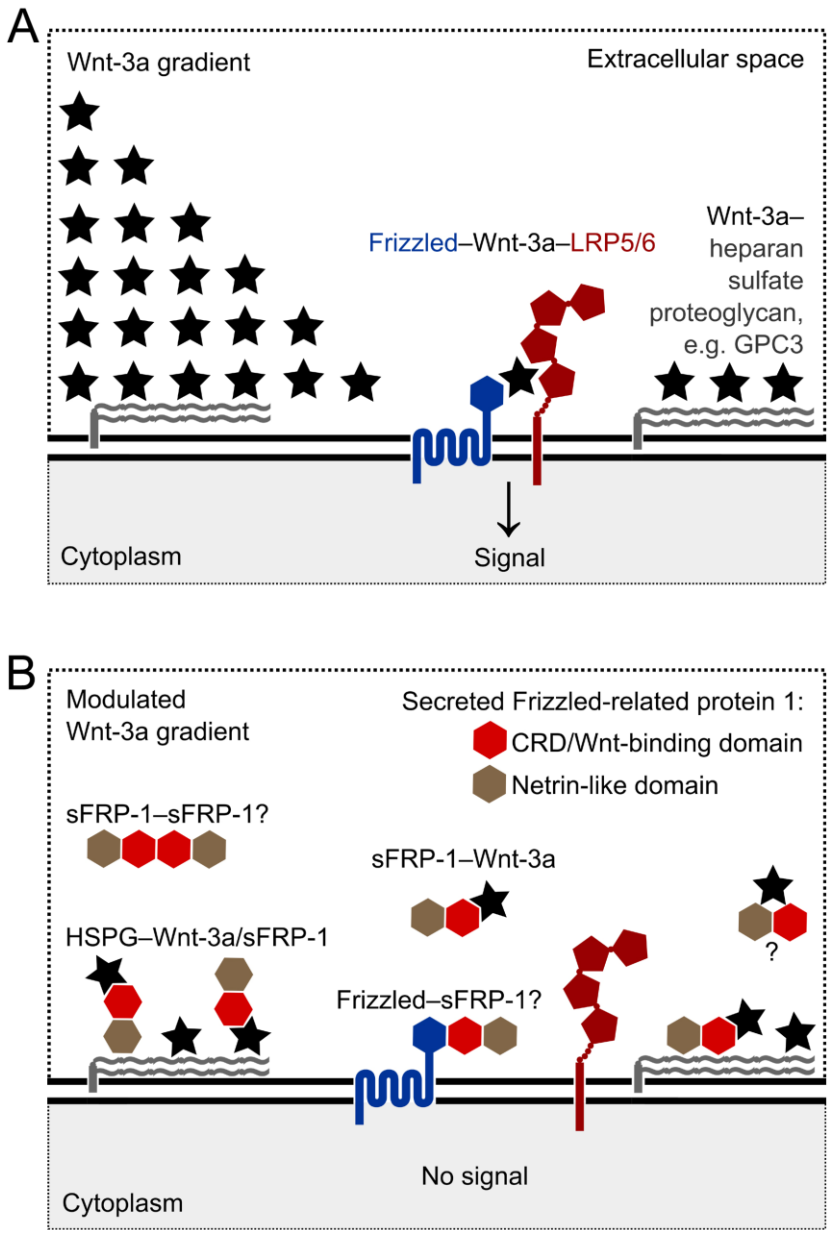


Figure 6.9. A, the canonical Wnt/β-catenin signalling pathway is initiated at the cell surface when secreted Wnt-3a (black star) forms a ternary complex with the cell-surface receptors Frizzled (blue) and LRP5/6 (dark red). This results in a downstream signal transduction and activation of the Wnt target genes. Glycosaminoglycan (e.g. heparan sulfate) chains are covalently linked to heparan sulfate proteoglycans (HSPGs, e.g. GPC3) on the cell surface, bind to Wnts and modulate their extracellular gradients. **B**, sFRP-1 may modulate Wnt signalling at least via three mechanisms: (1) inhibition by binding to Wnts and sequestering them from their receptors, (2) inhibition by forming non-signalling complexes with Frizzled receptors via Frizzled CRD (shown as a blue hexagon) interaction with the sFRP-1 CRD (shown as an orange hexagon) and (3) activation by bind and inhibiting function of the other sFRPs via CRD–CRD interactions (not shown). In addition, sFRP-1 binds to heparan sulfate chains via either its CRD or netrin-like domain (shown as a brown hexagon), or both domains. CRD–Wnt interaction might be further modulated by heparan sulfate chains, **Figure 6.8**, **C–D**. Whether sFRP-1 forms dimers *in vivo*, and whether the netrin-like domain and the CRD form an extended Wnt-binding surface remains unknown (indicated with a question mark (?)).

6.8 Conclusions and future perspectives

Nineteen 3.1–3.8 Å resolution crystal structures from four different crystal forms of the human sFRP-1 cysteine-rich domain (CRD) reveal the following: (1) a conserved hydrophobic homodimer interface, (2) conformationally variable N- and C-termini, (3) conformationally invariable, putative Wnt-3-binding site and (4) a positively charged surface region located proximal to a putative Wnt-3-binding site. The sFRP-1 CRD binds both to mouse Wnt-3a (K_d 11 μM) and glycosaminoglycans, heparin (K_d 6.2 μM) and heparan sulfate (K_d 13.1 μM). The sFRP-1 CRD binds weaker (K_d 45.2 μM) to Wnt-3a in the presence of 1 mM heparin, consistent with proximity and potential overlap of a putative Wnt-3-binding site and a positively charged CRD surface, a potential glycosaminoglycan-interaction site. The structural and binding results presented here provide a basis for further structure-guided studies to elucidate the role of the sFRP-1 CRD dimerisation and Wnt signalling modulation by glycosaminoglycans.

Chapter 7 – Further discussions

The canonical Wnt/β-catenin signalling pathway is a fundamental morphogenetic pathway of multicellular organisms that is used to regulate diverse processes such as primary body axis (head-to-tail) formation (Petersen and Reddien, 2008) and stem cell self-renewal (Reya et al., 2003). Diverse intra- and intercellular molecular mechanisms have evolved to limit Wnt/β-catenin signalling to specific locations and time as discussed in **Chapter 1**. The sFRP-1 and WIF-1 represent extracellular antagonists of the Wnt/β-catenin signalling pathway (**Chapters 2** to **6**). Some of the similarities and differences between sFRP-1 and WIF-1 are briefly highlighted in **Chapter 7**. The Wnt/β-catenin signalling pathway remained an active area of research and attracted attention both from academia and the biotechnology industry during preparation of this thesis (2007–2011). Some of the current discoveries, challenges and questions are briefly mentioned in **Chapter 7**.

7.1 Similarities between sFRP-1 and WIF-1

The sFRP-1 and WIF-1 share a number of common features listed below. (1) Both proteins contain multiple domains (WIF domain (WD) plus 5 EGF-like domains in WIF-1; cysteine-rich (CRD) and netrin-like domains in sFRP-1). (2) Both antagonists contain domains that are also present in the transmembrane, cell surface Wnt receptors (Frizzled receptor contains an extracellular CRD whereas the RYK receptor contains a WD). (3) Both antagonists directly bind to Wnt proteins via receptor-like, ligand-binding domains, WD and CRD, and prevent Wnt/β-catenin signalling. In both cases, the interplay between the WD and EGF-like domains in WIF-1, the CRD and netrin-like domain in sFRP-1, and the contribution of this interplay to the protein function is poorly understood. This thesis describes the first insights into the interplay between the WD and EGF-like domains in WIF-1, and the implications of this interplay for the antagonism of Wnt/β-catenin signalling. The interplay between the netrin-

like domain and CRD in sFRPs remains unknown. (4) Both sFRP-1 and WIF-1 are expressed during embryonic development as well as in adult organisms. In *Xenopus*, WIF-1 is expressed during somitogenesis (Hsieh et al., 1999; in the vertebrate embryo, somites will eventually become dermis, skeletal muscle, and vertebrae). Mouse sFRP-1 is expressed in specific regions of the embryonic forebrain and midbrain (Hoang et al., 1998). (5) In mouse, knockouts of either WIF-1 or sFRP-1 result in viable organisms with relatively mild disorders. Homozygous sFRP-1-null mice are fertile but display decreased brain weight and slight hematological abnormalities in males (Mouse Genome Database, Blake et al., 2011). Targeted deletion of the mouse Wif1 gene accelerates development of radiation-induced osteosarcomas in vivo (Kansara et al., 2009). Relatively mild disorders observed in either WIF-1 or sFRP-1 single knockout organisms might be explained by the presence of multiple Wnt/β-catenin signalling antagonists with possibly overlapping function. In contrast to the single sFRP-1 knockout, double homozygous mutations in *Sfrp1* and *Sfrp2* lead to severe shortening of the thoracic region (Satoh et al., 2006). By contrast, a homozygous mutation in one or the other exerts no effect on embryogenesis, suggesting that sFRP-1 and sFRP-2 are functionally redundant (Satoh et al., 2006). (6) Both WIF-1 and sFRP-1 are heparan sulfate proteoglycan (HSPG)-binding proteins and (7) are absent in *Drosophila*, the model organism widely used to study HSPG-shaped Wnt morphogen gradients (Yan and Lin, 2009).

7.2 Differences between sFRP-1 and WIF-1

Although both WIF-1 and sFRP-1 were initially described as extracellular proteins that bind to Wnt morphogens in the extracellular space and prevent them from triggering signalling, more elaborate molecular mechanisms are emerging. Since the discovery of WIF-1 (Hsieh et al., 1999), the only known function of vertebrate WIF-1 remains Wnt binding and Wnt/β-catenin signalling inhibition. (1) Surprisingly, in contrast to WIF-1, sFRP-1 acts as an agonist of Wingless signalling when it is present at low concentration (Üren et al., 2000). However,

similarly to WIF-1, sFRP-1 inhibits Wingless signalling when is present at high concentration (Üren et al., 2000). The molecular mechanism responsible for the biphasic modulation of Wingless signalling by sFRP-1 remains unknown. (2) WIF-1 has never been reported to form oligomers or bind to the transmembrane RYK receptors via WD–WD interactions, whereas the CRDs of sFRP-1, -3, and Frizzled 8 form homodimers in all known crystal structures (Dann et al., 2001; Chapter 6). (3) sFRP-1 also interacts with Frizzled 2 and regulates the growth of retinal ganglion cell axons (Rodriguez et al., 2005). The CRD is largely responsible for the effect of sFRP-1 on axon outgrowth (Rodriguez et al., 2005). (4) Both WIF-1 and sFRP-1 recognise different proteins that do not belong to the Wnt/β-catenin signalling pathway. Vertebrate WIF-1 binds to a range of Wnt proteins whereas *Drosophila* WIF-1 (Shifted) does not bind to Wnts but possibly interacts with another morphogen, Hedgehog (Glise et al., 2005; Gorfinkiel et al., 2005). sFRP-1 binds via its CRD to RANKL, a receptor activator of nuclear factor kappa-B ligand which activates osteoclast differentiation (Häusler et al., 2004). (5) sFRPs, possibly, are the first extracellular regulators of Wnt signalling that emerged during evolution. They are present in *Amphimedon queenslandica*, a sponge from the Great Barrier Reef, whereas WIF-1 is a vertebrate specific Wnt-binding protein (Adamska et al., 2010; Hsieh et al., 1999). (6) Probably the most distinctive feature of WIF-1, discussed in detail in this thesis, is its lipid-binding properties. Although, it has been noted that CRDs of sFRPs and Frizzled receptors are structurally similar to the cholesterol-binding protein NPC (Bazan and de Sauvage, 2009), no interactions between CRDs and lipids have been reported so far.

7.3 Future directions

Wnt proteins may act as morphogens in a concentration-dependent manner and differ in that respect from most extracellular signalling molecules. Thus a better understanding of Wnt/β-catenin signalling pathway requires an understanding of morphogen gradients (Rogers and

Schier, 2011). Wnt morphogen gradients provide a mechanism by which the secretion of Wnt signalling proteins from one part of an embryo (or an adult organism) can determine the location, differentiation and fate of surrounding cells.

In general, an understanding of morphogen gradients requires answers to two different questions (Gurdon and Bourillot, 2001). The first question is how a morphogen gradient (in this case, Wnt gradient) is formed. To answer this question we need to know (1) the properties of Wnts, (2) the shape and absolute concentration of the gradient, and (3) the factors that create and maintain the gradient in its various positions (Gurdon and Bourillot, 2001). Unfortunately, Wnts are poorly expressed, hydrophobic proteins (Willert et al., 2003; Chapters 1 and 2) as a result biophysical and biochemical characterisation of Wnt properties (1) remains a major challenge. Thus future experiments on how to produce pure Wnt proteins for biophysical studies in high quantities (milligrams) is of great importance. Detection of Wnt gradients in vivo (2) remains experimentally challenging as well because Wnts (and other morphogens) are present at low concentration in the extracellular space and are hard to fix for visualisation in situ. However, the emerging super-resolution light microscopy techniques (such as stimulated emission depletion (STED) or stochastic optical reconstruction microscopy (STORM), reviewed by Hell, 2003, and Huang et al., 2010, respectively) are likely to be informative when applied to the study of Wnt and other morphogen gradients. (3) The work presented in this thesis extends our understanding of some factors, such as WIF-1, sFRP-1 and HSPGs, and the contribution of their interplay to the modulation of the Wnt/β-catenin signalling pathway.

The second major question in morphogen research is how cells interpret morphogen, in this case Wnt, concentration (Gurdon and Bourillot, 2001). To understand it, we need to know (4) how cells recognize different threshold concentrations of Wnts through multiple receptors and (5) how they transduce this information to the nucleus to create the specific cell fate

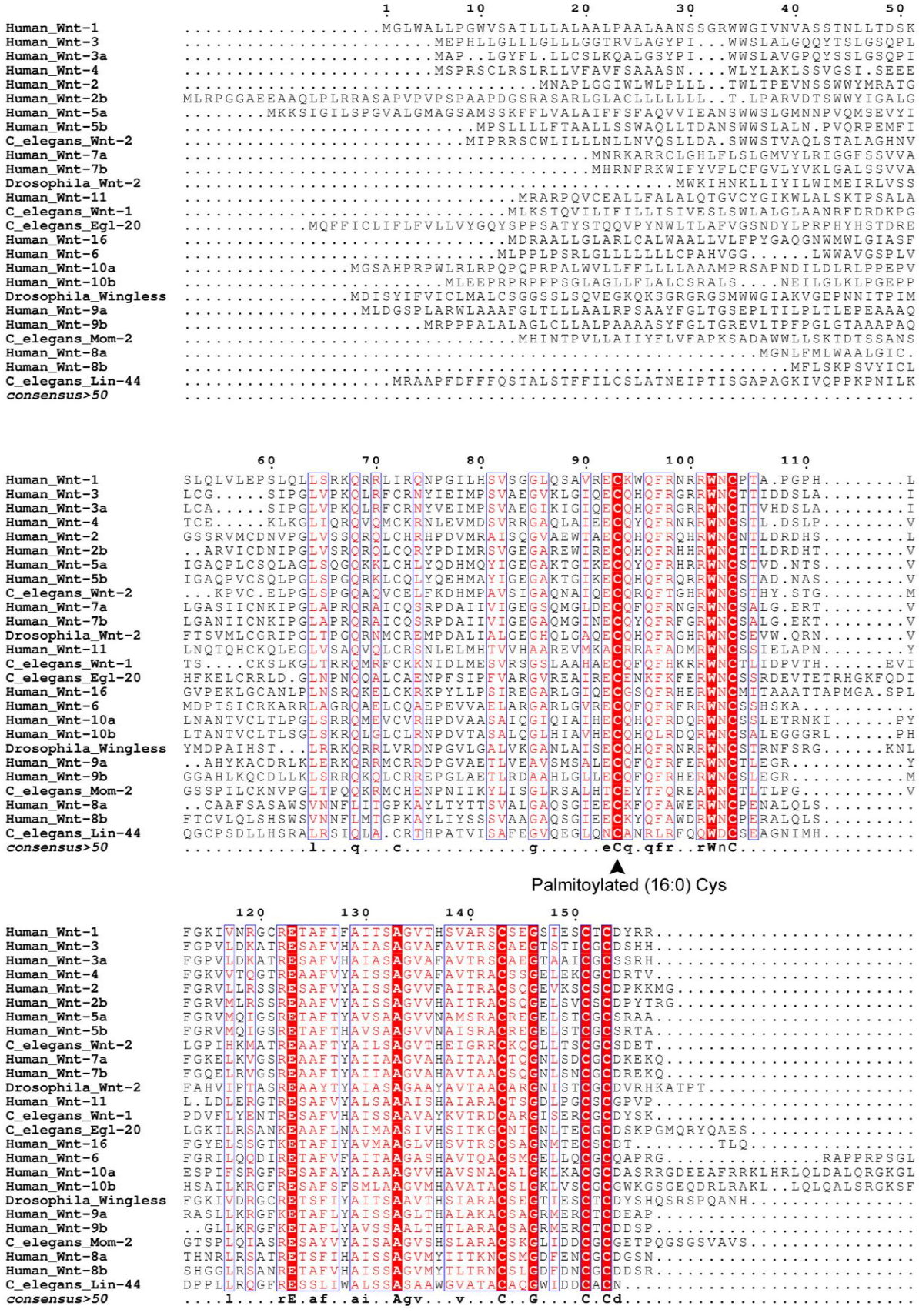
response. The specificity of multiple Wnt-receptor interactions remains poorly defined because purified Wnts and the extracellular domains of their receptors are typically not available in sufficient quantities for detailed biochemical studies. Crystallographic studies of the LRP5/6–Wnt–Frizzled complexes remain a daunting challenge in the foreseeable future. However, work on Wnt–WIF-1 interactions provides a better understanding of both Wnt–antagonist and Wnt–receptor interactions because the Wnt-binding WIF domain is present in both WIF-1 and the transmembrane Wnt receptor RYK. Recent structural studies of interactions between the Wnt co-receptor LRP6 and different ligands (Wnt-1, Wnt-3a, the antagonist Dickkopf and the LRP5/6 chaperone MESD) complements previous studies on Wnt–CRD interactions and provides a clearer picture of how the ternary LRP5/6–Wnt–Frizzled complex may assemble on the cell surface and trigger downstream signalling (Chen et al., 2011; Ahn et al., 2011; Cheng et al., 2011). However, how cells recognize different threshold concentrations of Wnts (4) and how they transduce this information to the nucleus to create specific cell fate response (5), remains largely elusive.

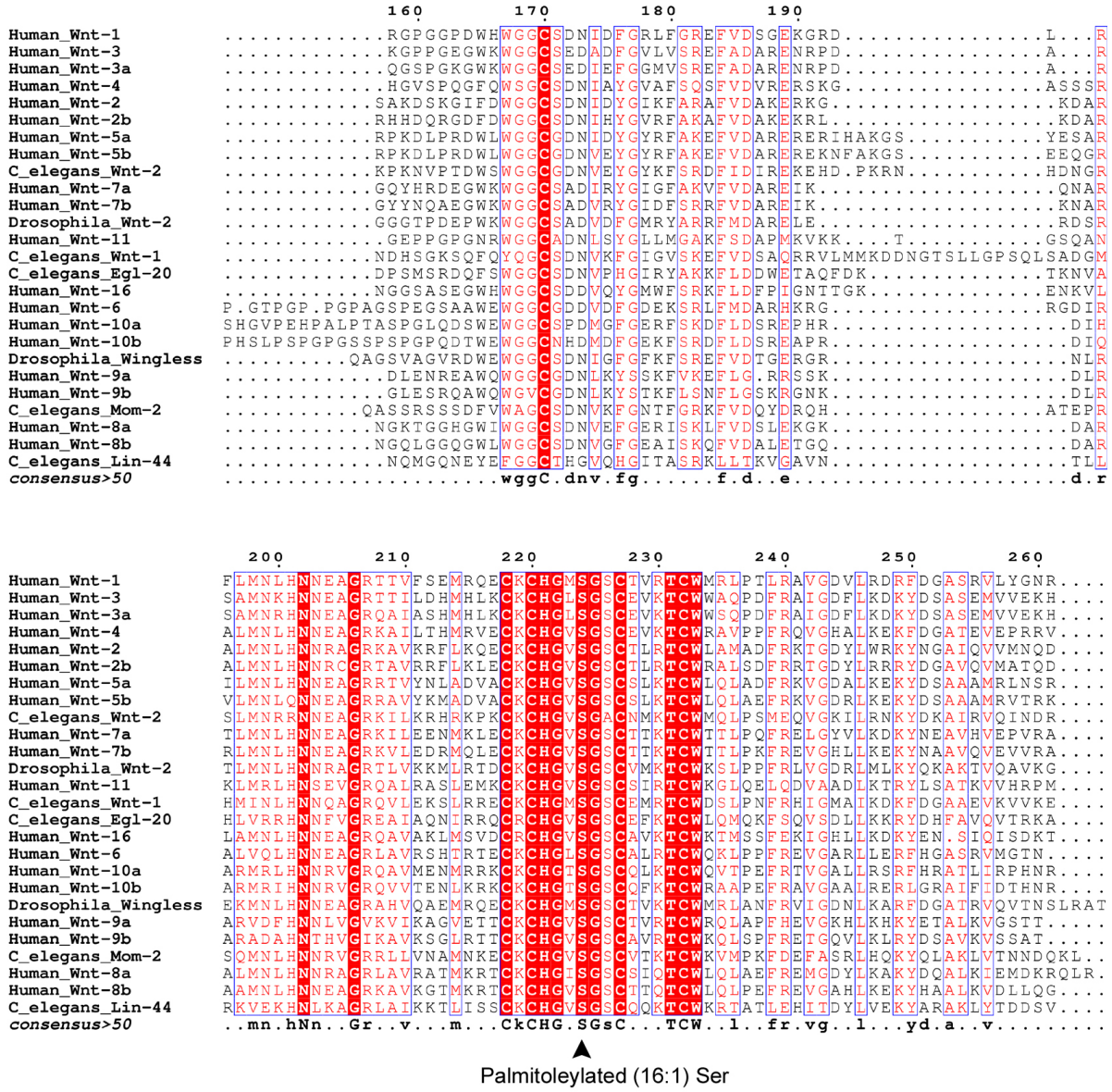
Deregulation of the canonical Wnt/β-catenin signalling pathway leads to multiple life-threatening human diseases. The Wnt/β-catenin pathway was first linked to cancer formation when it was found that mutations in APC can cause colon cancer. (Kinzler et al., 1991; Nishisho et al., 1991). Since then, aberrant Wnt signalling has been implicated in different kinds of cancers, osteoporosis and, recently, in neurological diseases (Polakis, 2007; Krishnan et al., 2006; De Ferrari and Moon, 2006). However, there are no drugs in the market that specifically target the aberrant Wnt/β-catenin signalling pathway. The lack of progress is largely due to the lack of "druggable" enzyme targets in the pathway (Garber, 2009). Despite this, researchers from the biotechnology companies are making a progress by targeting Wnts, their extracellular antagonists (notably, monoclonal antibodies against Sclerostin, an LRP5/6-binding protein, increase bone formation and show potential in human clinical trials as a

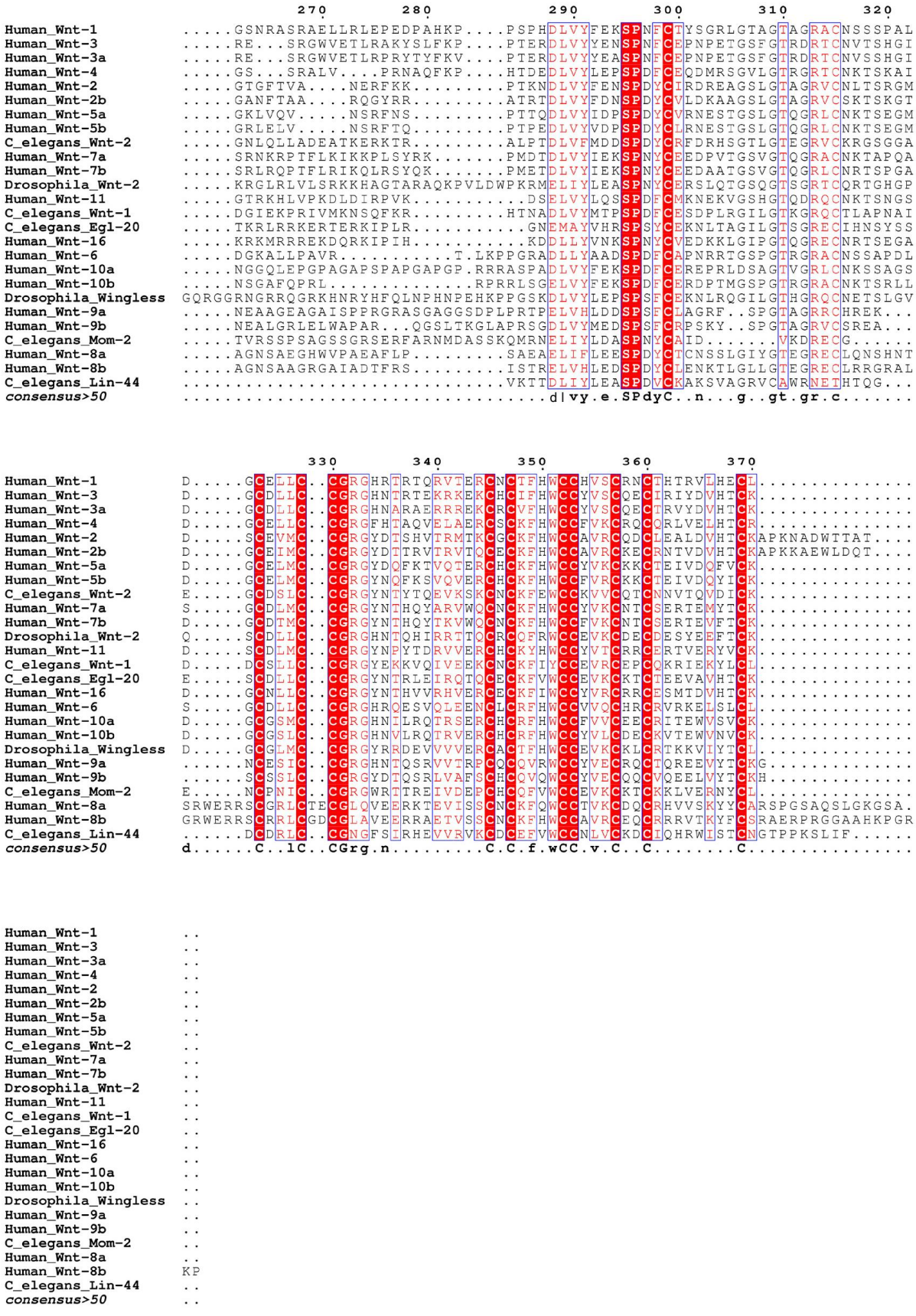
treatment of osteoporosis) and recently identified cytoplasmic Wnt/β-catenin pathway components (such as tankyrase, an enzyme which destabilizes Axin and thus promotes β-catenin target gene activation) (DeAlmeida et al., 2007; Padhi et al., 2011; Huang et al., 2009). In addition, modulation of stem cell self-renewal using Wnt proteins holds promise for regenerative medicine (Willert et al., 2003; Reya et al., 2003; Ito et al., 2007). To summarise, the canonical Wnt/β-catenin signalling pathway remains an active area of research, and a better understaning of this fundamental pathway will help us to solve specific problems in human health.

Appendices

Appendix 1 – Amino acid sequence alignment of Wnt proteins

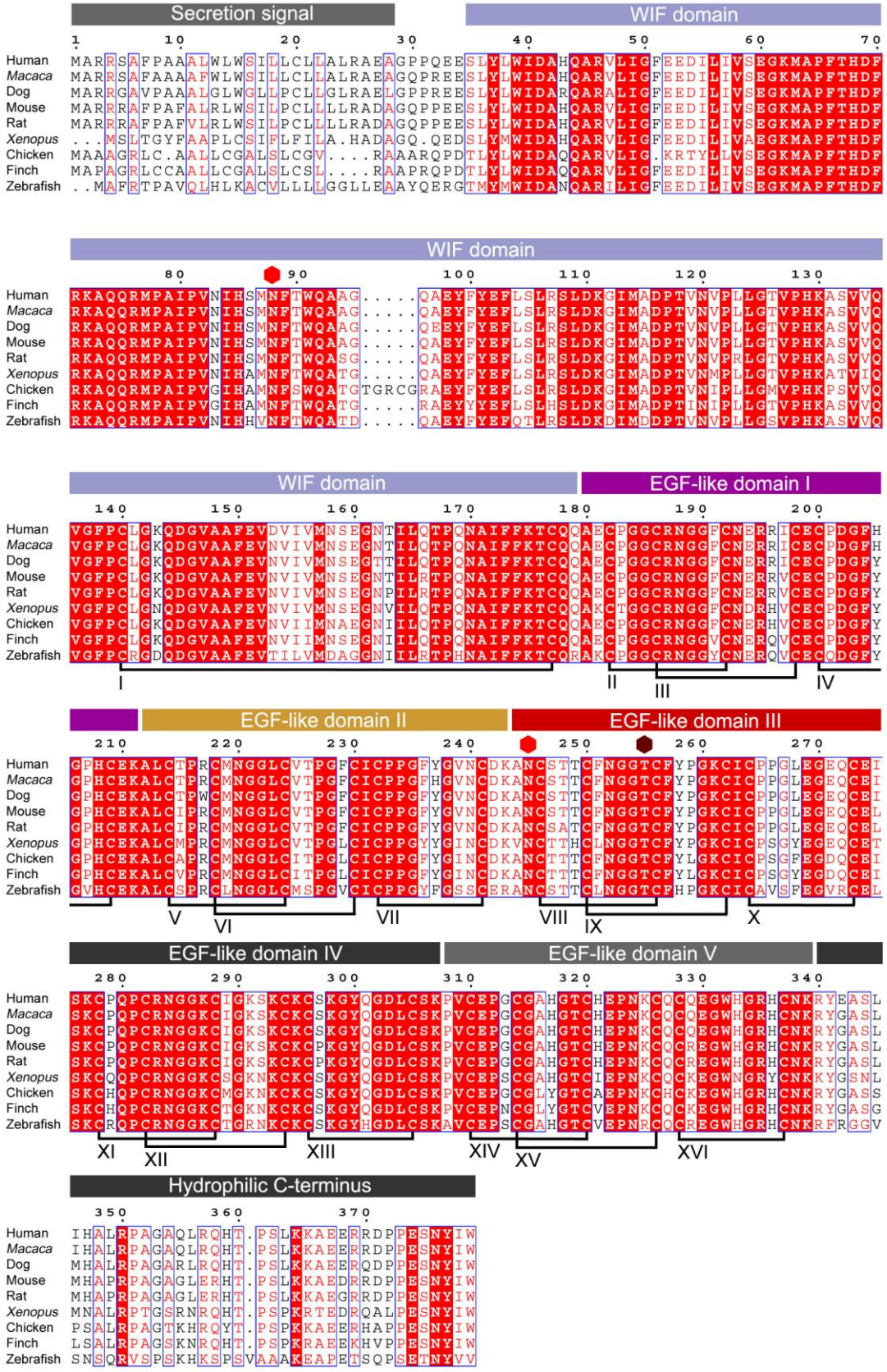



Palmitoleylated (16:1) Ser


270 280 290 300 310 320
Human_Wnt-1 .....GSNRASRAELLRLEPEDPAHKP....PSPHDLVYFEKSPNFCTYSGRLGTAGTAGRACNSSSPAL
Human_Wnt-3 .....RE...SRGWVETLRAKYSLFKP....PTERDLVYYENSPNFCEPNPETGSFGTRDRTCNVTSHGI
Human_Wnt-3a .....RE...SRGWVETLRPRYTYFKV....PTERDLVYYEASPNFCEPNPETGSFGTRDRTCNVSSHGI
Human_Wnt-4 .....GS...SRALV....PRNAQFKP....HTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAI
Human_Wnt-2 .....GTGFTVA....NERFKK.........PTKNDLVYFENSPDYCIRDREAGSLGTAGRVCNLTSRGM
Human_Wnt-2b .....GANFTAA....RQGYRR.........ATRTDLVYFDNSPDYCVLDKAAGSLGTAGRVCSKTSKGT
Human_Wnt-5a .....GKLVQV.....NSRFNS.........PTTQDLVYIDPSPDYCVRNESTGSLGTQGRLCNKTSEGM
Human_Wnt-5b .....GRLELV.....NSRFTQ.........PTPEDLVYVDPSPDYCLRNESTGSLGTQGRLCNKTSEGM
C_elegans_Wnt-2 .....GNLQLLADEATKERKTR.........ALPTDLVFMDDSPDYCRFDRHSGTLGTEGRVCKRGSGGA
Human_Wnt-7a .....SRNKRPTFLKIKKPLSYRK.......PMDTDLVYIEKSPNYCEEDPVTGSVGTQGRACNKTAPQA
Human_Wnt-7b .....SRLRQPTFLRIKQLRSYQK.......PMETDLVYIEKSPNYCEEDAATGSVGTQGRLCNRTSPGA
Drosophila_Wnt-2 .....KRGLRLVLSRKKHAGTARAQKPVLDWPKRMELIYLEASPNYCERSLQTGSQGTSGRTCQRTGHGP
Human_Wnt-11 .....GTRKHLVPKDLDIRPVK..........DSELVYLQSSPDFCMKNEKVGSHGTQDRQCNKTSNGS
C_elegans_Wnt-1 .....DGIEKPRIVMKNSQFKR.........HTNADLVYMTPSPDFCESDPLRGILGTKGRQCTLAPNAI
C_elegans_Egl-20 .....TKRLRRKERTERKIPLR..........GNEMAYVHRSPSYCEKNLTAGILGTSGRECIHNSYSS
Human_Wnt-16 .....KRKMRRREKDQRKIPIH..........KDDLLYVNKSPNYCVEDKKLGIPGTQGRECNRTSEGA
Human_Wnt-6 .....DGKALLPAVR..........T.LKPPGRADLLYAADSPDFCAPNRRTGSPGTRGRACNSSAPDL
Human_Wnt-10a .....NGGQLEPGPAGAPSPAPGAPGP.RRRASPADLVYFEKSPDFCEREPRLDSAGTVGRLCNKSSAGS
Human_Wnt-10b .....NSGAFQPRL.............RPRRLSGELVYFEKSPDFCERDPTMGSPGTRGRACNKTSRLL
Drosophila_Wingless GQRGGRNGRRQGRKHNRYHFQLNPHNPEHKPPGSKDLVYLEPSPSFCEKNLRQGILGTHGRQCNETSLGV
Human_Wnt-9a .....NEAAGEAGAISPPRGRASGAGGSDPLPRTPELVHLDDSPSFCLAGRF..SPGTAGRRCHREK...
Human_Wnt-9b .....NEALGRLELWAPAR...QGSLTKGLAPRSGDLVYMEDSPSFCRPSKY..SPGTAGRVCSREA...
C_elegans_Mom-2 .....TVRSSPSAGSSGRSERFARNMDASSKQMRNELIYLDASPNYCAID.......VKDRECG......
Human_Wnt-8a .....AGNSAEGHWVPAEAFLP.........SAEAELIFLEESPDYCTCNSSLGIYGTEGRECLQNSHNT
Human_Wnt-8b .....AGNSAAGRGAIADTFRS.........ISTRELVHLEDSPDYCLENKTLGLLGTEGRECLRRGRAL
C_elegans_Lin-44 .....VKTTDLIYLEASPDVCKAKSVAGRVCAWRNETHTQG...
consensus>50 .....d|vy.e.SPdyC..n...g..gt.gr.c.....
330 340 350 360 370
Human_Wnt-1 D.....GCELLC..CGRGHRTRTQRVTERCNCTFHWCCHVSCRNCTHTRVLHECL.....
Human_Wnt-3 D.....GCDLLC..CGRGHNTRTEKRKEKCHCIFHWCCYVSCQECIRIYDVHTCK.....
Human_Wnt-3a D.....GCDLLC..CGRGHNARAERRREKCRCVFHWCCYVSCQECTRVYDVHTCK.....
Human_Wnt-4 D.....GCELLC..CGRGFHTAQVELAERCSCKFHWCCFVKCRQCQRLVELHTCR.....
Human_Wnt-2 D.....SCEVMC..CGRGYDTSHVTRMTKCGCKFHWCCAVRCQDCLEALDVHTCKAPKNADWTTAT....
Human_Wnt-2b D.....GCEIMC..CGRGYDTTRVTRVTQCECKFHWCCAVRCKECRNTVDVHTCKAPKKAEWLDQT....
Human_Wnt-5a D.....GCELMC..CGRGYDQFKTVQTERCHCKFHWCCYVKCKKCTEIVDQFVCK.....
Human_Wnt-5b D.....GCELMC..CGRGYNQFKSVQVERCHCKFHWCCFVRCKKCTEIVDQYICK.....
C_elegans_Wnt-2 E.....GCDSLC..CGRGYNTYTQEVKSKCNCKFEWCCKVVCQTCNNVTQVDICK.....
Human_Wnt-7a S.....GCDLMC..CGRGYNTHQYARVWQCNCKFHWCCYVKCNTCSERTEMYTCK.....
Human_Wnt-7b D.....GCDTMC..CGRGYNTHQYTKVWQCNCKFHWCCFVKCNTCSERTEVFTCK.....
Drosophila_Wnt-2 Q.....SCDLLC..CGRGHNTQHIRRTTQCRCQFRWCCEVKCDECDESYEEFTCK.....
Human_Wnt-11 D.....SCDLMC..CGRGYNPYTDRVVERCHCKYHWCCYVTCRRCERTVERYVCK.....
C_elegans_Wnt-1 D.....DCSLLC..CGRGYEKKVQIVERKCNCKFIYCCEVRCEPCQKRIEKYLCL.....
C_elegans_Egl-20 E.....SCDLLC..CGRGYNTRLEIRQTQCECKFVWCCEVKCKTCTEEVAVHTCK.....
Human_Wnt-16 D.....GCNLLC..CGRGYNTHVVRHVERCECKFIWCCYVRCRRCESMTDVHTCK.....
Human_Wnt-6 S.....GCDLLC..CGRGHRQESVQLEENCLCRFHWCCVVQCHRCRVRKELSLCL.....
Human_Wnt-10a D.....GCGSMC..CGRGHNILRQTRSERCHCRFHWCCFVVCEECRITEWVSVCK.....
Human_Wnt-10b D.....GCGSLC..CGRGHNVLRQTRVERCHCRFHWCCYVLCDECKVTEWVNVCK.....
Drosophila_Wingless D.....GCGLMC..CGRGYRRDEVVVVERCACTFHWCCEVKCKLCRTKKVIYTCL.....
Human_Wnt-9a .....NCESIC..CGRGHNTQSRVVTRPCQCQVRWCCYVECRQCTQREEVYTCKG.....
Human_Wnt-9b .....SCSSLC..CGRGYDTQSRLVAFSCHCQVQWCCYVECQQCVQEELVYTCKH.....
C_elegans_Mom-2 E.....NCPNIC..CGRGWRTTREIVDEPCHCQFVWCCEVKCKTCKKLVERNYCL.....
Human_Wnt-8a SRWERRSCGRLCTECGLQVEERKTEVISSCNCKFQWCCTVKCDQCRHVVSKYYCARSPGSAQSLGKGSA.
Human_Wnt-8b GRWERRSCRRLCGDCGLAVEERRAETVSSCNCKFHWCCAVRCEQCRRVTKYFCSRAERPRGGAAHKPGR
C_elegans_Lin-44 .....DCDRLC..CGNGFSIRHEVVRVKCDCEFVWCCNLVCKDCIQHRWISTCNGTPPKSLIF.....
consensus>50 d.....C..lC..CGrg.n.....C.C.f.wCC.v.C..C.....C.....
Human_Wnt-1 ..
Human_Wnt-3 ..
Human_Wnt-3a ..
Human_Wnt-4 ..
Human_Wnt-2 ..
Human_Wnt-2b ..
Human_Wnt-5a ..
Human_Wnt-5b ..
C_elegans_Wnt-2 ..
Human_Wnt-7a ..
Human_Wnt-7b ..
Drosophila_Wnt-2 ..
Human_Wnt-11 ..
C_elegans_Wnt-1 ..
C_elegans_Egl-20 ..
Human_Wnt-16 ..
Human_Wnt-6 ..
Human_Wnt-10a ..
Human_Wnt-10b ..
Drosophila_Wingless ..
Human_Wnt-9a ..
Human_Wnt-9b ..
C_elegans_Mom-2 ..
Human_Wnt-8a ..
Human_Wnt-8b KP
C_elegans_Lin-44 ..
consensus>50 ..

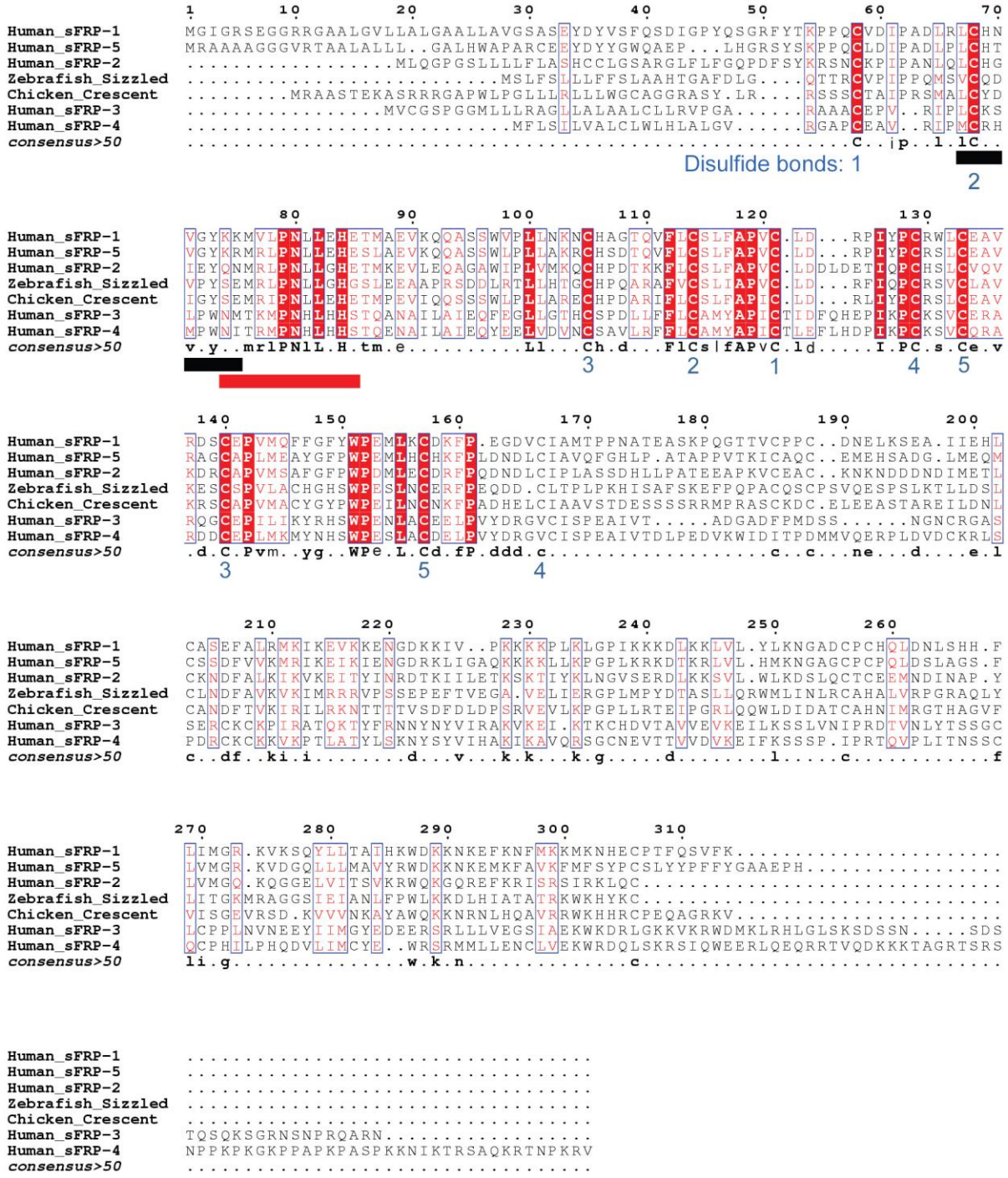
Appendix 1. Amino acid sequence alignment of all human Wnt proteins with selected *Drosophila melanogaster* and *Caenorhabditis elegans* Wnts. Alignment of human Wnt-1 (UniProtKB/Swiss-Prot identification accession number P04628; UniProt (Universal Protein Resource) is a comprehensive resource for protein sequence and annotation data, maintained by the European Bioinformatics Institute (EBI), the Swiss Institute of Bioinformatics (SIB) and the Protein Information Resource (PIR)), Wnt-2 (P09544), Wnt-2b (Q93097), Wnt-3 (P56703), Wnt-3a (P56704), Wnt-4 (P56705), Wnt-5a (P41221), Wnt-5b (Q9H1J7), Wnt-6 (Q9Y6F9), Wnt-7a (O00755), Wnt-7b (P56706), Wnt-8a (Q9H1J5), Wnt-8b (Q93098), Wnt-9a (O14904), Wnt-9b (O14905), Wnt-10a (Q9GZT5), Wnt-10b (O00744), Wnt-11 (O96014), Wnt-16 (Q9UBV4); *Drosophila melanogaster* Wingless (P09615, also called Wnt-1), Wnt-2 (P28465); *Caenorhabditis elegans* Wnt-1 (P34888), Wnt-2 (P34889), Egl-20 (Q9TVJ1, *eg*g *l*aying defective-20), Mom-2 (Q10459, a member (-2) of the *m*ore *o*f *M*S gene class) and Lin-44 (Q27886, abnormal cell lineage protein-44). Numbering corresponds to the full length human Wnt-1. Evolutionarily conserved residues are shown in a red background. The sequences were aligned using MultAlin (Corpet, 1988; http://multalin.toulouse.inra.fr/) and formatted with ESPript (Gouet et al., 1999; http://espript.ibcp.fr/).

Appendix 2 – Amino acid sequence alignment of the vertebrate WIF-1 family members



Appendix 2. Amino acid sequence alignment of the vertebrate Wnt inhibitory factor 1 (WIF-1) family members. Alignment of human (SwissProt identification accession number Q9Y5W5), *Macaca mulatta* (NCBI reference sequence XP_538269.2), dog (NCBI XP_538269.2), mouse (SwissProt Q9WUA1), rat (SwissProt Q6IN38), *Xenopus laevis* (SwissProt Q9W6F8), chicken (NCBI NP_001186536), zebra finch (*Taeniopygia guttata*, NCBI XP_002188540.1), zebrafish (SwissProt Q9W6F9) WIF-1, respectively. Numbering corresponds to the full length human WIF-1. Evolutionarily conserved residues are shown in a red background. Domain boundaries are indicated by overbars above the alignments; the colour scheme corresponds to that in **Figure 5.4**. Disulfide bonds are indicated with black lines and marked with Roman numerals. Following WIF-$1_{\Delta C}$ crystal structure refinement, electron density corresponding to the N-linked N-acetylglucosamine could be seen at two glycosylation sites: Asn88 and Asn245 (marked by filled red hexagons) (described in **Chapter 5**). O-linked fucose could be seen at one glycosylation site, Thr255 (marked by filled brown hexagon) (described in **Chapter 5**). Evolutionarily conserved residues are shown in a red background. The sequences were aligned using MultAlin (Corpet, 1988; http://multalin.toulouse.inra.fr/) and formatted with ESPript (Gouet et al., 1999; http://espript.ibcp.fr/).

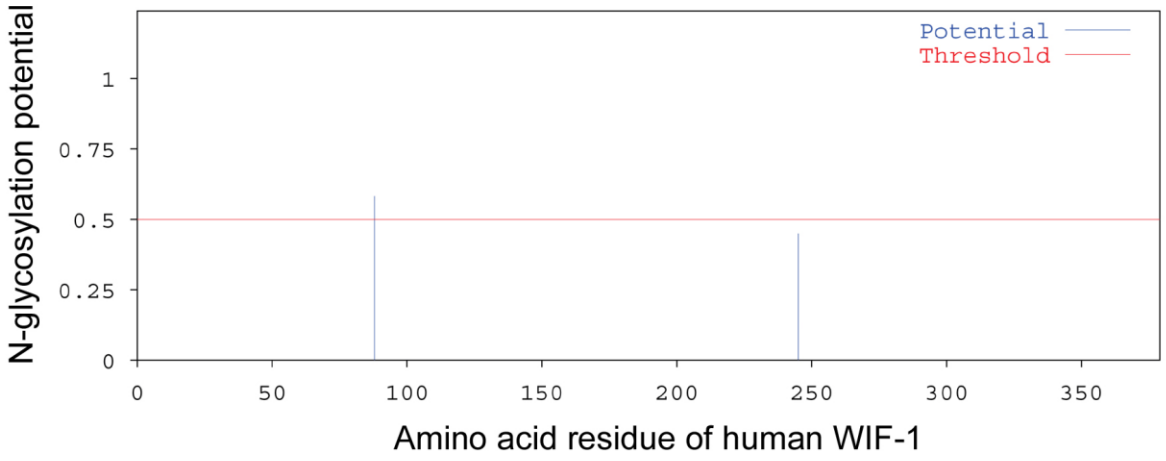
Appendix 3 – Amino acid sequence alignment of the vertebrate sFRP family members



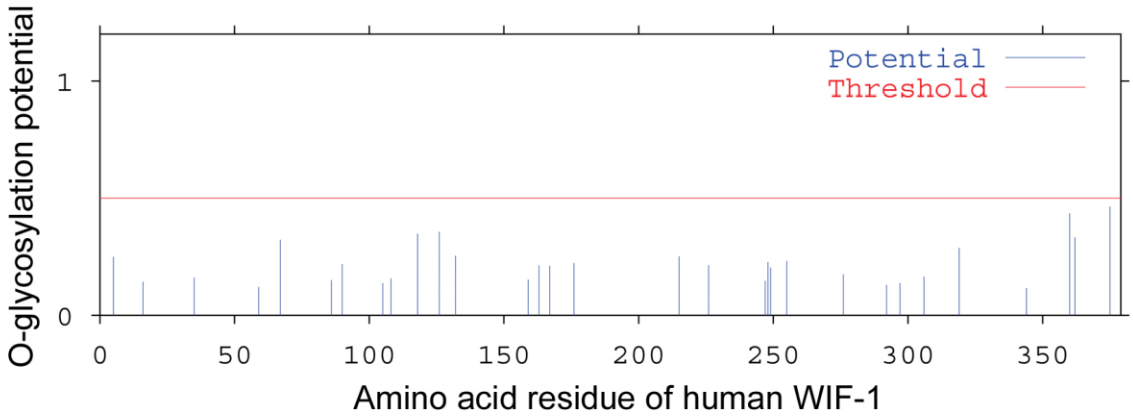
Appendix 3. Amino acid sequence alignment of human secreted Frizzled-related proteins 1–5, *Danio rerio* (zebrafish) Sizzled and *Gallus gallus* (chicken) Crescent proteins. Alignment of human sFRP-1 (UniProtKB/Swiss-Prot identification accession number Q8N474), sFRP-2 (Q96HF1), sFRP-3 (Q92765), sFRP-4 (Q6FHJ7), sFRP-5 (Q5T4F7), zebrafish Sizzled (Q7SX78) and chicken Crescent (O42397). Cysteine residues forming the five disulfide bonds of the N-terminal cysteine-rich domain (CRD) are numbered. Amino acid residues implicated in *Drosophila* Frizzled-2 CRD– and mouse Frizzled-6 CRD–*Xenopus* Wnt-8 binding are highlighted with a black bar below the sequence alignments (Dann et al., 2001). Amino acid residues implicated in the sFRP-1 antagonism of Wnt-3 signalling are highlighted with a red bar below the sequence alignments (Bhat et al., 2007). Numbering corresponds to the full length human sFRP-1. Evolutionarily conserved residues are shown in a red background. The sequences were aligned using MultAlin (Corpet, 1988; http://multalin.toulouse.inra.fr/) and formatted with ESPript (Gouet et al., 1999; http://espript.ibcp.fr/).

Appendix 4 – Analysis of human WIF-1 amino acid sequence

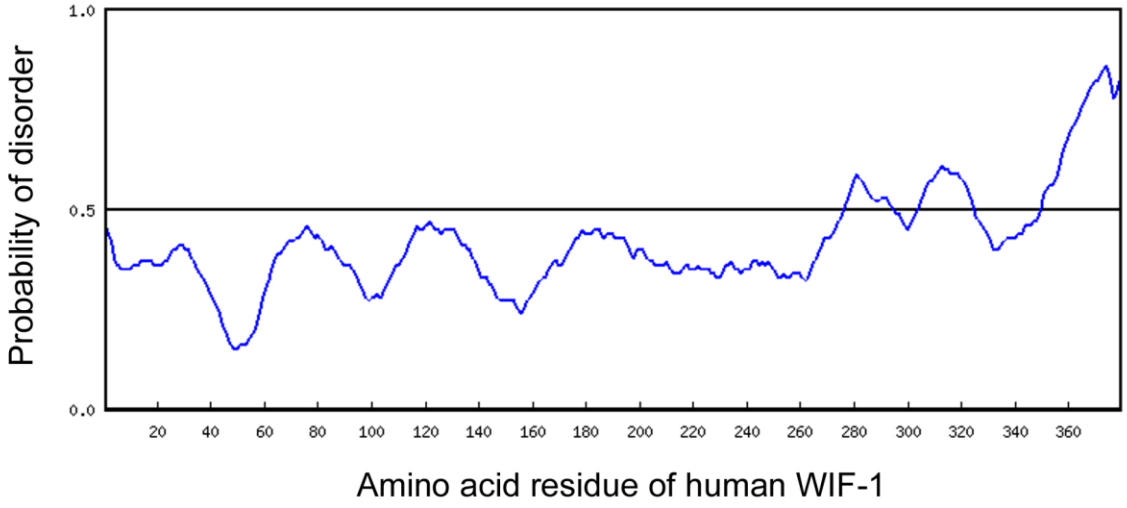
A



B



C



Appendix 4. Analysis of human WIF-1 amino acid sequence. **A**, analysis of human WIF-1 amino acid sequence using the NetNGlyc 1.0 Server (Gupta et al., 2004; http://www.cbs.dtu.dk/services/NetNGlyc/) suggests putative glycosylation sites at Asn88 (potential 0.58) and Asn245 (potential 0.45). The NetNGlyc predicts N-linked glycosylation sites in human proteins using artificial neural networks that examine the sequence context of Asn-Xaa-Ser/Thr sequons (Gupta et al., 2004). Glycosylation of both Asn88 (see **Section 3.6**) and Asn245 (see **Section 5.5**) was confirmed by X-ray crystallographic studies. **B**, analysis of human WIF-1 amino acid sequence using the NetOGlyc 3.1 Server (Julenius et al., 2005; http://www.cbs.dtu.dk/services/NetOGlyc/) suggests no putative O-linked glycosylation sites (the highest G-score of 0.435 for Ser375 is below the threshold level). The NetOglyc server produces neural network predictions of mucin type GalNAc O-glycosylation sites in mammalian proteins (Julenius et al., 2005). NetOGlyc 3.1 was not able to identify O-linked glycosylation (fucosylation) at Thr255 observed in the WIF-$1_{\Delta C}$ crystal structure (see **Section 5.5**). **C**, analysis of human WIF-1 amino acid sequence using the disorder prediction server RONN (Yang et al., 2005; http://www.strubi.ox.ac.uk/RONN) suggests three disordered regions of WIF-1: Lys277–Cys294, Cys305–Asn324, Arg350–Trp379. The first two regions are parts of the EGF-like domains IV and IV–V, respectively. Thus, regions Lys277–Cys294 and Cys305–Asn324 are likely to represent relatively ordered parts of the protein (similar to the EGF-like domain I as observed in the WIF-$1_{WD\text{-}EGF\text{-}I}$ crystal structure, see **Section 3.6**). The third region Arg350–Trp379 does not exhibit significant sequence similarity to any known domain fold and is possibly lacking any secondary structure elements (is disordered as predicted by RONN).

Appendix 5 – 3D alignment of the WD crystal structure and entries of the PDB

Pairwise comparison and 3D alignment of WIF domain crystal structure (from WIF-$1_{WD\text{-}EGF\text{-}I}$ construct, residues Ser35–Gln179) and 69266 entries (170247 chains) of the Protein Data Bank (PDB); 20 identified entries with the highest Q-score (see description below) are listed.

No.	Scoring (Q)[a]	RMSD[b]	N_{algn}[c]	PDB ID (chain)	Title of the target PDB entry
1	0.58	1.90	136	2D3J (A)	NMR STRUCTURE OF THE WIF DOMAIN FROM HUMAN WIF-1
2	0.24	2.97	79	2BP3 (B)	CRYSTAL STRUCTURE OF FILAMIN A DOMAIN 17 AND GPIB ALPHA CYTOPLASMIC DOMAIN COMPLEX
3	0.24	3.00	79	2BP3 (A)	CRYSTAL STRUCTURE OF FILAMIN A DOMAIN 17 AND GPIB ALPHA CYTOPLASMIC DOMAIN COMPLEX
4	0.23	3.83	99	3EAD (D)	CRYSTAL STRUCTURE OF CALX-CBD1
5	0.23	2.72	73	3CNK (B)	CRYSTAL STRUCTURE OF THE DIMERIZATION DOMAIN OF HUMAN FILAMIN A
6	0.23	3.79	98	3E9T (B)	CRYSTAL STRUCTURE OF APO-FORM CALX CBD1 DOMAIN
7	0.22	2.83	74	3CNK (A)	CRYSTAL STRUCTURE OF THE DIMERIZATION DOMAIN OF HUMAN FILAMIN A
8	0.22	3.87	109	1FT0 (B)	CRYSTAL STRUCTURE OF TRUNCATED HUMAN RHOGDI K113A MUTANT
9	0.22	3.82	108	2JHT (D)	CRYSTAL STRUCTURE OF RHOGDI K135T,K138T,K141T MUTANT
10	0.22	3.82	108	1FT3 (B)	CRYSTAL STRUCTURE OF TRUNCATED RHOGDI K141A MUTANT
11	0.22	3.04	75	2JF1 (A)	CRYSTAL STRUCTURE OF THE FILAMIN A REPEAT 21 COMPLEXED WITH THE INTEGRIN BETA2 CYTOPLASMIC TAIL PEPTIDE
12	0.22	3.66	94	3EAD (C)	CRYSTAL STRUCTURE OF CALX-CBD1
13	0.22	3.73	105	1QVY (C)	CRYSTAL STRUCTURE OF RHOGDI K(199,200)R DOUBLE MUTANT
14	0.22	2.97	76	3ISW (A)	CRYSTAL STRUCTURE OF FILAMIN-A IMMUNOGLOBULIN-LIKE REPEAT 21 BOUND TO AN N-TERMINAL PEPTIDE OF CFTR
15	0.22	3.79	97	3GIN (A)	CRYSTAL STRUCTURE OF E454K-CBD1
16	0.22	3.60	92	3EAD (A)	CRYSTAL STRUCTURE OF CALX-CBD1
17	0.22	3.86	107	1FST (A)	CRYSTAL STRUCTURE OF TRUNCATED HUMAN RHOGDI TRIPLE MUTANT
18	0.22	3.84	96	3E9T (A)	CRYSTAL STRUCTURE OF APO-FORM CALX CBD1 DOMAIN
19	0.21	2.83	74	2W0P (A)	CRYSTAL STRUCTURE OF THE FILAMIN A REPEAT 21 COMPLEXED WITH THE MIGFILIN PEPTIDE
20	0.21	3.01	88	1MSP (A)	MAJOR SPERM PROTEIN, ALPHA ISOFORM (RECOMBINANT), PH 4.6
...	...	...	...	...	...
718	0.14	3.78	104	1DOA (B)	STRUCTURE OF THE RHO FAMILY GTP-BINDING PROTEIN CDC42 IN COMPLEX WITH THE MULTIFUNCTIONAL REGULATOR RHOGDI

[a]**Q-score** represents the quality function of Cα alignment, maximised by the secondary structure matching (SSM) alignment algorithm. Quality of alignment is a complex characteristic, which takes both the alignment length and RMSD into account. In general, more meaningful 3D alignments correspond to lower RMSD and higher number of aligned residues N_{algn}. However, it is almost always possible to achieve lower RMSD at the expense of the alignment length and vice versa. Q-score, given by expression

$$Q = (N_{algn} \times N_{algn})/((1+(RMSD/R_0)^2) \times N_{res1} \times N_{res2}),$$

reduces the effect of RMSD–N_{algn} balance on the estimation of alignments (N_{res1} and N_{res2} stand for the number of residues in the aligned proteins, and empirical parameter R_0 is set to 3 Å).

Q-score reaches 1 only in the case of identical structures, and drops down with increasing RMSD or decreasing alignment length. In practice higher Q-scores indicate "better" matches far more often then RMSD and alignment length alone.

[b]**RMSD** stands for the Root Mean Square Deviation, calculated between Cα atoms of matched residues for the best 3D superposition of the query and target structures. RMSD is presented in angstroms. In simple words, RMSD gives you an idea how separated, at best 3D superposition, a "typical" pair of matched Cα atoms is.

Generally, the larger RMSD, the more distant the matched structures are.

[c]**N_{algn}** is a number of aligned residues.

Definitions **a**–**c** are taken from the PDBeFold (also known as SSM) server web page (http://www.ebi.ac.uk/msd-srv/ssm/ssmresults.html) (Krissinel and Henrick, 2004).

Appendix 6 – List of Formulae

Formula 1

$I_{hkl} \propto |F_{hkl}|^2$

$I_{hkl} = k_1 k_2 I_0 LPTE|F_{hkl}|^2$,

where I_0 is the intensity of the incident beam,

$k_1 = e^4/(m^2c^4)$ (includes constants from the Thomson formula),

$k_2 = \lambda^3\Omega/V^2$ (Ω is the volume of the crystal, V is the volume of the unit cell),

L is the Lorentz factor,

P is the polarization factor,

T is the transmission factor and depends on the capacity of the crystal to absorb X-rays,

E is the extinction coefficient,

F_{hkl} is the structure factor,

$$F_{hkl} = \sum_j f_j \exp(2\pi i(hx_j + ky_j + lz_j)) \exp(-B_j \sin^2\theta/\lambda^2)$$

and f_j is the atomic scattering factor; also see Formula 3 and a textbook by Giacovazzo et al., 2002 (pages 178–179).

Formula 2

$$\rho(xyz) = \frac{1}{V}\sum_{hkl} |F_{hkl}| \exp(i\alpha_{hkl}) \exp(-2\pi i(hx + ky + lz))$$

where $\rho(xyz)$ is the electron density at the position x, y, z in the unit cell,

V is the volume of the unit cell,

$|F_{hkl}|$ is the amplitude of the structure factor for each reflection hkl,

α_{hkl} is the phase for each reflection hkl.

Formula 3

$$F_{hkl} = \int_V \rho(xyz) \exp(2\pi i(hx + ky + lz)) dV$$

where F_{hkl} is the structure factor (also see Formula 1),

V is the volume of the unit cell,

$\rho(xyz)$ is the electron density at the position x, y, z in the unit cell.

Formula 4

$$R_{work} = \frac{\sum_{hkl \notin free} \| F_{obs} | - | F_{calc} \|}{\sum_{hkl \notin free} | F_{obs} |}$$

R_{work} is the cross-validation value between observed (F_{obs}; F_{obs} is calculated from the observed intensities, see Formula 1) and calculated (F_{calc}) structure factor amplitudes for the data set used in refinement. Typically, 95 % of all data is used in refinement; the remaining 5 % of data is not used in refinement directly but is used to calculate cross-validation value R_{free} (Formula 5; Brünger, 1992). Cross-validation and R-values used during model building and refinement are discussed in a textbook by B. Rupp, 2010 (Chapter 12).

Formula 5

$$R_{free} = \frac{\sum_{hkl \in free} \| F_{obs} | - | F_{calc} \|}{\sum_{hkl \in free} | F_{obs} |}$$

R_{free} is the cross-validation value for the excluded data set (the "free" data set) (Brünger, 1992; Rupp, 2010).

Formula 6

$$B_{iso} = 8\pi^2 \left\langle u_{iso}^2 \right\rangle$$

where B_{iso} is the isotropic displacement parameter or B-factor; $<u^2_{iso}>$ is the mean square isotropic displacement of the atom from its equilibrium or mean position (Formula 6 is taken from the book by Rupp, 2010, page 262).

Appendix 7 – List of abbreviations and symbols

°C	degrees Celsius
APC	adenomatous polyposis coli
ATP	adenosine-5'-triphosphate
β-TRCP1	β-transducin repeat-containing protein 1
Bicine	2-(Bis(2-hydroxyethyl)amino)acetic acid, $C_6H_{13}NO_4$
Bis-Tris	2-[Bis-(2-hydroxyethyl)-amino]-2-hydroxymethyl-propane-1,3-diol, $C_8H_{19}NO_5$
B_{max}	the SPR signal value (in RUs) obtained at the receptor saturation level
BMP	bone morphogenetic protein
CHAPS	a zwitterionic detergent, 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate, $C_{32}H_{58}N_2O_7S$
CHO	Chinese hamster ovary (cells)
CK1α	casein kinase 1 α
CK1γ	casein kinase 1 γ
CM	conditioned medium
CRD	cysteine-rich domain
Dkk	Dickkopf
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide, $(CH_3)_2SO$
DNA	deoxyribonucleic acid
DPPC	1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine, $C_{40}H_{80}NO_8P$
DTT	dithiothreitol, (2S,3S)-1,4-bis(sulfanyl)butane-2,3-diol, $C_4H_{10}O_2S_2$
Dvl	Dishevelled
e	the electric charge carried by a single proton, $1.602176565(35)\times10^{-19}$ C

EDTA	ethylenediaminetetraacetic acid, $C_{10}H_{16}N_2O_8$
EGF	epidermal growth factor
ER	endoplasmic reticulum
EXT1	Exostosin-1, glucuronosyl-N-acetylglucosaminyl-proteoglycan/N-acetylglucosaminyl-proteoglycan 4-alpha-N-acetylglucosaminyltransferase, E.C. 2.4.1.224
FCS	foetal calf serum
FGF	fibroblast growth factors
g	acceleration of gravity at the Earth's surface, g = 9.81 m/s^2
G-418	an aminoglycoside antibiotic, also known as Geneticin, $C_{20}H_{40}N_4O_{10}$
GAG	glycosaminoglycans
GlcNAc	N-acetyl-D-glucosamine, 2-(acetylamino)-2-deoxy-D-glucose
Glycerol	propan-1,2,3-triol
GPC3	Glypican 3
GPI	glycosylphosphatidylinositol
GSK3β	glycogen synthase kinase 3 β
HDL	high-density lipoproteins
HEK	human embryonic kidney (cells)
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid, $C_8H_{18}N_2O_4S$
HPLC	high-performance liquid chromatography, high-pressure liquid chromatography
HRP	horseradish peroxidase
HS	heparan sulfate
HSPG	heparan sulfate proteoglycan
Ig	immunoglobulin

IMAC	immobilized metal ion affinity chromatography
k	Boltzmann constant, $1.3806488(13)\times10^{-23}$ J/K
K_d	dissociation constant
kDa	kiloDalton, 1 Da = $1.660538921(73)\times10^{-27}$ kg
LB	lysogeny broth
LEF	lymphoid enhancer-binding factor
LRP	low-density lipoprotein receptor-related protein
MALDI-TOF-MS	matrix assisted laser desorption ionisation time-of-flight mass spectrometry
MALLS	multi-angle laser light scattering
MES	2-(N-morpholino)ethanesulfonic acid, $C_6H_{13}NO_4S$
MuSK	muscle-specific kinase
MW	molecular weight
NMR	nuclear magnetic resonance
PBS	phosphate buffered saline
PC	phosphocholine
PCP	planar cell polarity
PCR	polymerase chain reaction
PDB	Protein Data Bank
RMSD	root-mean-square deviation
RNA	ribonucleic acid
ROR	tyrosine-protein kinase transmembrane receptor ROR1
RU	response units
RYK	related to receptor tyrosine kinase
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis

SEC	size-exclusion chromatography
sFRP	secreted Frizzled-related protein
SPR	surface plasmon resonance
TCF	T cell specific transcription factor
Tris	tris(hydroxymethyl)aminomethane, $(HOCH_2)_3CNH_2$
Triton X-100	a nonionic surfactant, $C_{14}H_{22}O(C_2H_4O)_n$ (n = 9-10)
Tween 20	a polysorbate surfactant, polyoxyethylene (20) sorbitan monolaurate
WD	WIF domain
WIF-1	Wnt inhibitory factor 1
Wnt	an acronym for the two founding members of a family, Wingless from *Drosophila* and Integrated-1 from mouse

References

- Aberle, H., Bauer, A., Stappert, J., Kispert, A., and Kemler, R. (1997) β-catenin is a target for the ubiquitin-proteasome pathway. EMBO J. 16, 3797–3804.
- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., and Zwart, P.H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221.
- Adamska, M., Larroux, C., Adamski, M., Green, K., Lovas, E., Koop, D., Richards, G.S., Zwafink, C., and Degnan, B.M. (2010) Structure and expression of conserved Wnt pathway components in the demosponge Amphimedon queenslandica. Evol. Dev. 12, 494–518.
- Adell, T., Thakur, A.N., and Müller, W.E. (2007) Isolation and characterization of Wnt pathway-related genes from Porifera. Cell Biol. Int. 31, 939–949.
- Ahn, J., Lüdecke, H.J., Lindow, S., Horton, W.A., Lee, B., Wagner, M.J., Horsthemke, B., and Wells, D.E. (1995) Cloning of the putative tumour suppressor gene for hereditary multiple exostoses (EXT1). Nat Genet. 11, 137–143.
- Ahn, V.E., Chu, M.L., Choi, H.J., Tran, D., Abo, A., and Weis, W.I. (2011) Structural basis of Wnt signaling inhibition by Dickkopf binding to LRP5/6. Dev. Cell 21, 862–873.
- Ai, L., Tao, Q., Zhong, S., Fields, C.R., Kim, W.J., Lee, M.W., Cui, Y., Brown, K.D., and Robertson, K.D. (2006) Inactivation of Wnt inhibitory factor-1 (WIF1) expression by epigenetic silencing is a common event in breast cancer. Carcinogenesis 27, 1341–1348.
- Amit, S., Hatzubai, A., Birman, Y., Andersen, J.S., Ben-Shushan, E., Mann, M., Ben-Neriah, Y., and Alkalay, I. (2002) Axin-mediated CKI phosphorylation of β-catenin at Ser 45: a molecular switch for the Wnt pathway. Genes Dev. 16, 1066–1076.

- Aricescu, A.R., Lu, W., and Jones, E.Y. (2006) A time- and cost-efficient system for high-level protein production in mammalian cells. Acta Crystallogr. D Biol. Crystallogr. 62, 1243–1250.
- Baeg, G.H., Lin, X., Khare, N., Baumgartner, S., and Perrimon, N. (2001) Heparan sulfate proteoglycans are critical for the organization of the extracellular distribution of Wingless. Development 128, 87–94.
- Bafico, A., Gazit, A., Pramila, T., Finch, P.W., Yaniv, A., and Aaronson, S.A. (1999) Interaction of frizzled related protein (FRP) with Wnt ligands and the frizzled receptor suggests alternative mechanisms for FRP inhibition of Wnt signaling. J. Biol. Chem. 274. 16180–16187.
- Batra, S., Shi, Y., Kuchenbecker, K.M., He, B., Reguart, N., Mikami, I., You, L., Xu, Z., Lin, Y.C., Clément, G., and Jablons, D.M. (2006) Wnt inhibitory factor-1, a Wnt antagonist, is silenced by promoter hypermethylation in malignant pleural mesothelioma. Biochem. Biophys. Res. Commun. 342, 1228–1232.
- Bazan, J.F. and de Sauvage, F.J. (2009) Structural ties between cholesterol transport and morphogen signaling. Cell 138, 1055–1056.
- Bhanot, P., Brink, M., Samos, C.H., Hsieh, J.C., Wang, Y., Macke, J.P., Andrew, D., Nathans, J., and Nusse, R. (1996) A new member of the frizzled family from *Drosophila* functions as a Wingless receptor. Nature. 382, 225–230.
- Bhat, R.A., Stauffer, B., Komm, B.S., and Bodine, P.V. (2007) Structure-function analysis of secreted frizzled-related protein-1 for its Wnt antagonist function. J. Cell. Biochem. 102, 1519–1528.
- Biben, C., Stanley, E., Fabri, L., Kotecha, S., Rhinn, M., Drinkwater, C., Lah, M., Wang, C.C., Nash, A., Hilton, D., Ang, S.L., Mohun, T., and Harvey, R.P. (1998) Murine cerberus homologue mCer-1: a candidate anterior patterning molecule. Dev. Biol. 194, 135–

151.

- Binari, R.C., Staveley, B.E., Johnson, W.A., Godavarti, R., Sasisekharan, R., and Manoukian, A.S. (1997) Genetic evidence that heparin-like glycosaminoglycans are involved in wingless signaling. Development 124, 2623–2632.

- Bishop, B., Aricescu, A.R., Harlos, K., O'Callaghan, C.A., Jones, E.Y., and Siebold, C. (2008) Structural insights into hedgehog ligand sequestration by the human hedgehog-interacting protein HHIP. Nat. Struct. Mol. Biol. 16, 698–703.

- Blake, J.A., Bult, C.J., Kadin, J.A., Richardson, J.E., Eppig, J.T., Mouse Genome Database Group. (2011) The Mouse Genome Database (MGD): premier model organism resource for mammalian genomics and genetics. Nucleic Acids Res. 39 (Database issue), D842–D848.

- Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S.M., and Bricogne, G. (2004) Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. Acta Crystallogr. D Biol. Crystallogr. 60, 2210–2221.

- Bork, P., Downing, A.K., Kieffer, B., and Campbell, I.D. (1996) Structure and distribution of modules in extracellular proteins. Q. Rev. Biophys. 29, 119–167.

- Bourhis, E., Tam, C., Franke, Y., Bazan, J.F., Ernst, J., Hwang, J., Costa, M., Cochran, A.G., and Hannoush, R.N. (2010) Reconstitution of a frizzled8 · Wnt3a · LRP6 signaling complex reveals multiple Wnt and Dkk1 binding sites on LRP6. J. Biol. Chem. 285, 9172–9179.

- Bouwmeester, T., Kim, S., Sasai, Y., Lu, B., and De Robertis, E.M. (1996) Cerberus is a head-inducing secreted factor expressed in the anterior endoderm of Spemann's organizer. Nature 382, 595–601.

- Brannon, M., Gomperts, M., Sumoy, L., Moon, R.T., and Kimelman, D. (1997) A β-catenin/XTcf-3 complex binds to the siamois promoter to regulate dorsal axis specification in

Xenopus. Genes Dev. 11, 2359–2370.

- Brünger, A.T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355, 472–475.
- Cadigan, K.M. and Nusse, R. (1997) Wnt signaling: a common theme in animal development. Genes Dev. 11, 3286–2305.
- Cajanek, L., Adlerz, L., Bryja, V., and Arenas, E. (2010) WNT unrelated activities in commercially available preparations of recombinant WNT3a. J. Cell. Biochem. 111, 1077–1079.
- Camejo, G., Hurt-Camejo, E., Wiklund, O., and Bondjers, G. (1998) Association of apo B lipoproteins with arterial proteoglycans: pathological significance and molecular basis. Atherosclerosis 139, 205–222.
- Capurro, M.I., Xu, P., Shi, W., Li, F., Jia, A., and Filmus, J. (2008) Glypican-3 inhibits Hedgehog signaling during development by competing with patched for Hedgehog binding. Dev. Cell. 14, 700–711.
- Carron, C., Pascal, A., Djiane, A., Boucaut, J.C., Shi, D.L., and Umbhauer, M. (2003) Frizzled receptor dimerization is sufficient to activate the Wnt/beta-catenin pathway. J. Cell Sci. 116, 2541–2550.
- Chan, S.L., Cui, Y., van Hasselt, A., Li, H., Srivastava, G., Jin, H., Ng, K.M., Wang, Y., Lee, K.Y., Tsao, G.S., Zhong, S., Robertson, K.D., Rha, S.Y., Chan, A.T., and Tao, Q. (2007) The tumor suppressor Wnt inhibitory factor 1 is frequently methylated in nasopharyngeal and esophageal carcinomas. Lab. Invest. 87, 644–650.
- Chang, V.T., Crispin, M., Aricescu, A.R., Harvey, D.J., Nettleship, J.E., Fennelly, J.A., Yu, C., Boles, K.S., Evans, E.J., Stuart, D.I., Dwek, R.A., Jones, E.Y., Owens, R.J., and Davis, S.J. (2007) Glycoprotein structural genomics: solving the glycosylation problem. Structure 15, 267–273.

- Chao, J., Tillman, D.M., Wang, M.Y., Margolius, H.S., and Chao, L. (1986) Identification of a new tissue-kallikrein-binding protein. Biochem. J. 239, 325–331.
- Chen, L., Wang, K., Shao, Y., Huang, J., Li, X., Shan, J., Wu, D., and Zheng, J.J. (2008) Structural insight into the mechanisms of Wnt signaling antagonism by Dkk. J. Biol. Chem. 283, 23364–23370.
- Chen, S., Bubeck, D., MacDonald, B.T., Liang, W.X., Mao, J.H., **Malinauskas, T.**, Llorca, O., Aricescu, A.R., Siebold, C., He, X., and Jones, E.Y. (2011) Structural and functional studies of LRP6 ectodomain reveal a platform for Wnt signaling. Dev. Cell. 21, 848–861.
- Cheng, Z., Biechele, T., Wei, Z., Morrone, S., Moon, R.T., Wang, L., and Xu, W. (2011) Crystal structures of the extracellular domain of LRP6 and its complex with DKK1. Nat. Struct. Mol. Biol. 18, 1204–1210.
- Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. (2003) Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. 31, 3497–3500.
- Chien, A.J., Conrad, W.H., and Moon, R.T. (2009) A Wnt survival guide: from flies to human disease. J. Invest. Dermatol. 129, 1614–1627.
- Chim, C.S., Chan, W.W., Pang, A., and Kwong, Y.L. (2006) Preferential methylation of Wnt inhibitory factor-1 in acute promyelocytic leukemia: an independent poor prognostic factor. Leukemia 20, 907–909.
- Ching, W., Hang, H.C., and Nusse, R. (2008) Lipid-independent secretion of a *Drosophila* Wnt protein. J. Biol. Chem. 283, 17092–17098.
- Chuman, Y., Üren, A., Cahill, J., Regan, C., Wolf, V., Kay, B.K., and Rubin, J.S. (2004) Identification of a peptide binding motif for secreted frizzled-related protein-1. Peptides 25, 1831–1838.

- Clevers, H. (2006) Wnt/β-catenin signaling in development and disease. Cell 127, 469–480.
- Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. 16, 10881–10890.
- Dann, C.E., Hsieh, J.C., Rattner, A., Sharma, D., Nathans, J., and Leahy, D.J. (2001) Insights into Wnt binding and signalling from the structures of two Frizzled cysteine-rich domains. Nature 412, 86–90.
- Darwin, C.G. (1922) The reflection of X-rays from imperfect crystals. Phil. Mag. 43, 800–829.
- DasGupta, R., Kaykas, A., Moon, R.T., and Perrimon, N. (2005) Functional genomic analysis of the Wnt-wingless signaling pathway. Science. 308, 826–833.
- Davidson, G., Wu, W., Shen, J., Bilic, J., Fenger, U., Stannek, P., Glinka, A., and Niehrs, C. (2005) Casein kinase 1 gamma couples Wnt receptor activation to cytoplasmic signal transduction. Nature 438, 867–872.
- Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B. 3rd, Snoeyink, J., Richardson, J.S., Richardson, D.C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 35, W375–383.
- De Ferrari, G.V. and Moon, R.T. (2006) The ups and downs of Wnt signaling in prevalent neurological disorders. Oncogene 25, 7545–7553.
- DeAlmeida, V.I., Miao, L., Ernst, J.A., Koeppen, H., Polakis, P., and Rubinfeld, B. (2007) The soluble wnt receptor Frizzled8CRD-hFc inhibits the growth of teratocarcinomas in vivo. Cancer Res. 67, 5371–5379.
- Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical

mapping of functionally annotated residues. Nucleic Acids Res. 34, W116–W118.

- Ellies, D.L., Viviano, B., McCarthy, J., Rey, J.P., Itasaki, N., Saunders, S., and Krumlauf, R. (2006) Bone density ligand, Sclerostin, directly interacts with LRP5 but not LRP5G171V to modulate Wnt activity. J. Bone Miner. Res. 21, 1738–1749.

- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010) Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501.

- Esko, J.D., Stewart, T.E., and Taylor, W.H. (1985) Animal cell mutants defective in glycosaminoglycan biosynthesis. Proc. Natl. Acad. Sci. U. S. A. 82, 3197–3201.

- Evans, P. (2006) Scaling and assessment of data quality. Acta Crystallogr. D Biol. Crystallogr. 62, 72–82.

- Fedi, P., Bafico, A., Nieto Soria, A., Burgess, W.H., Miki, T., Bottaro, D.P., Kraus, M.H., and Aaronson, S.A. (1999) Isolation and biochemical characterization of the human Dkk-1 homologue, a novel inhibitor of mammalian Wnt signaling. J. Biol. Chem. 274, 19465–19472.

- Felsenstein, J. (2004) Inferring Phylogenies. Published by Sinauer Associates Inc., U.S.A, ISBN 0878931775.

- Fiedler, M., Mendoza-Topaz, C., Rutherford, T.J., Mieszczanek, J., and Bienz, M. (2011) Dishevelled interacts with the DIX domain polymerization interface of Axin to interfere with its function in down-regulating β-catenin. Proc. Natl. Acad. Sci. U. S. A. 108, 1937–1942.

- Finch, P.W., He, X., Kelley, M.J., Üren, A., Schaudies, R.P., Popescu, N.C., Rudikoff, S., Aaronson, S.A., Varmus, H.E., and Rubin, J.S. (1997) Purification and molecular cloning of a secreted, Frizzled-related antagonist of Wnt action. Proc. Natl. Acad. Sci. U. S. A. 94, 6770–6775.

- Firth, S.M. and Baxter, R.C. (2002) Cellular actions of the insulin-like growth factor

binding proteins. Endocr. Rev. 23, 824–854.

- Fradkin, L.G., Dura, J.M., and Noordermeer, J.N. (2010) Ryks: new partners for Wnts in the developing and regenerating nervous system. Trends Neurosci. 33, 84–92.
- Gallet, A., Staccini-Lavenant, L., and Thérond, P.P. (2008) Cellular trafficking of the glypican Dally-like is required for full-strength Hedgehog signaling and wingless transcytosis. Dev. Cell. 14, 712–725.
- Gao, G., Shao, C., Zhang, S.X., Dudley, A., Fant, J., and Ma, J.X. (2003) Kallikrein-binding protein inhibits retinal neovascularization and decreases vascular leakage. Diabetologia 46, 689–698.
- Garber, K. (2009) Drugging the Wnt pathway: problems and progress. J. Natl. Cancer Inst. 101, 548–550.
- Garman, E.F. (2010) Radiation damage in macromolecular crystallography: what is it and why should we care? Acta Crystallogr. D Biol. Crystallogr. 66, 339–351.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. John M. Walker: The Proteomics Protocols Handbook (Humana Press) 571–607.
- Gerhard, D.S., Wagner, L., Feingold, E.A., Shenmen, C.M., Grouse, L.H., Schuler, G., Klein, S.L., Old, S., Rasooly, R., Good, P., Guyer, M., Peck, A.M., Derge, J.G., Lipman, D., Collins, F.S., Jang, W., Sherry, S., Feolo, M., Misquitta, L., Lee, E., Rotmistrovsky, K., Greenhut, S.F., Schaefer, C.F., Buetow, K., Bonner, T.I., Haussler, D., Kent, J., Kiekhaus, M., Furey, T., Brent, M., Prange, C., Schreiber, K., Shapiro, N., Bhat, N.K., Hopkins, R.F., Hsie, F., Driscoll, T., Soares, M.B., Casavant, T.L., Scheetz, T.E., Brown-stein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Piao, Y., Dudekula, D.B., Ko, M.S., Kawakami, K., Suzuki, Y., Sugano, S., Gruber, C.E., Smith, M.R., Simmons, B., Moore, T., Waterman, R., Johnson, S.L., Ruan, Y., Wei, C.L., Mathavan, S., Gunaratne, P.H., Wu, J., Garcia, A.M., Hulyk, S.W.,

Fuh, E., Yuan, Y., Sneed, A., Kowis, C., Hodgson, A., Muzny, D.M., McPherson, J., Gibbs, R.A., Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madari, A., Young, A.C., Wetherby, K.D., Granite, S.J., Kwong, P.N., Brinkley, C.P., Pearson, R.L., Bouffard, G.G., Blakesly, R.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Griffith, M., Griffith, O.L., Krzywinski, M.I., Liao, N., Morin, R., Palmquist, D., Petrescu, A.S., Skalska, U., Smailus, D.E., Stott, J.M., Schnerch, A., Schein, J.E., Jones, S.J., Holt, R.A., Baross, A., Marra, M.A., Clifton, S., Makowski, K.A., Bosak, and S., Malek, J.; MGC Project Team. (2004) The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). Genome Res. 14, 2121–2127.

- Giacovazzo, C., Monaco, H.L., Artioli, G., Viterbo, D., Ferraris, G., Gilli, G., Zanotti, G. and Catti, M. (2002) Fundamentals of Crystallography (second edition, Oxford University Press).
- Gilbert, S.F. (2000) Developmental Biology, 6th edition (Sinauer Associates; ISBN-10: 0-87893-243-7), Chapter 14.
- Giráldez, A.J., Copley, R.R., and Cohen, S.M. (2002) HSPG modification by the secreted enzyme Notum shapes the Wingless morphogen gradient. Dev. Cell 2, 667–676.
- Glinka, A., Wu, W., Delius, H., Monaghan, A.P., Blumenstock, C., and Niehrs, C. (1998) Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. Nature 391, 357–362.
- Glise, B., Miller, C.A., Crozatier, M., Halbisen, M.A., Wise, S., Olson, D.J., Vincent, A., and Blair, S.S. (2005) Shifted, the *Drosophila* ortholog of Wnt inhibitory factor-1, controls the distribution and movement of Hedgehog. Dev. Cell 8, 255–266.
- Gorfinkiel, N., Sierra, J., Callejo, A., Ibañez, C., and Guerrero, I. (2005) The *Drosophila* ortholog of the human Wnt inhibitor factor Shifted controls the diffusion of lipid-

modified Hedgehog. Dev. Cell 8, 241–253.

- Gouet, P., Courcelle, E., Stuart, D.I., and Métoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15, 305–308.
- Gupta, R., Jung, E., and Brunak, S. (2004) Prediction of N-glycosylation sites in human proteins. In preparation.
- Gurdon, J.B. and Bourillot, P.Y. (2001) Morphogen gradient interpretation. Nature 413, 797–803.
- Gurley, K.A., Rink, J.C., and Sánchez Alvarado, A. (2008) β-catenin defines head versus tail identity during planarian regeneration and homeostasis. Science 319, 323–327.
- Häcker, U., Nybakken, K., and Perrimon, N. (2005) Heparan sulphate proteoglycans: the sweet side of development. Nat. Rev. Mol. Cell Biol. 6, 530–541.
- Häusler, K.D., Horwood, N.J., Chuman, Y., Fisher, J.L., Ellis, J., Martin, T.J., Rubin, J.S., and Gillespie, M.T. (2004) Secreted frizzled-related protein-1 inhibits RANKL-dependent osteoclast formation. J. Bone Miner. Res. 19, 1873–1881.
- Hell, S.W. (2003) Toward fluorescence nanoscopy. Nat. Biotechnol. 21, 1347–1355.
- Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77, 51–59.
- Hoang, B., Moos, M. Jr., Vukicevic, S., and Luyten, F.P. (1996) Primary structure and tissue distribution of FRZB, a novel protein related to *Drosophila* frizzled, suggest a role in skeletal morphogenesis. J. Biol. Chem. 271, 26131–26137.
- Hoang, B.H., Thomas, J.T., Abdul-Karim, F.W., Correia, K.M., Conlon, R.A., Luyten, F.P., and Ballock, R.T. (1998) Expression pattern of two Frizzled-related genes, *Frzb-1* and *Sfrp-1*, during mouse embryogenesis suggests a role for modulating action of Wnt family members. Dev. Dyn. 212, 364–372.
- Holm, L. and Rosenström, P. (2010) Dali server: conservation mapping in 3D.

Nucleic Acids Res. 38, W545–W549.

- Holton, J.M. and Frankel, K.A. (2010) The minimum crystal size needed for a complete diffraction data set. Acta Crystallogr. D Biol. Crystallogr. 66, 393–408.
- Hsieh, J.C., Kodjabachian, L., Rebbert, M.L., Rattner, A., Smallwood, P.M., Samos, C.H., Nusse, R., Dawid, I.B., and Nathans, J. (1999) A new secreted protein that binds to Wnt proteins and inhibits their activities. Nature 398, 431–436.
- Hsin, K., Sheng, Y., Harding, M.M., Taylor, P., and Walkinshaw, M.D. (2008) MESPEUS: a database of the geometry of metal sites in proteins. J. Appl. Cryst. 41, 963–968.
- Hsu, W., Zeng, L., and Costantini, F. (1999) Identification of a domain of Axin that binds to the serine/threonine protein phosphatase 2A and a self-binding domain. J. Biol. Chem. 274, 3439–3445.
- Huang, B., Babcock, H., and Zhuang, X. (2010) Breaking the diffraction barrier: super-resolution imaging of cells. Cell 143, 1047–1058.
- Huang, S.M., Mishina, Y.M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G.A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., Hild, M., Shi, X., Wilson, C.J., Mickanin, C., Myer, V., Fazal, A., Tomlinson, R., Serluca, F., Shao, W., Cheng, H., Shultz, M., Rau, C., Schirle, M., Schlegl, J., Ghidelli, S., Fawell, S., Lu, C., Curtis, D., Kirschner, M.W., Lengauer, C., Finan, P.M., Tallarico, J.A., Bouwmeester, T., Porter, J.A., Bauer, A., and Cong, F. (2009) Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. Nature 461, 614–620.
- Ikeda, S., Kishida, S., Yamamoto, H., Murai, H., Koyama, S., and Kikuchi, A. (1998) Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3β and β-catenin and promotes GSK-3β-dependent phosphorylation of β-catenin. EMBO J. 17, 1371–1384.
- Ito, M., Yang, Z., Andl, T., Cui, C., Kim, N., Millar, S.E., and Cotsarelis, G. (2007)

Wnt-dependent de novo hair follicle regeneration in adult mouse skin after wounding. Nature 447, 316–320.

- Jones, S. and Thornton, J.M. (1996) Principles of protein-protein interactions. Proc. Natl. Acad. Sci. U. S. A. 93, 13–20.
- Jope, R.S., Yuskaitis, C.J., and Beurel, E. (2007) Glycogen synthase kinase-3 (GSK3): inflammation, diseases, and therapeutics. Neurochem. Res. 32, 577–595.
- Julenius, K., Mølgaard, A., Gupta, R., and Brunak, S. (2005) Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. Glycobiology 15, 153–164.
- Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D Biol. Crystallogr. 66, 133–144.
- Kansara, M., Tsang, M., Kodjabachian, L., Sims, N.A., Trivett, M.K., Ehrich, M., Dobrovic, A., Slavin, J., Choong, P.F., Simmons, P.J., Dawid, I.B., and Thomas, D.M. (2009) Wnt inhibitory factor 1 is epigenetically silenced in human osteosarcoma, and targeted disruption accelerates osteosarcomagenesis in mice. J. Clin. Invest. 119, 837–851.
- Kantardjieff, K.A. and Rupp, B. (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. Protein Sci. 12, 1865–1871.
- Kawakami, K., Hirata, H., Yamamura, S., Kikuno, N., Saini, S., Majid, S., Tanaka, Y., Kawamoto, K., Enokida, H., Nakagawa, M., and Dahiya, R. (2009) Functional significance of Wnt inhibitory factor-1 gene in kidney cancer. Cancer Res. 69, 8603–8610.
- Keep, N.H., Barnes, M., Barsukov, I., Badii, R., Lian, L.Y., Segal, A.W., Moody, P.C., and Roberts, G.C. (1997) A modulator of rho family G proteins, rhoGDI, binds these G proteins via an immunoglobulin-like domain and a flexible N-terminal arm. Structure 5, 623–633.

- Kim, J., You, L., Xu, Z., Kuchenbecker, K., Raz, D., He, B., and Jablons, D. (2007) Wnt inhibitory factor inhibits lung cancer cell growth. J. Thorac. Cardiovasc. Surg. 133, 733–737.
- Kim, J.B., Leucht, P., Lam, K., Luppen, C., Ten Berge, D., Nusse, R., and Helms, J.A. (2007) Bone regeneration is regulated by wnt signaling. J. Bone Miner. Res. 22, 1913–1923.
- Kimelman, D. and Xu, W. (2006) β-catenin destruction complex: insights and questions from a structural perspective. Oncogene 25, 7482–7491.
- Kinzler, K.W., Nilbert, M.C., Su, L.K., Vogelstein, B., Bryan, T.M., Levy, D.B., Smith, K.J., Preisinger, A.C., Hedge, P., McKechnie, D., et al. (1991) Identification of FAP locus genes from chromosome 5q21. Science 253, 661–665.
- Kleywegt, G.J. (1996) Use of non-crystallographic symmetry in protein structure refinement. Acta Crystallogr. D Biol. Crystallogr. 52, 842–857.
- Kleywegt, G.J. and Jones, T.A. (1996) Phi/psi-chology: Ramachandran revisited. Structure 4, 1395–1400.
- Komekado, H., Yamamoto, H., Chiba, T., and Kikuchi, A. (2007) Glycosylation and palmitoylation of Wnt-3a are coupled to produce an active form of Wnt-3a. Genes Cells 12, 521–534.
- Kraft, P., Bergamaschi, A., Broennimann, Ch., Dinapoli, R., Eikenberry, E.F., Henrich, B., Johnson, I., Mozzanica, A., Schlepütz, C.M., Willmott, P.R., and Schmitt, B. (2009) Performance of single-photon-counting PILATUS detector modules. J. Synchrotron. Radiat. 16, 368–375.
- Krishnan, V., Bryant, H.U., and Macdougald, O.A. (2006) Regulation of bone mass by Wnt signaling. J. Clin. Invest. 116, 1202–1209.
- Krissinel, E. and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr.

60, 2256–2268.

- Kühl, M., Sheldahl, L.C., Park, M., Miller, J.R., and Moon, R.T. (2000) The Wnt/Ca^{2+} pathway: a new vertebrate Wnt signaling pathway takes shape. Trends Genet. 16, 279–283.
- Kuntal, B.K., Aparoy, P., and Reddanna, P. (2010) EasyModeller: A graphical interface to MODELLER. BMC Res. Notes 3, 226.
- Lapraz, F., Röttinger, E., Duboc, V., Range, R., Duloquin, L., Walton, K., Wu, S.Y., Bradham, C., Loza, M.A., Hibino, T., Wilson, K., Poustka, A., McClay, D., Angerer, L., Gache, C., and Lepage, T. (2006) RTK and TGF-β signaling pathways genes in the sea urchin genome. Dev. Biol. 300, 132–152.
- Laskowski, R.A. (2009) PDBsum new things. Nucleic Acids Res. 37, D355–D359.
- Lee, E., Salic, A., Krüger, R., Heinrich, R., and Kirschner, M.W. (2003) The roles of APC and Axin derived from experimental and theoretical analysis of the Wnt pathway. PLoS Biol. 1, E10.
- Leslie, A.G. (2006) The integration of macromolecular diffraction data. Acta Crystallogr. D Biol. Crystallogr. 62, 48–57.
- Leyns, L., Bouwmeester, T., Kim, S.H., Piccolo, S., and De Robertis, E.M. (1997) Frzb-1 is a secreted antagonist of Wnt signaling expressed in the Spemann organizer. Cell. 88, 747–756.
- Li, X., Ominsky, M.S., Niu, Q.T., Sun, N., Daugherty, B., D'Agostin, D., Kurahara, C., Gao, Y., Cao, J., Gong, J., Asuncion, F., Barrero, M., Warmington, K., Dwyer, D., Stolina, M., Morony, S., Sarosi, I., Kostenuik, P.J., Lacey, D.L., Simonet, W.S., Ke, H.Z., and Paszty, C. (2008) Targeted deletion of the sclerostin gene in mice results in increased bone formation and bone strength. J. Bone Miner. Res. 23, 860–869.
- Li, X., Zhang, Y., Kang, H., Liu, W., Liu, P., Zhang, J., Harris, S.E., and Wu, D. (2005) Sclerostin binds to LRP5/6 and antagonizes canonical Wnt signaling. J. Biol. Chem.

280, 19883–19887.

- Liepinsh, E., Bányai, L., Patthy, L., and Otting, G. (2006) NMR structure of the WIF domain of the human Wnt-inhibitory factor-1. J. Mol. Biol. 357, 942–950.
- Lin, K., Wang, S., Julius, M.A., Kitajewski, J., Moos, M. Jr., and Luyten, F.P. (1997) The cysteine-rich frizzled domain of Frzb-1 is required and sufficient for modulation of Wnt signaling. Proc. Natl. Acad. Sci. U. S. A. 94, 11196–11200.
- Lin, X. and Perrimon, N. (1999) Dally cooperates with *Drosophila* Frizzled 2 to transduce Wingless signalling. Nature 400, 281–284.
- Liu, C., Li, Y., Semenov, M., Han, C., Baeg, G.H., Tan, Y., Zhang, Z., Lin, X., and He, X. (2002) Control of β-catenin phosphorylation/degradation by a dual-kinase mechanism. Cell 108, 837–847.
- Logan, C.Y. and Nusse, R. (2004) The Wnt signaling pathway in development and disease. Annu. Rev. Cell Dev. Biol. 20, 781–810.
- Lopez-Rios, J., Esteve, P., Ruiz, J.M., and Bovolenta, P. (2008) The Netrin-related domain of Sfrp1 interacts with Wnt ligands and antagonizes their activity in the anterior neural plate. Neural Dev. 3, 19.
- Lu, W.G., Yamamoto, V., Ortega, B., and Baltimore, D. (2004) Mammalian Ryk is a Wnt coreceptor required for stimulation of neurite outgrowth. Cell 119, 97–108.
- MacDonald, B.T., Tamai, K., and He, X. (2009) Wnt/β-catenin signaling: components, mechanisms, and diseases. Dev. Cell 17, 9–26.
- **Malinauskas, T.** (2008) Docking of fatty acids into the WIF domain of the human Wnt inhibitory factor-1. Lipids 43, 227–230.
- **Malinauskas, T.**, Aricescu, A.R., Lu, W., Siebold, C., and Jones, E.Y. (2011) Modular mechanism of Wnt signaling inhibition by Wnt inhibitory factor 1. Nat. Struct. Mol. Biol. 18, 886–893.

- Martí-Renom, M.A., Stuart, A.C., Fiser, A., Sánchez, R., Melo, F., and Sali, A. (2000) Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct. 29, 291–325.
- Mayo, C.J., Diprose, J.M., Walter, T.S., Berry, I.M., Wilson, J., Owens, R.J., Jones, E.Y., Harlos, K., Stuart, D.I., and Esnouf, R.M. (2005) Benefits of automated crystallization plate tracking, imaging, and analysis. Structure 13, 175–182.
- Mazieres, J., He, B., You, L., Xu, Z., Lee, A.Y., Mikami, I., Reguart, N., Rosell, R., McCormick, F., and Jablons, D.M. (2004) Wnt inhibitory factor-1 is silenced by promoter hypermethylation in human lung cancer. Cancer Res. 64, 4717–4720.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007) Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674.
- Meitinger, T., Meindl, A., Bork, P., Rost, B., Sander, C., Haasemann, M., and Murken, J. (1993) Molecular modelling of the Norrie disease protein predicts a cystine knot growth factor tertiary structure. Nat. Genet. 5, 376–380.
- Moreira, I.S., Fernandes, P.A., and Ramos, M.J. (2007) Hot spots-a review of the protein-protein interface determinant amino-acid residues. Proteins 68, 803–812.
- Mukhopadhyay, M., Shtrom, S., Rodriguez-Esteban, C., Chen, L., Tsukui, T., Gomer, L., Dorward, D.W., Glinka, A., Grinberg, A., Huang, S.P., Niehrs, C., Izpisúa Belmonte, J.C. and Westphal, H. (2001) Dickkopf1 is required for embryonic head induction and limb morphogenesis in the mouse. Dev. Cell 1, 423–434.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb. Symp. Quant. Biol. 51, 263–273.
- Munemitsu, S., Albert, I., Souza, B., Rubinfeld, B., and Polakis, P. (1995) Regulation of intracellular β-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor

protein. Proc. Natl. Acad. Sci. U. S. A. 92, 3046–3050.

- Nakamura, F., Pudas, R., Heikkinen, O., Permi, P., Kilpeläinen, I., Munday, A.D., Hartwig, J.H., Stossel, T.P., and Ylänne, J. (2006) The structure of the GPIb-filamin A complex. Blood 107, 1925–1932.

- Nakato, H., Futch, T.A., and Selleck, S.B. (1995) The division abnormally delayed (dally) gene: a putative integral membrane proteoglycan required for cell division patterning during postembryonic development of the nervous system in *Drosophila*. Development 121, 3687–3702.

- Nakaya, N., Lee, H.S., Takada, Y., Tzchori, I., and Tomarev, S.I. (2008) Zebrafish olfactomedin 1 regulates retinal axon elongation in vivo and is a modulator of Wnt signaling pathway. J. Neurosci. 28, 7900–7910.

- Neumann, C.J. and Cohen, S.M. (1997) Long-range action of Wingless organizes the dorsal-ventral axis of the *Drosophila* wing. Development 124, 871–880.

- Neumann, S., Coudreuse, D.Y., van der Westhuyzen, D.R., Eckhardt, E.R., Korswagen, H.C., Schmitz, G., and Sprong, H. (2009) Mammalian Wnt3a is released on lipoprotein particles. Traffic 10, 334–343.

- Niehrs, C. (2010) On growth and form: a Cartesian coordinate system of Wnt and BMP signaling specifies bilaterian body axes. Development 137, 845–857.

- Nishisho, I., Nakamura, Y., Miyoshi, Y., Miki, Y., Ando, H., Horii, A., Koyama, K., Utsunomiya, J., Baba, S., and Hedge, P. (1991) Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. Science 253, 665–669.

- Nusse, R. and Varmus, H.E. (1992) Wnt genes. Cell 69, 1073–1087.

- Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326.

- Padhi, D., Jang, G., Stouch, B., Fang, L., and Posvar, E. (2011) Single-dose, placebo-

controlled, randomized study of AMG 785, a sclerostin monoclonal antibody. J. Bone Miner. Res. 26, 19–26.

- Panáková, D., Sprong, H., Marois, E., Thiele, C., and Eaton, S. (2005) Lipoprotein particles are required for Hedgehog and Wingless signalling. Nature 435, 58–65.
- Patthy, L. (2000) The WIF module. Trends Biochem. Sci. 25, 1213.
- Pearce, J.J., Penny, G., and Rossant, J. (1999) A mouse cerberus/Dan-related gene family. Dev. Biol. 209, 98–110.
- Petersen, C.P. and Reddien, P.W. (2008) Smed-βcatenin-1 is required for anteroposterior blastema polarity in planarian regeneration. Science. 319, 327–330.
- Pfeffer, P.L., De Robertis, E.M., and Izpisua-Belmonte, J.C. (1997) Crescent, a novel chick gene encoding a Frizzled-like cysteine-rich domain, is expressed in anterior regions during early embryogenesis. Int. J. Dev. Biol. 41, 449–458.
- Phillips, B.T. and Kimble, J. (2009) A new look at TCF and β-catenin through the lens of a divergent *C. elegans* Wnt pathway. Dev. Cell 17, 27–34.
- Phillips, W.C., Stanton, M., Stewart, A., Qian, H., Ingersoll, C., and Sweet, R.M. (2000) Multiple CCD detector for macromolecular X-ray crystallography. J. Appl. Cryst. 33, 243–251.
- Piao, S., Lee, S.H., Kim, H., Yum, S., Stamos, J.L., Xu, Y., Lee, S.J., Lee, J., Oh, S., Han, J.K., Park, B.J., Weis, W.I., and Ha, N.C. (2008) Direct inhibition of GSK3β by the phosphorylated cytoplasmic domain of LRP6 in Wnt/β-catenin signaling. PLoS One 3, e4046.
- Piccolo, S., Agius, E., Leyns, L., Bhattacharyya, S., Grunz, H., Bouwmeester, T., and De Robertis, E.M. (1999) The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. Nature 397, 707–710.
- Pilia, G., Hughes-Benzie, R.M., MacKenzie, A., Baybayan, P., Chen, E.Y., Huber, R.,

Neri, G., Cao, A., Forabosco, A., and Schlessinger, D. (1996) Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome. Nat. Genet. 12, 241–247.

- Pinson, K.I., Brennan, J., Monkley, S., Avery, B.J., and Skarnes, W.C. (2000) An LDL-receptor-related protein mediates Wnt signalling in mice. Nature. 407, 535–538.
- Polakis P. (2007) The many ways of Wnt in cancer. Curr. Opin. Genet. Dev. 17, 45–51.
- Powell, H.R. (1999) The Rossmann Fourier autoindexing algorithm in MOSFLM. Acta Crystallogr. D Biol. Crystallogr. 55, 1690–1695.
- Qin, X., Strong, D.D., Baylink, D.J., and Mohan, S. (1998) Structure-function analysis of the human insulin-like growth factor binding protein-4. J. Biol. Chem. 273, 23509–23516.
- Rabenstein, D.L. (2002) Heparin and heparan sulfate: structure and function. Nat. Prod. Rep. 19, 312–331.
- Ramachandran, G.N., Ramakrishnan, C., and Sasisekharan, V. (1963) Stereochemistry of polypeptide chain configurations. J. Mol. Biol. 7, 95–99.
- Rao, Z., Handford, P., Mayhew, M., Knott, V., Brownlee, G.G., and Stuart D. (1995) The structure of a Ca^{2+}-binding epidermal growth factor-like domain: its role in protein-protein interactions. Cell 82, 131–141.
- Rattner, A., Hsieh, J.C., Smallwood, P.M., Gilbert, D.J., Copeland, N.G., Jenkins, N.A., and Nathans, J. (1997) A family of secreted proteins contains homology to the cysteine-rich ligand-binding domain of frizzled receptors. Proc. Natl. Acad. Sci. U. S. A. 94, 2859–2863.
- Reddien, P.W. and Sánchez Alvarado, A. (2004) Fundamentals of planarian regeneration. Annu. Rev. Cell Dev. Biol. 20, 725–757.

- Reeves, P.J., Callewaert, N., Contreras, R., and Khorana, H.G. (2002) Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. Proc. Natl. Acad. Sci. U. S. A. 99, 13419–13424.
- Reya, T., Duncan, A.W., Ailles, L., Domen, J., Scherer, D.C., Willert, K., Hintz, L., Nusse, R., and Weissman, I.L. (2003) A role for Wnt signalling in self-renewal of haematopoietic stem cells. Nature 423, 409–414.
- Riese, J., Yu, X., Munnerlyn, A., Eresh, S., Hsu, S.C., Grosschedl, R., and Bienz, M. (1997) LEF-1, a nuclear factor coordinating signaling inputs from wingless and decapentaplegic. Cell 88, 777–787.
- Rodriguez, J., Esteve, P., Weinl, C., Ruiz, J.M., Fermin, Y., Trousse, F., Dwivedy, A., Holt, C., and Bovolenta, P. (2005) SFRP1 regulates the growth of retinal ganglion cell axons through the Fz2 receptor. Nat. Neurosci. 8, 1301–1309.
- Rogers, K.W. and Schier, A.F. (2011) Morphogen Gradients: From Generation to Interpretation. Annu. Rev. Cell Dev. Biol. 2011. 27, 28.1–28.31.
- Rossmann, M.G. and Blow, D. (1962) The detection of subunits within the crystallographic asymmetric unit. Acta Crystallogr. 15, 24–31.
- Rubin, E.M., Guo, Y., Tu, K., Xie, J., Zi, X., and Hoang, B.H. (2010) Wnt inhibitory factor 1 decreases tumorigenesis and metastasis in osteosarcoma. Mol. Cancer Ther. 9, 731–741.
- Rupp, B. (2009) Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology (Garland Science, U. S. A.).
- Sakanaka, C., Weiss, J.B., and Williams, L.T. (1998) Bridging of β-catenin and glycogen synthase kinase-3β by axin and inhibition of β-catenin-mediated transcription. Proc. Natl. Acad. Sci. U. S. A. 95, 3020–3023.

- Salic, A.N., Kroll, K.L., Evans, L.M., and Kirschner, M.W. (1997) Sizzled: a secreted Xwnt8 antagonist expressed in the ventral marginal zone of Xenopus embryos. Development 124, 4739–4748.
- Sambrook, J. and Russell, D. (2001) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory Press, U. S. A.).
- Satoh, W., Gotoh, T., Tsunematsu, Y., Aizawa, S., and Shimono, A. (2006) Sfrp1 and Sfrp2 regulate anteroposterior axis elongation and somite segmentation during mouse embryogenesis. Development 133, 989–999.
- Schiller, J., Arnhold, J., Benard, S., Müller, M., Reichl, S., and Arnold, K. (1999) Lipid analysis by matrix-assisted laser desorption and ionization mass spectrometry: A methodological approach. Anal Biochem. 267, 46–56.
- Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys. J. 78, 1606–1619.
- Schwarz-Romond, T., Fiedler, M., Shibata, N., Butler, P.J., Kikuchi, A., Higuchi, Y., and Bienz, M. (2007) The DIX domain of Dishevelled confers Wnt signaling by dynamic polymerization. Nat. Struct. Mol. Biol. 14, 484–492.
- Seifert, J.R. and Mlodzik, M. (2007) Frizzled/PCP signalling: a conserved mechanism regulating cell polarity and directed motility. Nat. Rev. Genet. 8, 126–138.
- Semenov, M.V. and He, X. (2000) Secreted Antagonists/Modulators of Wnt Signaling. Madame Curie Bioscience Database, Landes Bioscience, National Center for Biotechnology Information, U.S. Bookshelf ID NBK6536.
- Shimomura, Y., Agalliu, D., Vonica, A., Luria, V., Wajid, M., Baumer, A., Belli, S., Petukhova, L., Schinzel, A., Brivanlou, A.H., Barres, B.A., and Christiano, A.M. (2010) APCDD1 is a novel Wnt inhibitor mutated in hereditary hypotrichosis simplex. Nature 464, 1043–1047.

- Stiegler, A.L., Burden, S.J., and Hubbard, S.R. (2009) Crystal structure of the frizzled-like cysteine-rich domain of the receptor tyrosine kinase MuSK. J. Mol. Biol. 393, 1–9.
- Surmann-Schmitt, C., Widmann, N., Dietz, U., Saeger, B., Eitzinger, N., Nakamura, Y., Rattel, M., Latham, R., Hartmann, C., von der Mark, H., Schett, G., von der Mark, K., and Stock, M. (2009) Wif-1 is expressed at cartilage-mesenchyme interfaces and impedes Wnt3a-mediated inhibition of chondrogenesis. J. Cell Sci. 122, 3627–3637.
- Tabata, T. and Takei, Y. (2004) Morphogens, their identification and regulation. Development 131, 703–712.
- Takada, R., Satomi, Y., Kurata, T., Ueno, N., Norioka, S., Kondoh, H., Takao, T., and Takada, S. (2006) Monounsaturated fatty acid modification of Wnt protein: its role in Wnt secretion. Dev. Cell 11, 791–801.
- Takahashi, M., Fujita, M., Furukawa, Y., Hamamoto, R., Shimokawa, T., Miwa, N., Ogawa, M., and Nakamura, Y. (2002) Isolation of a novel human gene, APCDD1, as a direct target of the β-Catenin/T-cell factor 4 complex with probable involvement in colorectal carcinogenesis. Cancer Res. 62, 5651–5656.
- Takei, Y., Ozawa, Y., Sato, M., Watanabe, A., and Tabata, T. (2004) Three *Drosophila* EXT genes shape morphogen gradients through synthesis of heparan sulfate proteoglycans. Development 131, 73–82.
- Tamai, K., Semenov, M., Kato, Y., Spokony, R., Liu, C., Katsuyama, Y., Hess, F., Saint-Jeannet, J.P., and He, X. (2000) LDL-receptor-related proteins in Wnt signal transduction. Nature 407, 530–535.
- Tanaka, E.M. and Weidinger, G. (2008) Heads or tails: can Wnt tell which one is up? Nat. Cell Biol. 10, 122–124.
- Tang, Y., Simoneau, A.R., Liao, W.X., Yi, G., Hope, C., Liu, F., Li, S., Xie, J.,

Holcombe, R.F., Jurnak, F.A., Mercola, D., Hoang, B.H., and Zi, X. (2009) WIF1, a Wnt pathway inhibitor, regulates SKP2 and c-myc expression leading to G1 arrest and growth inhibition of human invasive urinary bladder cancer cells. Mol. Cancer Ther. 8, 458–468.

- Taniguchi, H., Yamamoto, H., Hirata, T., Miyamoto, N., Oki, M., Nosho, K., Adachi, Y., Endo, T., Imai, K., and Shinomura, Y. (2005) Frequent epigenetic inactivation of Wnt inhibitory factor-1 in human gastrointestinal cancers. Oncogene 24, 7946–7952.
- Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Adams, P.D., Read, R.J., Zwart, P.H., and Hung, L.W. (2008) Iterative-build OMIT maps: map improvement by iterative model building and refinement without model bias. Acta Crystallogr. D Biol. Crystallogr. 64, 515–524.
- Trombetta, E.S. and Parodi, A.J. (2003) Quality control and protein folding in the secretory pathway. Annu. Rev. Cell Dev. Biol. 19, 649–676.
- Tsai, C.J., Lin, S.L., Wolfson, H.J., and Nussinov, R. (1997) Studies of protein-protein interfaces: a statistical analysis of the hydrophobic effect. Protein Sci. 6, 53–64.
- Tsuda, M., Kamimura, K., Nakato, H., Archer, M., Staatz, W., Fox, B., Humphrey, M., Olson, S., Futch, T., Kaluza, V., Siegfried, E., Stam, L., and Selleck, S.B. (1999) The cell-surface proteoglycan Dally regulates Wingless signalling in *Drosophila*. Nature 400, 276–280.
- UniProt Consortium. (2010) The Universal Protein Resource (UniProt) in 2010. Nucleic Acids Res. 38, D142–D148.
- Urakami, S., Shiina, H., Enokida, H., Kawakami, T., Tokizane, T., Ogishima, T., Tanaka, Y., Li, L.C., Ribeiro-Filho, L.A., Terashima, M., Kikuno, N., Adachi, H., Yoneda, T., Kishi, H., Shigeno, K., Konety, B.R., Igawa, M., and Dahiya, R. (2006) Epigenetic inactivation of Wnt inhibitory factor-1 plays an important role in bladder cancer through aberrant canonical Wnt/beta-catenin signaling pathway. Clin. Cancer Res. 12, 383–391.

- Üren, A., Reichsman, F., Anest, V., Taylor, W.G., Muraiso, K., Bottaro, D.P., Cumberledge, S., and Rubin, J.S. (2000) Secreted frizzled-related protein-1 binds directly to Wingless and is a biphasic modulator of Wnt signaling. J. Biol. Chem. 275, 4374–4382.
- van Amerongen, R. and Nusse, R. (2009) Towards an integrated view of Wnt signaling in development. Development 136, 3205–3214.
- Veverka, V., Henry, A.J., Slocombe, P.M., Ventom, A., Mulloy, B., Muskett, F.W., Muzylak, M., Greenslade, K., Moore, A., Zhang, L., Gong, J., Qian, X., Paszty, C., Taylor, R.J., Robinson, M.K., and Carr, M.D. (2009) Characterization of the structural features and interactions of sclerostin: molecular insight into a key regulator of Wnt-mediated bone formation. J. Biol. Chem. 284, 10890–10900.
- Walter, T.S., Diprose, J.M., Mayo, C.J., Siebold, C., Pickford, M.G., Carter, L., Sutton, G.C., Berrow, N.S., Brown, J., Berry, I.M., Stewart-Jones, G.B., Grimes, J.M., Stammers, D.K., Esnouf, R.M., Jones, E.Y., Owens, R.J., Stuart, D.I., and Harlos, K. (2005) A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization. Acta Crystallogr. D Biol. Crystallogr. 61, 651–657.
- Walter, T.S., Meier, C., Assenberg, R., Au, K.F., Ren, J., Verma, A., Nettleship, J.E., Owens, R.J., Stuart, D.I., and Grimes, J.M. (2006) Lysine methylation as a routine rescue strategy for protein crystallization. Structure 14, 1617–1622.
- Wang, S., Krinks, M., Lin, K., Luyten, F.P., and Moos, M. Jr. (1997) Frzb, a secreted protein expressed in the Spemann organizer, binds and inhibits Wnt-8. Cell 88, 757–766.
- Wang, Y. and Nathans, J. (2007) Tissue/planar cell polarity in vertebrates: new insights and new questions. Development 134, 647–658.
- Wang, Y., Macke, J.P., Abella, B.S., Andreasson, K., Worley, P., Gilbert, D.J., Copeland, N.G., Jenkins, N.A., and Nathans, J. (1996) A large family of putative

transmembrane receptors homologous to the product of the *Drosophila* tissue polarity gene frizzled. J. Biol. Chem. 271, 4468–4476.

- Wawrzak, D., Métioui, M., Willems, E., Hendrickx, M., de Genst, E., and Leyns, L. (2007) Wnt3a binds to several sFRPs in the nanomolar range. Biochem. Biophys. Res. Commun. 357, 1119–1123.

- Wehrli, M., Dougan, S.T., Caldwell, K., O'Keefe, L., Schwartz, S., Vaizel-Ohayon, D., Schejter, E., Tomlinson, A., and DiNardo, S. (2000) *arrow* encodes an LDL-receptor-related protein essential for Wingless signalling. Nature. 407, 527–530.

- Weidauer, S.E., Schmieder, P., Beerbaum, M., Schmitz, W., Oschkinat, H., and Mueller, T.D. (2009) NMR structure of the Wnt modulator protein Sclerostin. Biochem. Biophys. Res. Commun. 380, 160–165.

- Willert, K., Brown, J.D., Danenberg, E., Duncan, A.W., Weissman, I.L., Reya, T., Yates, J.R. 3rd, and Nusse, R. (2003) Wnt proteins are lipid-modified and can act as stem cell growth factors. Nature 423, 448–452.

- Wong, H.C., Bourdelas, A., Krauss, A., Lee, H.J., Shao, Y., Wu, D., Mlodzik, M., Shi, D.L., and Zheng, J. (2003) Direct binding of the PDZ domain of Dishevelled to a conserved internal sequence in the C-terminal region of Frizzled. Mol. Cell 12, 1251–1260.

- Wu, G., Huang, H., Garcia Abreu, J., and He, X. (2009) Inhibition of GSK3 phosphorylation of β-catenin via phosphorylated PPPSPXS motifs of Wnt coreceptor LRP6. PLoS One 4, e4926.

- Wu, M., Le, H.D., Wang, M., Yurkov, V., Omelchenko, A., Hnatowich, M., Nix, J., Hryshko, L.V., and Zheng, L. (2010) Crystal structures of progressive Ca^{2+} binding states of the Ca^{2+} sensor Ca^{2+} binding domain 1 (CBD1) from the CALX Na^{+}/Ca^{2+} exchanger reveal incremental conformational transitions. J. Biol. Chem. 285, 2554–2561.

- Yabe, T., Shimizu, T., Muraoka, O., Bae, Y.K., Hirata, T., Nojima, H., Kawakami, A.,

Hirano, T., and Hibi, M. (2003) Ogon/Secreted Frizzled functions as a negative feedback regulator of Bmp signaling. Development 130, 2705–2716.

- Yan, D. and Lin, X. (2009) Shaping morphogen gradients by proteoglycans. Cold Spring Harb. Perspect. Biol. 1, a002493.

- Yan, D., Wu, Y., Feng, Y., Lin, S.C., and Lin, X. (2009) The core protein of glypican Dally-like determines its biphasic activity in wingless morphogen signaling. Dev. Cell. 17, 470–481.

- Yang, Z., Wang, Y., Fang, J., Chen, F., Liu, J., Wu, J., and Wang, Y. (2010) Expression and aberrant promoter methylation of Wnt inhibitory factor-1 in human astrocytomas. J. Exp. Clin. Cancer Res. 29, 26.

- Yang, Z.R., Thomson, R., McNeil, P., and Esnouf, R.M. (2005) RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. Bioinformatics 21, 3369–3376.

- Yee, D.S., Tang, Y., Li, X., Liu, Z., Guo, Y., Ghaffar, S., McQueen, P., Atreya, D., Xie, J., Simoneau, A.R., Hoang, B.H., and Zi, X. (2010) The Wnt inhibitory factor 1 restoration in prostate cancer cells was associated with reduced tumor growth, decreased capacity of cell migration and invasion and a reversal of epithelial to mesenchymal transition. Mol. Cancer 9, 162.

- Yoshikawa, S., McKinnon, R.D., Kokel, M., and Thomas, J.B. (2003) Wnt-mediated axon guidance via the *Drosophila* Derailed receptor. Nature 422, 583–588.

- Zeng, L., Fagotto, F., Zhang, T., Hsu, W., Vasicek, T.J., Perry, W.L. 3rd, Lee, J.J., Tilghman, S.M., Gumbiner, B.M., and Costantini, F. (1997) The mouse Fused locus encodes Axin, an inhibitor of the Wnt signaling pathway that regulates embryonic axis formation. Cell 90, 181–192.

- Zeng, Y.A. and Nusse, R. (2010) Wnt proteins are self-renewal factors for mammary

stem cells and promote their long-term expansion in culture. Cell Stem Cell 6, 568–577.

- Zhai, L., Chaturvedi, D., and Cumberledge, S. (2004) *Drosophila* wnt-1 undergoes a hydrophobic modification and is targeted to lipid rafts, a process that requires porcupine. J. Biol. Chem. 279, 33220–33227.
- Zhang, B., Abreu, J.G., Zhou, K., Chen, Y., Hu, Y., Zhou, T., He, X., and Ma, J.X. (2010) Blocking the Wnt pathway, a unifying mechanism for an angiogenic inhibitor in the serine proteinase inhibitor family. Proc. Natl. Acad. Sci. U. S. A. 107, 6900–6905.
- Zhang, F., McLellan, J.S., Ayala, A.M., Leahy, D.J., and Linhardt, R.J. (2007) Kinetic and structural studies on interactions between heparin or heparan sulfate and proteins of the hedgehog signaling pathway. Biochemistry 46, 3933–3941.
- Zhang, K.Y., Cowtan, K., and Main, P. (1997) Combining constraints for electron-density modification. Methods Enzymol. 277, 53–64.
- Zhu, W., Shiojima, I., Ito, Y., Li, Z., Ikeda, H., Yoshida, M., Naito, A.T., Nishi, J., Ueno, H., Umezawa, A., Minamino, T., Nagai, T., Kikuchi, A., Asashima, M., and Komuro, I. (2008) IGFBP-4 is an inhibitor of canonical Wnt signalling required for cardiogenesis. Nature 454, 345–349.
- Zoltewicz, J.S., Ashique, A.M., Choe, Y., Lee, G., Taylor, S., Phamluong, K., Solloway, M., and Peterson, A.S. (2009) Wnt signaling is regulated by endoplasmic reticulum retention. PLoS One 4, e6191.