LETTER

https://doi.org/10.1038/s41586-018-0447-x

Crystal structure of the Frizzled 4 receptor in a ligand-free state

Shifan Yang[1], Yiran Wu[1,11], Ting-Hai Xu[2,11], Parker W. de Waal[2,11], Yuanzheng He[3], Mengchen Pu[1], Yuxiang Chen[1,4,5,6], Zachary J. DeBruine[2], Bingjie Zhang[1], Saheem A. Zaidi[7], Petr Popov[7,8], Yu Guo[1,4,5,6], Gye Won Han[7], Yang Lu[3], Kelly Suino-Powell[2], Shaowei Dong[1,4,5,6], Kaleeckal G. Harikumar[9], Laurence J. Miller[9], Vsevolod Katritch[7,8], H. Eric Xu[2,10], Wenqing Shui[1,4], Raymond C. Stevens[1,4], Karsten Melcher[2], Suwen Zhao[1,4] & Fei Xu[1,4]*

Frizzled receptors (FZDs) are class-F G-protein-coupled receptors (GPCRs) that function in Wnt signalling and are essential for developing and adult organisms[1,2]. As central mediators in this complex signalling pathway, FZDs serve as gatekeeping proteins both for drug intervention and for the development of probes in basic and in therapeutic research. Here we present an atomic-resolution structure of the human Frizzled 4 receptor (FZD4) transmembrane domain in the absence of a bound ligand. The structure reveals an unusual transmembrane architecture in which helix VI is short and tightly packed, and is distinct from all other GPCR structures reported so far. Within this unique transmembrane fold is an extremely narrow and highly hydrophilic pocket that is not amenable to the binding of traditional GPCR ligands. We show that such a pocket is conserved across all FZDs, which may explain the long-standing difficulties in the development of ligands for these receptors. Molecular dynamics simulations on the microsecond timescale and mutational analysis uncovered two coupled, dynamic kinks located at helix VII that are involved in FZD4 activation. The stability of the structure in its ligand-free form, an unfavourable pocket for ligand binding and the two unusual kinks on helix VII suggest that FZDs may have evolved a novel ligand-recognition and activation mechanism that is distinct from that of other GPCRs.

FZDs have been grouped as class-F GPCRs on the basis of their sequence homology with other GPCR families. Although the main intracellular FZD-binding protein and Wnt signal-transducer is Dishevelled, several FZD receptors also have the ability to couple to G proteins and induce G-protein nucleotide exchange, which defines them as GPCRs[3]. However, the mechanistic details of FZD signalling through G proteins are poorly understood, and it is not clear whether all FZDs can couple to G proteins. FZDs are involved in the regulation of many biological processes during embryonic development and tissue homeostasis[4]. As the major cell-surface receptor for the Wnt pathway, ten FZDs have central roles in Wnt signalling, and their aberrance is linked to numerous diseases[5–7]. Consequently, the structural and functional characterization of FZD and the exploration of FZD-targeted therapeutics have received considerable attention in recent years. The elucidation of the structures of an extracellular soluble region in FZDs—the cysteine-rich domain (CRD)[8,9]—and its complex with Wnt protein[10–12] led to the discovery of Wnt surrogates, or antibody fragments, that target this region[13,14]. However, ligands that target the transmembrane domain (TMD) of FZDs—the traditional pocket in which many GPCR-targeting drugs act—are scarce, and the dynamics and signalling of these receptors are still largely unknown. To provide guidance on the development of FZD ligands and to understand the activation and signalling of FZDs, we solved the TMD structure of the human FZD4 receptor (hereafter referred to as ΔCRD-FZD4)—which shares 37–61% sequence homology with the other nine FZDs (Extended Data Fig. 1)—at a resolution of 2.4 Å in a ligand-free state (Extended Data Table 1a).

The construct of ΔCRD-FZD4 contains residues 178–517, with eight residues (420–427) of the third intracellular loop (ICL3) replaced by rubredoxin to facilitate crystallization (see Methods and Extended Data Fig. 2). The overall structure of ΔCRD-FZD4 reveals a canonical GPCR fold with seven transmembrane helices (7TM; helices I–VII) and a short helix 8 (H8) containing the conserved signalling-related KTXXXW motif, which is essential for the recruitment of the intracellular signal transducer Dishevelled[15], packed parallel to the membrane bilayer (Fig. 1a). We crystallized ΔCRD-FZD4 in its ligand-free state, and to our knowledge this is the first apo structure of a ligand-regulated GPCR (Fig. 1b).

Compared to the smoothened receptor (SMO), which belongs to the class-F receptors and is an important target for antitumour therapy[16–18], the 7TM bundle fold of FZD4 is similar, with a C_α root mean square deviation (r.m.s.d.) of 1.2 Å. Despite the high structural conservation of the TMDs, the helix-VI extension in FZD4 is much shorter than that in SMO. Additionally, the intracellular end of helix V moves outwards from the helical bundle by about 13° (Fig. 1c). To assess the stability of helices V and VI and their role in receptor activation, we ran three independent three-microsecond molecular dynamics simulations based on our structure (Extended Data Fig. 3a). We observed that helix VI of FZD4 remained closely packed within the TMD helical bundle of the receptor throughout simulation. This locked conformation of helix VI is primarily mediated by substantial hydrophobic interactions and two unique backbone hydrogen bonds (W320$^{3.43f}$ and W327$^{3.50f}$; superscripts denote Ballesteros–Weinstein numbering for GPCRs) that connect helix III to helix VI (Extended Data Fig. 3b). Unlike most of the other class-A GPCRs and SMO[19], in which helix VI moves outwards upon activation, helix VI in FZD4 remained relatively stable in an inward orientation (Extended Data Fig. 3b).

Among the ten FZDs, FZD4 is the only receptor that can be activated by binding of the protein Norrin. It is involved in retinal angiogenesis and in maintaining the integrity of the blood–retinal barrier, and mutations of FZD4 are found in familial exudative vitreoretinopathy[20,21]. To gain insight into this structure–function relationship, we mutated some disease- and signalling-related 'hotspot' residues[22] and assessed their function using the TOPflash reporter assay. We found that these mutations were in key positions of our ΔCRD-FZD4 structure and led to aberrant downstream signalling (Extended Data Fig. 4a). We also observed that family-conserved amino acids (Y250$^{2.39f}$ and W494$^{7.55f}$)

[1]iHuman Institute, ShanghaiTech University, Shanghai, China. [2]Center for Cancer and Cell Biology, Innovation and Integration Program, Van Andel Research Institute, Grand Rapids, MI, USA. [3]HIT Center for Life Sciences, School of Life Science and Technology, Harbin Institute of Technology, Harbin, China. [4]School of Life Science and Technology, ShanghaiTech University, Shanghai, China. [5]Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China. [6]University of Chinese Academy of Sciences, Beijing, China. [7]Departments of Biological Sciences and Chemistry, Bridge Institute, University of Southern California, Los Angeles, CA, USA. [8]Moscow Institute of Physics and Technology, Dolgoprudny, Russia. [9]Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Scottsdale, AZ, USA. [10]Key Laboratory of Receptor Research, VARI-SIMM Center, Center for Structure and Function of Drug Targets, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China. [11]These authors contributed equally: Yiran Wu, Ting-Hai Xu, Parker W. de Waal. *e-mail: xufei@shanghaitech.edu.cn

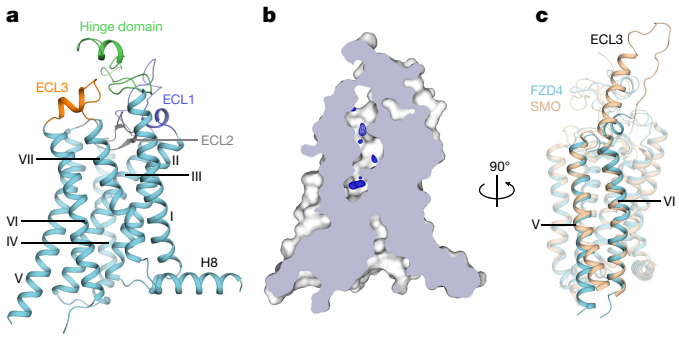


Fig. 1 | Overall structure of ΔCRD-FZD4 and the comparison with SMO. a, Side view of ΔCRD-FZD4 with the hinge domain, ECL1, ECL2 and ECL3 domains coloured in green, blue, grey, and orange, respectively. **b**, The electron density in the pocket represented as an $|F_o| - |F_c|$ maps contoured at 3.0σ in dark blue, which indicates that our structure is in a truly apo state (residual densities represent water molecules and an unknown atom or ion density). **c**, Comparison of ΔCRD-FZD4 (cyan) with ΔCRD-SMO (gold, truncated from RCSB Protein Data Bank code (PDB): 5L7D).

adopted different conformations in ΔCRD-FZD4 when compared with SMO (Extended Data Fig. 4b). These residues have been reported to have important roles in downstream signalling and in maintaining structural integrity, which indicates that these key residues have evolved distinct conformations in different proteins to correspond to their respective functions[23]. Our mutational analysis confirmed that the mutations Y250F and W494L led to reduced signalling activity (Extended Data Fig. 4a).

Despite the lack of the CRD, the arrangement of the extracellular side (hinge domain and extracellular loops (ECLs)) of this ΔCRD-FZD4 structure reveals some interesting features of the domain interaction and sheds light on the connection to the CRD. Therefore, we compared the domain organization of FZD4 with that of other multi-domain GPCRs. For the two class-F receptors, FZD4 and SMO, individual fragments of the extracellular side (hinge domain, ECL1, ECL2 and ECL3) pack together to form a compact structure (Fig. 2a, b). This unique tertiary structure in turn stabilizes the 7TM bundle with a continuous interface of 2,625 Å^2. In FZD4, the ECL2 β-hairpin acts as a plug, embedded deeply in the receptor, and occupies an abundance of space in the 7TM bundle cavity. ECL1 and ECL3 sandwich and stabilize the hinge domain through polar and nonpolar interactions (Fig. 2a). By contrast, although the ECL2 β-hairpin is located at a similar position in SMO, the hinge domain stacks on top of ECL1 and leans against ECL3, adopting an elongated conformation (Fig. 2b). This difference in organization of the hinge domain between FZD4 and SMO may result in a different arrangement of or different dynamics towards their respective CRDs. To illustrate this, we built a tentative full-length model of FZD4 (Extended Data Fig. 5a), which shows that the much shorter ECL3 of FZD4 markedly reduces the interface area between the CRD and TMD surfaces compared to SMO. Such a diminished interface results in a less stable orientation of the CRD, which we have confirmed by molecular dynamics simulations: substantial swinging of the domain is observed on a 1.5-microsecond scale (Extended Data Fig. 5b).

Distinct from class-F receptors, the extracellular regions (hinge domain and extracellular loops) in class-B and class-C receptors are less compact. In the glucagon receptor (GCGR)[24], the glucagon-like peptide 1 receptor (GLP-1R)[25] and the metabotropic glutamate-1 receptor (mGlu-1R)[26], the hinge domain interacts tightly with ECL1, the partially disordered ECL2 moves away from the binding cavity, and ECL3 is in an isolated position with no contact with other members from the extracellular side (Fig. 2c–e). Not surprisingly, such loose stacking leaves a spacious binding pocket for cognate ligands and supports an 'open–closed' conformational change of the extracellular domain in the activation of class-B GPCRs.

We next compared the arrangement of the intracellular side of FZD4 with that of SMO and representative class-A receptors (Extended

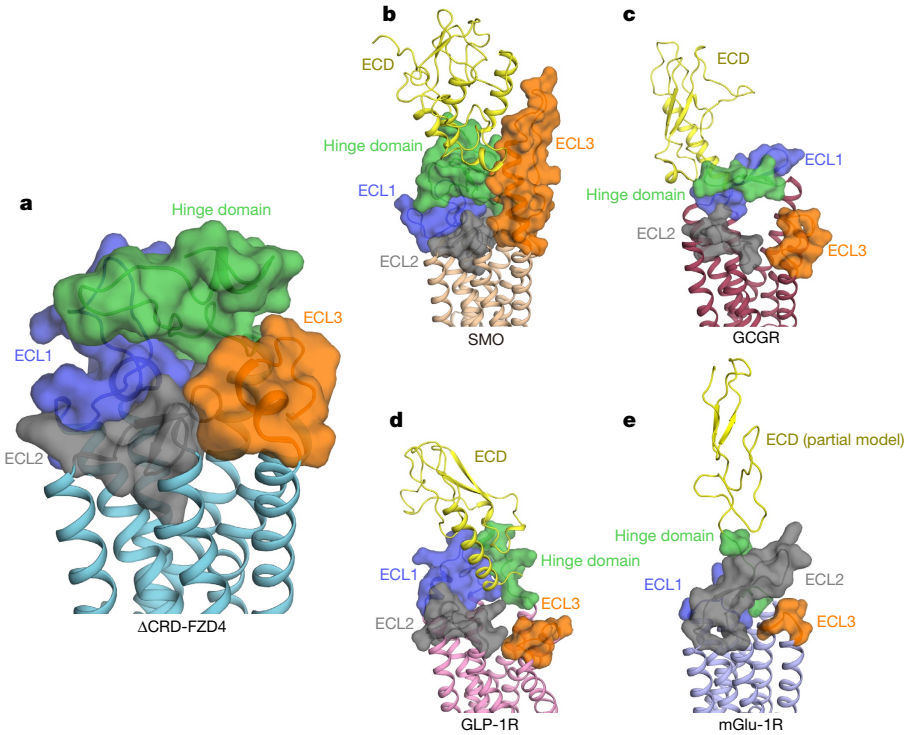


Fig. 2 | The arrangement of the extracellular side of FZD4 and other GPCRs. a, **b**, In FZD4 (**a**) and SMO (**b**; PDB: 5L7D), ECL2 acts as a plug and the individual fragments of the extracellular region (hinge domain, ECLs) form a compact structure. **c**–**e**, In GCGR (PDB: 5XF1), GLP-1R (PDB: 5NX2) and mGlu-1R (PDB: 4OR2), the organization of the extracellular region (hinge domain and ECLs) is less compact so either an antibody or a ligand is required for stabilization of the 7TM and/or the extracellular domain (ECD). ECDs for SMO, GCGR and GLP-1R are derived from full-length structures, and for mGlu-1R a partial model (PDB: 2E4W) is presented.

Data Fig. 3c). It was noted that FZD4 is flatter on the intracellular side compared to SMO and some class-A GPCRs, and that no cavity exists among helices III,VI or VII; such a cavity has been identified as an allosteric ligand-binding site in some class-A GPCRs[27]. The family-conserved Arg$^{3.50}$ in class-A GPCRs, which has previously been reported to be a G-protein-binding site[28], is also missing in FZD4 and SMO. The various different arrangements in FZD4 suggest that this receptor has a different mechanism for the recognition of allosteric ligands and downstream signalling molecules compared with other GPCRs.

Despite being a central intervention point of the Wnt–FZDs pathway, the discovery of a ligand that directly targets the traditional GPCR transmembrane pocket of FZDs has been challenging. Nonetheless, several drugs that target the SMO transmembrane pocket have been approved by the US Food and Drug Administration for the treatment of cancer. Compared to that of SMO, the transmembrane pocket of FZD4 is more constricted and cannot accommodate SMO ligands. When we superimposed SMO ligands, extensive clashes were found between the ligands and the FZD4 pocket (Fig. 3a and Extended Data Fig. 6a); this result was further confirmed by affinity mass spectrometry to assess the receptor–ligand interaction (Extended Data Table 1b). To understand these extensive clashes, we calculated the volume of the transmembrane pocket in FZD4, SMO and in all other GPCRs for which structures have been solved. The volume of the FZD4 pocket (882 Å^3) is larger than that of the SMO pocket (763 Å^3), and is in the middle of the range for all GPCRs with solved structures. However, there are two extremely narrow points—caused by the presence of bulky side chains—in which the area of the channel of the FZD4 pocket is reduced to 4.8 Å^2 and 8.0 Å^2 (Fig. 3b, c). Further molecular dynamics simulation analysis revealed that the side chains pointing to the pocket could only fluctuate over a small range, and that the cross-sectional area of the pocket does not change considerably (Fig. 3d and Extended Data Fig. 6b). When we mutated these residues and tested the effects by cell-based TOPflash reporter assay, we found that none of the mutations markedly reduced FZD4 activation either in the absence or in the presence of endogenous agonist (Fig. 3e and Extended Data Fig. 6c), which is consistent with our hypothesis that agonist binding to the pocket is not required for canonical signalling.

Besides being markedly different from its analogues in terms of shape, the pocket of FZD4 also contains a cluster of polar residues pointing to the cavity, which makes it the most hydrophilic pocket of all GPCRs with known structure (Fig. 3d and Extended Data Fig. 6d). Conservation analysis of homology models for the other nine FZDs based on our structure (Extended Data Fig. 7) indicated that all FZDs share high structural similarity and sequence homology at transmembrane helices, whereas the extracellular and intercellular regions are poorly conserved (Extended Data Fig. 7a). The aforementioned pocket-shape-related and polar residues form a conserved hydrophilic pocket in all FZDs that may substantially reduce the binding affinity of a ligand (Extended Data Fig. 7b–d). This may partially explain why it has been a challenge to develop good small-molecule ligands for FZDs in recent years. Other challenges arise from the redundancy of FZD in the genome and the high degree of homology among the ten FZDs, which might prohibit the identification of ligands as specific small-molecule inhibitors for each individual FZD.

To understand the activation and signalling of FZD4, we carried out a series of molecular dynamics simulations on a three-microsecond timescale, in addition to mutagenesis studies. Molecular dynamics results indicate a concerted movement at the proposed Dishevelled-binding site[23], where the receptor cycled between a 'closed' and a 'bent' conformation (Fig. 4a and Extended Data Fig. 8a). When cycling between these two states, the hydrogen bond between W352$^{4.50f}$ and H348$^{4.46f}$ breaks, which enables H348$^{4.46f}$ to swing by around 5 Å. Helix IV adopts a bent conformation enabling Y250$^{2.39f}$, which is sandwiched by R253$^{2.42f}$ and E341$^{4.39f}$ in the crystal structure (closed conformation), to swing by approximately 5 Å towards the newly exposed backbone carbonyl of S344$^{4.42f}$ (Fig. 4a). This opens a new

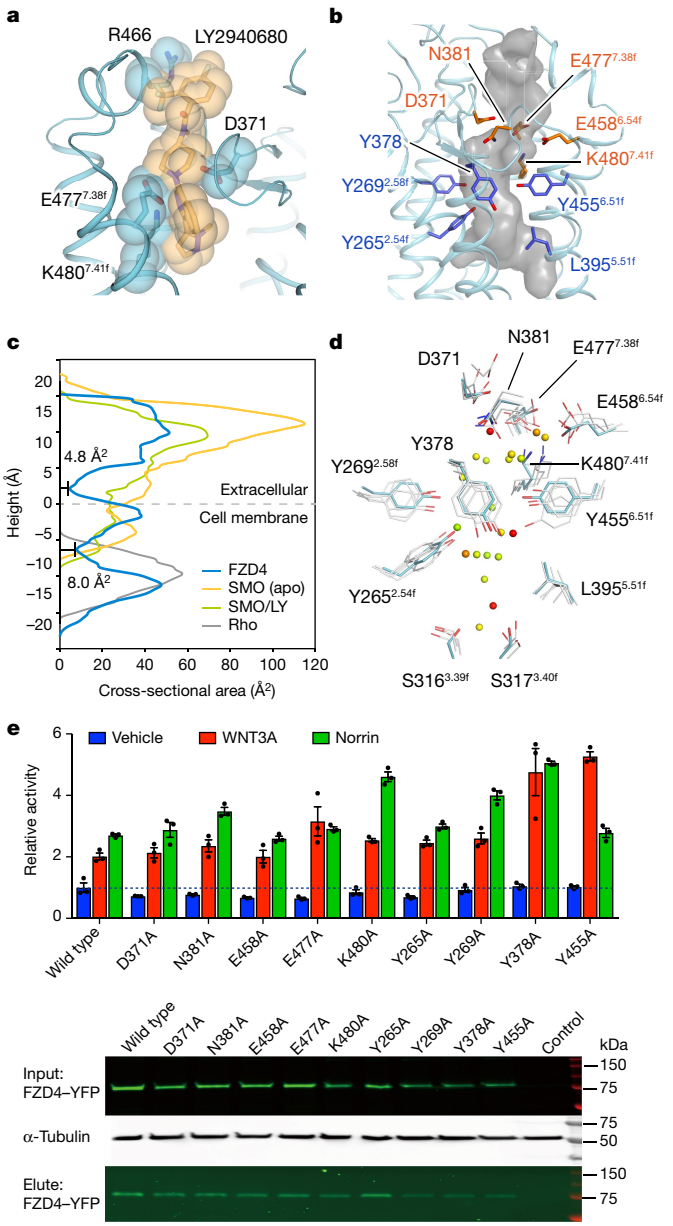


Fig. 3 | Transmembrane ligand pocket of FZD4. a, Superimposition of the SMO ligand LY2940680 (gold) in the pocket of FZD4 (cyan). Note the extensive clashes between LY2940680 and FZD4. **b**, Two narrow positions in the pocket of FZD4. The amino acids clamping position I (upper position) are shown in orange and those clamping position II (lower position) are shown in blue. **c**, The cross-sectional area of the ligand pockets of FZD4 (blue), the apo form of SMO (orange), SMO bound to LY2940680 (SMO/LY; green) and rhodopsin (Rho; grey). Note that there are two narrow positions in the pocket of FZD4. **d**, Molecular dynamics simulations of water molecules and key pocket-residue rotamers in the FZD4 pocket. Pocket-residue side chains in the crystal structure and in molecular dynamics simulations are coloured in cyan and grey, respectively. Simulated water molecules are represented as small spheres and coloured according to frequency: darker colours indicate higher frequencies. **e**, Mutational analysis of residues pointing towards the pocket indicated that agonist binding to the pocket is not required for signalling. The TOPflash reporter gene assay (top) was normalized to the activity of unstimulated wild-type FZD4 and each data point represents the mean ± s.e.m., repeated in triplicate. The cell-surface expression (bottom) for each mutant was determined by isolation of cell-surface-biotinylated proteins on avidin resin (see Methods). Cell-surface biotinylation was cross-validated by fluorescence microscopy analysis with similar results. Corresponding cell localization and total expression data are shown in Extended Data Fig. 6c.

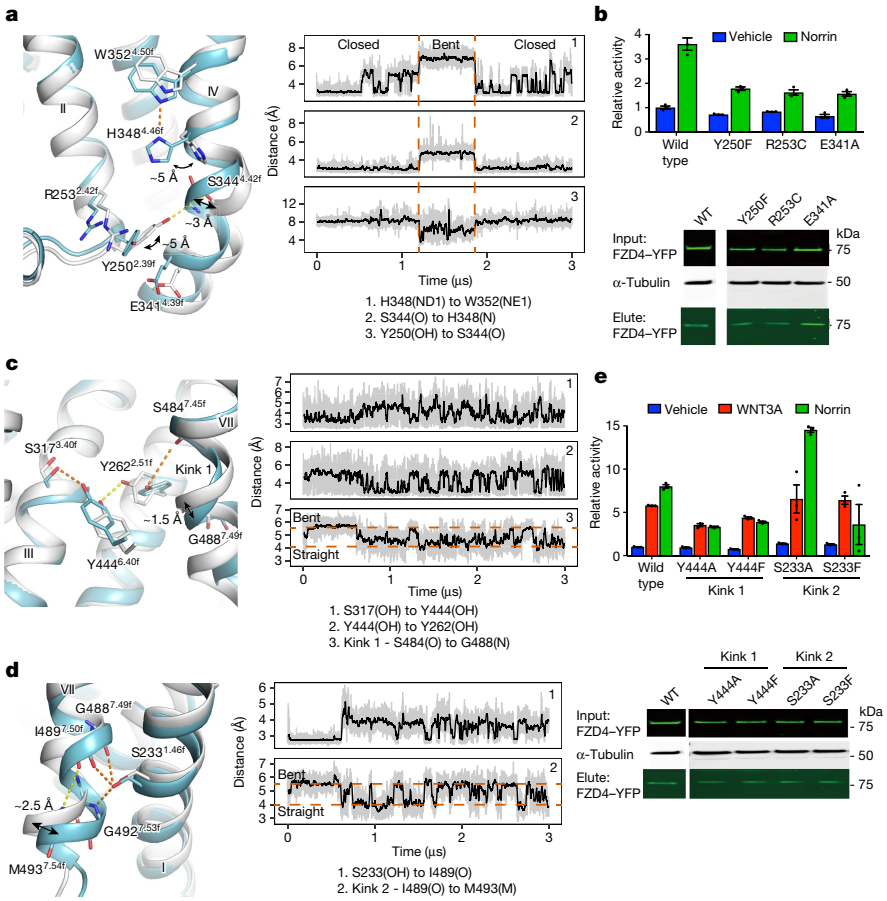


Fig. 4 | Activation (kinks) and Dishevelled-mediated signalling analysis of FZD4. a, Snapshots and distance traces of closed (cyan) and bent (grey) conformations observed during simulation at the Dishevelled-binding site. **b**, Activity assay (TOPflash reporter assay) and cell-surface expression data for mutations of the Y250 network residues. Each data point represents the mean ± s.e.m., repeated in triplicate. Cell-surface expression data are shown underneath the bar chart. Cell-surface biotinylation was cross-validated by fluorescence microscopy analysis with similar results. **c**, **d**, Snapshots and distance traces of bent (cyan) and straight (white) conformations for kink 1 (**c**) and kink 2 (**d**) in helix VII. **e**, Activity assay (TOPflash reporter assay, top) and cell-surface expression data (bottom) for mutations in kink1 and kink2. Each data point represents the mean ± s.e.m., repeated in triplicate. Cell-surface biotinylation was cross-validated by fluorescence microscopy analysis with similar results. Corresponding cell localization and total expression data are shown in Extended Data Figs. 8 and 9. Distance traces from molecular dynamics data are plotted as a moving average over a 10-ns window.

polar cavity that is flooded with additional water molecules (Extended Data Fig. 8a). Furthermore, mutations of any of the residues in the $R253^{2.42f}$–$Y250^{2.39f}$–$E341^{4.39f}$ sandwich caused Norrin-induced activation defects (Fig. 4b and Extended Data Figs. 4b, 8b). Our results suggest that this region can undergo a large conformational change in which the cytoplasmic portion of helix IV swings away from the receptor to form a pocket that is suitable for Dishevelled binding.

Besides the concerted movement of the site that is crucial for the binding of Dishevelled, two unusual and closely coupled, dynamic kinks (kink 1 and kink 2) can be seen in helix VII (Fig. 4c, d and Extended Data Fig. 9a). These two kinks are involved in conserved polar networks. At kink 1, the hydrogen-bonding network flips between two states: from $S317^{3.40f}$–$Y444^{6.40f}$ and $Y262^{2.51f}$–$S484^{7.45f}$ to $Y444^{6.40f}$–$Y262^{2.51f}$ (Fig. 4c and Extended Data Fig. 9b). Mutation of $Y444^{6.40f}$ in the centre of kink 1 reduced the activation of FZD4 (Fig. 4e and Extended Data Fig. 9c). Compared with kink 1, $S233^{1.46f}$ in kink 2 is involved in a tight hydrogen-bonding network in which it keeps the lower portion of helix VII in a bent conformation (Fig. 4d and Extended Data Fig. 9b). The loss of the polar side chain of $S233^{1.46f}$ would act to favour the linear conformation of kink 2 over the bent conformation. It is notable that S233A greatly increases Norrin-induced signalling with no effect on WNT3A signalling (Fig. 4e and Extended Data Fig. 9c). One possible explanation is that the dynamics seen at kink 2, and by extension at H8, are specific to Norrin signalling. These uniquely coupled kinks in the structure of ΔCRD-FZD4 undergo considerable movement in helix VII and H8 during simulations (Extended Data Fig. 9a), whereas helix VI remains relatively stable. The 'bent-to-linear' transformation in helix VII and H8 may relate to receptor activation, which is in contrast to the outward movement of helix VI that is seen in the receptors of other family members. In addition, FZD activation requires co-receptors such as LRP5 and LRP6, and may include the rearrangement of homo- or heterodimers. Previous studies have reported that an artificial ligand that induces the formation of a heterodimer between FZD and the co-receptor LRP6 is sufficient for activating the pathway[13]. The observation that dimerization of FZD and LRP6 (by several engineered ligands) is sufficient for signalling stands against conventional 'helix VI outward movement' of FZDs in the activation of the Wnt pathway and is consistent with the structural insight presented here.

The structure of a Frizzled receptor transmembrane domain and the activation mechanism we present here pave the way to revealing the function of FZDs as essential gatekeepers in the Wnt signalling pathway. As emerging cancer targets, the role of FZDs remains elusive. The more structural insight that is available for each FZD, the greater the likelihood of understanding their signalling pathways and developing successful drugs.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-018-0447-x.

Received: 12 October 2017; Accepted: 10 July 2018;
Published online 22 August 2018.

1. Nusse, R. & Clevers, H. Wnt/β-catenin signaling, disease, and emerging therapeutic modalities. *Cell* **169**, 985–999 (2017).
2. Clevers, H. & Nusse, R. Wnt/β-catenin signaling and disease. *Cell* **149**, 1192–1205 (2012).
3. Dijksterhuis, J. P., Petersen, J. & Schulte, G. WNT/Frizzled signalling: receptor-ligand selectivity with focus on FZD-G protein signalling and its physiological relevance: IUPHAR Review 3. *Br. J. Pharmacol*. **171**, 1195–1209 (2014).
4. Schulte, G. International union of basic and clinical pharmacology. LXXX. The class Frizzled receptors. *Pharmacol. Rev*. **62**, 632–667 (2010).
5. Wang, Y., Chang, H., Rattner, A. & Nathans, J. Frizzled receptors in development and disease. *Curr. Top. Dev. Biol*. **117**, 113–139 (2016).
6. Kahn, M. Can we safely target the WNT pathway? *Nat. Rev. Drug Discov*. **13**, 513–532 (2014).
7. Tao, L. et al. Frizzled proteins are colonic epithelial receptors for *C. difficile* toxin B. *Nature* **538**, 350–355 (2016).
8. Nile, A. H., Mukund, S., Stanger, K., Wang, W. & Hannoush, R. N. Unsaturated fatty acyl recognition by Frizzled receptors mediates dimerization upon Wnt ligand binding. *Proc. Natl Acad. Sci. USA* **114**, 4147–4152 (2017).
9. DeBruine, Z. J. et al. Wnt5a promotes Frizzled-4 signalosome assembly by stabilizing cysteine-rich domain dimerization. *Genes Dev*. **31**, 916–926 (2017).
10. Janda, C. Y., Waghray, D., Levin, A. M., Thomas, C. & Garcia, K. C. Structural basis of Wnt recognition by Frizzled. *Science* **337**, 59–64 (2012).
11. Chang, T. H. et al. Structure and functional properties of Norrin mimic Wnt for signalling with Frizzled4, Lrp5/6, and proteoglycan. *eLife* **4**, (2015).
12. Shen, G. et al. Structural basis of the Norrin–Frizzled 4 interaction. *Cell Res*. **25**, 1078–1081 (2015).
13. Janda, C. Y. et al. Surrogate Wnt agonists that phenocopy canonical Wnt and β-catenin signalling. *Nature* **545**, 234–237 (2017).
14. Gurney, A. et al. Wnt pathway inhibition via the targeting of Frizzled receptors results in decreased growth and tumorigenicity of human tumors. *Proc. Natl Acad. Sci. USA* **109**, 11717–11722 (2012).
15. Umbhauer, M. et al. The C-terminal cytoplasmic Lys-Thr-X-X-X-Trp motif in Frizzled receptors mediates Wnt/β-catenin signalling. *EMBO J*. **19**, 4944–4954 (2000).
16. Byrne, E. F. X. et al. Structural basis of Smoothened regulation by its extracellular domains. *Nature* **535**, 517–522 (2016).
17. Wang, C. et al. Structure of the human smoothened receptor bound to an antitumour agent. *Nature* **497**, 338–343 (2013).
18. Zhang, X. et al. Crystal structure of a multi-domain human smoothened receptor in complex with a super stabilizing ligand. *Nat. Commun*. **8**, 15383 (2017).
19. Huang, P. et al. Structural basis of Smoothened activation in Hedgehog signaling. *Cell* **174**, 312–324.e16 (2018).
20. Wang, Y. et al. Norrin/Frizzled4 signaling in retinal vascular development and blood brain barrier plasticity. *Cell* **151**, 1332–1344 (2012).
21. Zhang, C. et al. Norrin-induced Frizzled4 endocytosis and endo-lysosomal trafficking control retinal angiogenesis and barrier function. *Nat. Commun*. **8**, 16050 (2017).
22. Nikopoulos, K. et al. Overview of the mutation spectrum in familial exudative vitreoretinopathy and Norrie disease with identification of 21 novel variants in *FZD4*, *LRP5*, and *NDP*. *Hum. Mutat*. **31**, 656–666 (2010).
23. Strakova, K. et al. The tyrosine $Y250^{2.39}$ in Frizzled 4 defines a conserved motif important for structural integrity of the receptor and recruitment of Disheveled. *Cell. Signal*. **38**, 85–96 (2017).
24. Zhang, H. et al. Structure of the full-length glucagon class B G-protein-coupled receptor. *Nature* **546**, 259–264 (2017).
25. Jazayeri, A. et al. Crystal structure of the GLP-1 receptor bound to a peptide agonist. *Nature* **546**, 254–258 (2017).
26. Wu, H. et al. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science* **344**, 58–64 (2014).
27. Zheng, Y. et al. Structure of CC chemokine receptor 2 with orthosteric and allosteric antagonists. *Nature* **540**, 458–461 (2016).
28. Rasmussen, S. G. et al. Crystal structure of the β_2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549–555 (2011).

Acknowledgements This work was supported by the National Natural Science Foundation (NSF) of China grant 31670736 (F.X.), the National Key Research and Development Program of China grant 2018YFA0507004 (F.X.) and 2016YCF0905902 (S.Z.), the NSF of Shanghai grant 16ZR1448500 (S.Z.), the Russian Foundation for Basic Research (RFBR 18-34-00990) (P.P.) and Shanghai Municipal Government, ShanghaiTech University. The diffraction data were collected at BL41XU@Spring-8 with JASRI proposals 2016B2702. We thank J. Liu, X. Gu, N. Chen and L. Xue of the BV facility at the iHuman Institute, ShanghaiTech University for protein expression support; M. Hanson from the GPCR Consortium and K. Diederichs from the University of Konstanz for X-ray data processing; the mass spectrometry facility at the National Protein Science Center (Shanghai, China) for technical assistance; and A. Pautsch from Boehringer Ingelheim and W. Zhong from Amgen for discussions.

Reviewer information *Nature* thanks M. Filizola, X. He, A. K. Shukla and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions S.Y. performed cloning, protein purification, crystallization, data collection, structural analysis and figure preparation; Y.W. carried out structure analysis, molecular dynamics simulations and figure preparation; T.-H.X. performed mutagenesis, cellular localization, cell surface biotinylation, and TOPflash reporter assays and corresponding figure preparation; P.W.d.W. carried out molecular dynamics simulations of ΔCRD-FZD4, structure analysis and corresponding figure preparation; Y.H., Z.J.D., Y.L., K.S.-P. and K.G.H. performed mutagenesis and TOPflash reporter assays; M.P. carried out molecular replacement and structure refinement; B.Z. carried out affinity mass spectrometry; S.A.Z. performed full-length modelling; P.P. was responsible for stabilizing mutation design; G.W.H. was responsible for structure refinement, quality control and deposition; Y.C. and S.D. performed cloning and protein purification; Y.G. carried out computational analysis; V.K. supervised the stabilization of mutation design and full-length modelling; W.S. supervised the affinity mass spectrometry analysis and table preparation; L.J.M., K.M. and H.E.X. designed cell-based experiments, data analysis and interpretations; R.C.S. supervised the structure analysis; S.Z. supervised the structure analysis, simulation and figure preparation; F.X. designed and supervised experiments and performed data analysis; and S.Y. and F.X. wrote the manuscript with discussions and improvements from Y.W., P.W.d.W., H.E.X., K.M., S.Z. and R.C.S.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0447-x.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0447-x.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to F.X.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Purification of ΔCRD-FZD4. The construct of ΔCRD-FZD4 (residues 178–517) was designed with four mutations (M309L, C450I, C507F, S508Y) based on computational predictions of stabilizing mutations made with CompoMug tool[29] and with eight residues (420–427) at ICL3 replaced by rubredoxin. This construct was expressed in *Spodoptera frugiperda* (Sf9) cells with N-terminal haemagglutinin signal peptide, Flag tag and 10× His tag as previously described[30]. The cell membrane was washed with a low-salt buffer containing 10 mM HEPES (pH 7.5), 20 mM KCl, 10 mM $MgCl_2$ and protease inhibitor cocktail (Roche). Then the membrane was washed with a high-salt buffer containing 10 mM HEPES (pH 7.5), 1 M NaCl, 20 mM KCl, 10 mM $MgCl_2$ and protease inhibitor cocktail. To avoid aggregation by free cysteine, the membrane was incubated with 2 mg ml^{-1} iodoacetamide (Sigma) at 4 °C for 30 min. After incubation, the membrane solution was incubated with solubilization buffer containing 10 mM HEPES (pH 7.5), 20 mM KCl, 10 mM $MgCl_2$ and 2% (w/v) lauryl maltose neopentyl glycol (LMNG; Anatrace) at a ratio of 1:1 at 4 °C for 1 h. The sample was centrifuged at 160,000*g* for 40 min to remove debris, and the supernatant was incubated with TALON IMAC resin (Clontech) at 4 °C overnight. Then, the resin was washed with 50 column volumes of 50 mM HEPES (pH 7.5), 500 mM NaCl, 10% (v/v) glycerol, 1% (w/v) LMNG, 10 mM $MgCl_2$, 8 mM ATP and 10 mM imidazole, followed by 6 column volumes of 50 mM HEPES (pH 7.5), 500 mM NaCl, 10% (v/v) glycerol, 0.05% (w/v) LMNG and 40 mM imidazole. Fluorescent dye Cy3 NHS ester (GE Healthcare) was then added to the resin to a final concentration of 10 μM (to monitor the protein crystals). After incubation at 4 °C for 2 h, free fluorescent dye was removed by washing with 200 column volumes of 25 mM HEPES (pH 7.5), 500 mM NaCl, 10% (v/v) glycerol and 0.01% (w/v) LMNG. The receptor was then eluted with 25 mM HEPES (pH 7.5), 500 mM NaCl, 10% (v/v) glycerol, 0.01% (w/v) LMNG and 200 mM imidazole. Imidazole was removed using a PD MiniTrap G-25 column (GE Healthcare).

Crystallization of ΔCRD-FZD4. The receptor was concentrated to about 40 mg ml^{-1} with a 50-KDa cutoff concentrator (Millipore), and reconstituted into the lipidic cubic phase by mixing with monoolein at a protein/lipid ratio of 1:1.2. The crystallization trials were set up by crystallization robot NT8 (Formulatrix). The crystals emerged under the conditions of 100 mM sodium cacodylate trihydrate (pH 6.0), 80 mM $MgSO_4$, 30% PEG400 and 1.5–2.5% v/v (±)-2-methyl-2,4-pentanediol at 3 days in the sponge phase.

Data collection and structure determination. The X-ray diffraction data were collected at Spring-8 beam line 41XU, Hyogo, Japan, using a PILATUS detector (X-ray wavelength 1.0000 Å). A rastering system was used to find the best diffracting region of single crystals[31]. The crystals were exposed for 0.2 s and 0.2° oscillation per frame. XDS[32] was used for integrating and scaling data from the 36 crystals. A molecular replacement method with Phaser[33] was applied to obtain the initial phase information using the structures of the receptor portion of SMO (PDB ID: 4JKV) and of rubredoxin (PDB ID: 1IRN) as search models. The refinement was performed with Phenix[34] and Buster[35] followed by manual examination and rebuilding of the refined coordinates in Coot[36] using both $|2F_o| - |F_c|$ and $|F_o| - |F_c|$ maps. Ramachandran plot analysis of the final structures with MolProbity showed that 100% of the residues are in either favoured (95.2%) or allowed (4.8%) regions, with no outlier. The final data collection and refinement statistics are shown in Extended Data Table 1a.

Modelling and molecular dynamics simulations of full-length FZD4. Processing of the protein structure was performed with the Protein Preparation Wizard tool[37] in Schrödinger Suite 2015-4. ICL3 (cut in the construct) was built using the Prime tool[38] in Schrödinger. Calculations of pocket volume and cross-sectional area were performed using the program Cavity[39]. The positions of receptors in the membrane were obtained from the OPM database[40]. Amino acid conservation scores were obtained from the ConSurf server[41]. Generation of conservation scores, worm representation and calculations of hydrophilic/hydrophobic surface area ratio (based on non-carbon and carbon atoms) were performed with UCSF Chimera[42]. A full-length model of FZD4 was built using a multi-template homology modelling tool implemented in ICM-Pro (Molsoft). The model used the current structure of FZD4 TMD and CRD (PDB: 5CM4) structurally aligned to the full-length structure of SMO (PDB: 5L7D). The FZD4 polypeptide chain (residues 44–513) was threaded through the TMD and CRD structures and thoroughly optimized by conformational sampling.

Molecular dynamics simulations were performed with GROMACS 5.1.2[43] using force field CHARMM36[44]. Crystal structures of FZD4 were embedded into a pre-equilibrated POPC (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine) lipid bilayer with water using the membed tool in the program GROMACS. Sodium ions were added to a concentration of 0.15 M in water, and chloride ions were added to neutralize the system. Molecular dynamics simulations were run independently three times. First, atom velocity was generated at a temperature of 310 K. Then the system was relaxed in a canonical (NVT) ensemble for 300 ps and balanced in position-restrained molecular dynamics (isothermal–isobaric (NPT) ensemble with pressure of 1 atm, using semi-isotropic coupling) for 15 ns (total energy was stable). Finally, productive molecular dynamics with no position restraints was run for 50 ns. The coordinates of atoms were recorded every 1 ps. The input files for molecular dynamics simulations of the full-length FZD4 models were generated using the CHARMM-GUI[45] server, and SMO-derived lipid coordinates obtained from the OPM[46] server were used to embed the full-length model in the POPC layer. In three independent runs, the system was equilibrated under the NVT ensemble for 500 ps (1-fs time-step) and the NPT for 15 ns (2-fs time-step) with stepwise reduction in positional restraints on protein atoms, before the final production run of 1,500 ns.

Molecular dynamics simulations of ΔCRD-FZD4. All-atom atmospheric simulations of ΔCRD-FZD4 were performed using GROMACS5.0.6[43] in the NPT ensemble with periodic boundary conditions and the CHARMM36m force field[47]. The receptor was prepared for simulation by removing all heteroatoms and the rubredoxin fusion partner. Second, thermostabilizing mutants were reverted back to wild type with Modeller 9.18[48] and aligned for membrane insertion using the Orientations of Proteins database PPM server[46]. Titratable residues were left in their dominant state at pH 7.0 and all histidines were represented with a hydrogen on the epsilon nitrogen. The resulting ΔCRD-FZD4 was capped with neutral acetyl and methylamine groups and embedded into a pre-equilibrated POPC lipid bilayer solvated in a box of TIP3P waters allowing for at least 14 Å of padding on all sides with 150 mM NaCl, and neutralised by removing appropriate ions or counter ions using the Desmond system builder within Maestro (Schrödinger Release 2018-1: Maestro, Schrödinger). Final system dimensions were 96 × 96 × 114 Å and the system comprised 185 lipids, 16,722 water molecules, 46 chloride ions and 48 sodium ions.

Before production simulations, 50,000 steps of energy minimization were performed followed by equilibration in the NVT and NPT ensembles for 10 and 50 ns, respectively, with positional restraints (1,000 kJ mol^{-1} nm^{-2}) placed on heavy atoms. A second round of NTP equilibration for 50 ns was run with positional restraints (1,000 kJ mol^{-1} nm^{-2}) on backbone atoms to allow for sidechain relaxation. System temperature was maintained at 310 K using the v-rescale method with a coupling time of 0.1 ps and pressure was maintained at 1 bar using the Berendsen barostat with a coupling time of 1.0 ps and compressibility of 4.5×10^{-5} bar^{-1} with semi-isotropic coupling. Simulations were performed with a 2-fs timestep and all bond lengths were constrained using LINCS. Electrostatic interactions were computed using the particle mesh Ewald (PME) method with non-bonded interactions cut at 10.0 Å.

Three independent 3-μs production simulations of ΔCRD-FZD4 were performed using the Parrinello–Rahmam barostat with a coupling time of 5.0 ps for a combined total of 9 μs. During production, trajectory snapshots were saved every 10 ps. Simulation analysis was performed using MDTraj 1.7.2[49] and VMD 1.9.2[50]. Plots were generated using the R statistical package (http://www.R-project.org). System parameters and trajectories are available upon request.

Cell-based luciferase assay. Mutations were introduced by QuickChange site-directed mutagenesis (Stratagene). All constructs and mutations were sequence-verified. HEK293 cells were maintained in DMEM (Gibco) with 5% fetal bovine serum. Cell-based luciferase assays were performed as previously described[51] with small modifications. In brief, cells were split at 20,000 per well in a 24-well plate 24 h before transfection. Cells were transfected with 10 ng FZD4-YFP expression plasmid, 10 ng LRP6 expression plasmid, 100 ng SuperTOPflash TCF-luciferase reporter, 10 ng ligand expression vector (WNT3A or Norrin or vehicle) and 1 ng phRGtkRenilla control plasmid using X-tremeGene 9 (Roche) per well. Cells were collected and lysed 48 h after transfection. Luminescence activities were measured using the Dual Luciferase Kit (Promega) according to the manufacturer's instructions. Renilla luciferase serves as transfection control. All activities were normalized (relative activities) to the activity of basal wild-type FZD, which was set as 1.0.

Fluorescence microscopy and protein expression. Wild-type and mutant FZD4-YFP fluorescence were visualized in non-stimulated cells 48 h after transfection using a Nikon Eclipse TE300 mercury lamp microscope. YFP fluorescence was excited at 488 nm, and emission was detected at 496–518 nm with 5-s exposure. Each corresponding YFP view was also taken in bright field with 200-ms exposure. We used a plasmid expressing C-terminally sfGFP-tagged arrestin protein, which localizes largely to the cytoplasm, as negative control for cell membrane localization[52].

After fluorescence images were taken, the cells were lysed in CelLytic M (Sigma-Aldrich) with 1 × protease inhibitor mixture (Roche Diagnostics) for SDS–PAGE analysis. FZD4-YFP bands were imaged using a ChemiDoc MP Imaging System (Bio-Rad) with Alexa 488 filter setting and 600-s exposure time.

Cell surface biotinylation assay. Cell surface biotinylation assays were carried out as previously described[52]. In brief, HEK293 cells were seeded at a density of 0.8×10^6 per ml in 6-well plates and transfected the following day with 0.5 μg of FZD4-YFP expression plasmid and 0.5 μg LRP6 expression vector with

Lipofectamine 2000 (Invitrogen) transfection reagent. After 48 h, plates were placed on ice, medium was carefully aspirated and cells were washed with cold PBS (20 mM potassium phosphate, pH 7.4, 150 mM NaCl). The cells were then incubated with 1 ml per well of PBS with 0.25 mg ml^{-1} EZlink-Sulfo-NHS-SS-biotin (Pierce) for 40 min at 4 °C. Biotinylation was stopped by the addition of 100 mM glycine in PBS (1 ml). Cells were washed once with cold TBS and lysed in CelLytic M (Sigma-Aldrich) with 1× protease inhibitor mixture (Roche Diagnostics). Non-solubilized material was removed by centrifugation (10 min at 20,000*g* and 4 °C). 20 μl of supernatant plus 20 μl of 2× SDS loading buffer served as input. The remaining supernatant was incubated with prewashed Streptavidin MagBeads (GenScript) for about 30 min at 4 °C, followed by three washes with lysate buffer. Biotinylated proteins were eluted by reduction of the NHS-SS-biotin bond with SDS loading buffer containing 100 mM dithiothreitol for 20 min at 65 °C. Eluates were subjected to SDS–PAGE for fluorescent imaging and western blot analysis using anti-actin or anti-α-tubulin antibodies as loading control. For input samples, the fluorescent images were exposed for 30 s and for elute samples for 600 s.

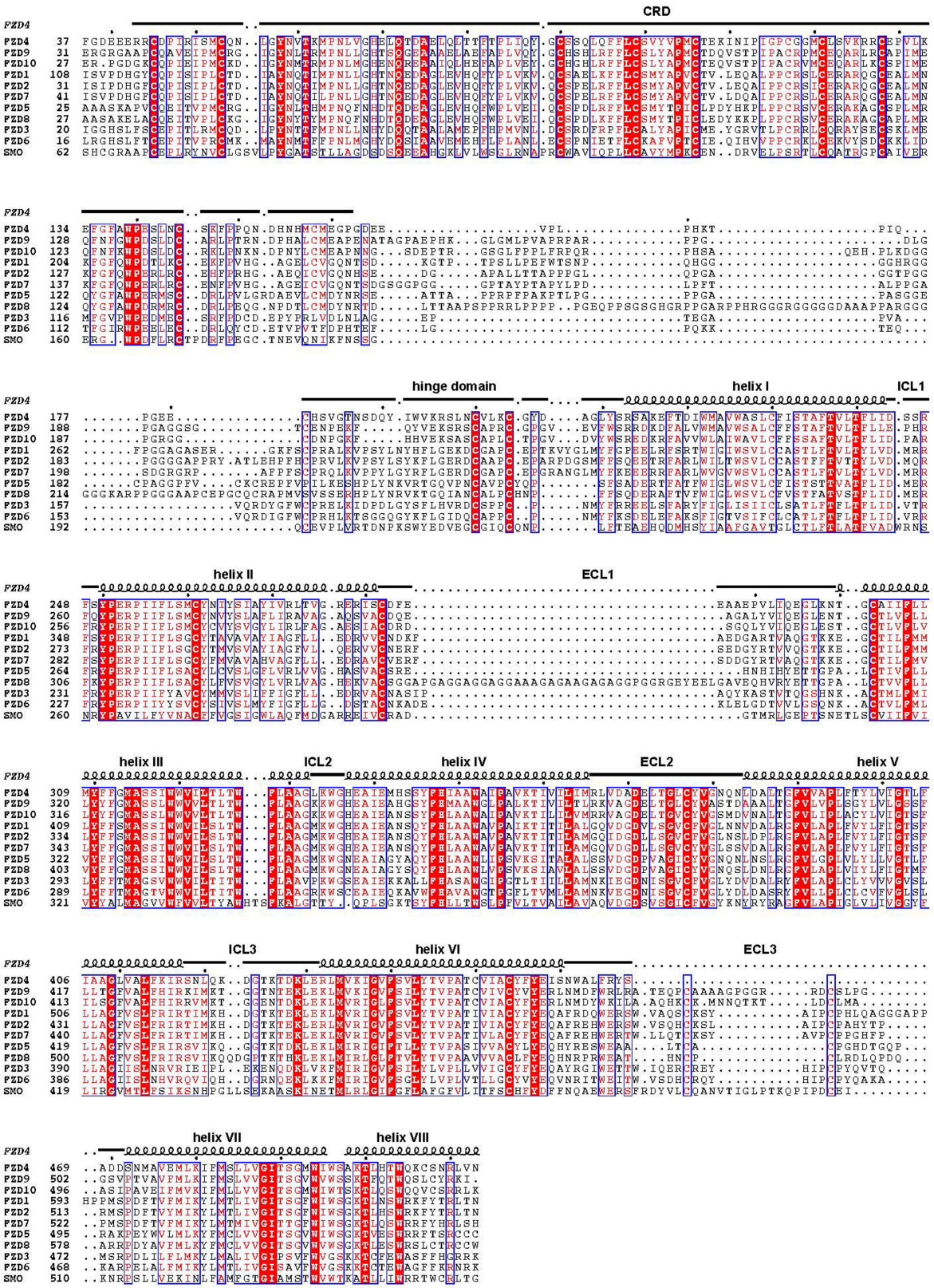
Detection of receptor and ligand interaction with affinity mass spectrometry analysis. The purified FZD4 or SMO proteins were incubated with LY2940680 at a final concentration of 250 nM (protein) and 50 nM (compound) at 4 °C for 60 min. Then the ligand-bound protein complexes were separated from free ligands by ultracentrifugation as previously described[53]. Purified A2a receptor that underwent the same process served as a negative control. Compounds released from protein complexes were analysed by an Agilent 6530 TOF system equipped with Agilent 1260 HPLC with a reported method[53]. Experimental triplicates were prepared for each pair of the receptor of interest and the negative control. Responses of specific compounds were extracted using MassHunter software (Agilent) based on the accurate mass measurement (<10 p.p.m. error) and matching retention time of the compound standard (<0.1 min shift). S/C ratios refer to the ratio of the mass spectrometry response of a specific compound detected in the protein incubation sample compared with the control. Average S/C >2 (r.s.d. <30%) or presence of the compound in the protein incubation sample only indicates positive binding of the compound to the protein target[53].

Statistics and reproducibility. For TOPflash reporter assays, each data point was determined from three independent transfectants ($n = 3$), represented as mean ± s.e.m. Sample sizes were selected by power analysis for a 95% confidence interval to detect a 1.5-fold change relative to wild type, for a standard deviation of wild-type values of 0.1 and mutant values of 0.2 (statistical power = 98.7%). Cell surface expression data were obtained using two independent methods—fluorescence microscopy and cell surface biotinylation—with similar results, and a subset of the fluorescence microscopy data was further validated independently by another laboratory member. The majority of transfections and TOPflash assays were independently repeated at least once in triplicate with similar results. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

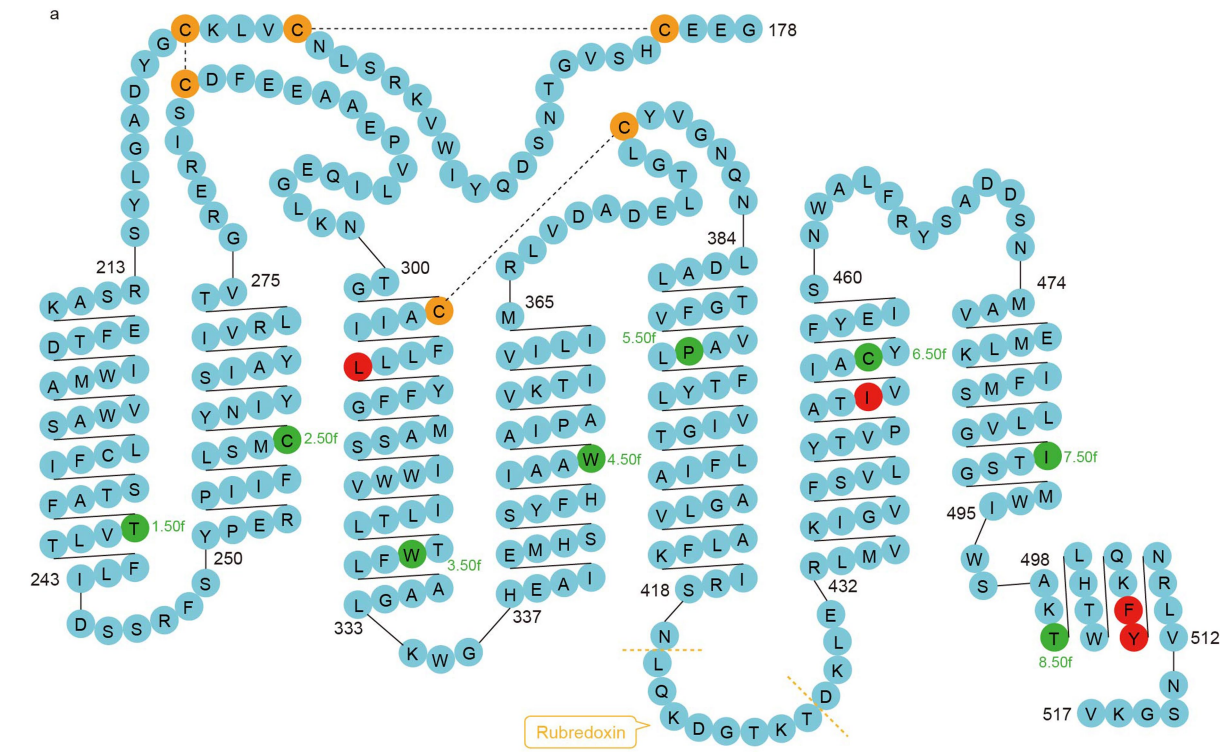
Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data availability. Coordinates and structure factors for ΔCRD-FZD4 have been deposited in the Protein Data Bank (PDB) with the accession number 6BD4. All other data relating to this study are available from the corresponding author on reasonable request.

29. Popov, P. et al. Computational design of thermostabilizing point mutations for G protein-coupled receptors. *eLife* **7**, e34729 (2018).
30. Ma, Y. et al. Structural basis for apelin control of the human apelin receptor. *Structure* **25**, 858–866.e4 (2017).
31. Cherezov, V. et al. Rastering strategy for screening and centring of microcrystal samples of human membrane proteins with a sub-μm size X-ray synchrotron beam. *J. R. Soc. Interface* **6**, S587–S597 (2009).
32. Kabsch, W. Xds. *Acta Crystallogr. D* **66**, 125–132 (2010).
33. McCoy, A. J. et al. *Phaser* crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
34. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
35. Smart, O. S. et al. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. D* **68**, 368–380 (2012).
36. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
37. Pau, M. S., Gao, S., Malbon, C. C., Wang, H. Y. & Bertalovitz, A. C. The intracellular loop 2 F328S Frizzled-4 mutation implicated in familial exudative vitreoretinopathy impairs Dishevelled recruitment. *J. Mol. Signal.* **10**, 5 (2015).
38. Kramer, G. D., Say, E. A. & Shields, C. L. Simultaneous novel mutations of LRP5 and TSPAN12 in a case of familial exudative vitreoretinopathy. *J. Pediatr. Ophthalmol. Strabismus* **53**, e1–e5 (2016).
39. Yuan, Y., Pei, J. & Lai, L. LigBuilder 2: a practical *de novo* drug design approach. *J. Chem. Inf. Model.* **51**, 1083–1091 (2011).
40. Lomize, M. A., Lomize, A. L., Pogozheva, I. D. & Mosberg, H. I. OPM: Orientations of Proteins in Membranes database. *Bioinformatics* **22**, 623–625 (2006).
41. Ashkenazy, H. et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).
42. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
43. Abraham, M. J. et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
44. MacKerell, A. D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
45. Lee, J. et al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).
46. Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370–D376 (2012).
47. Huang, J. et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).
48. Eswar, N. et al. Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinformatics* **15**, 5.6.1–5.6.30 (2006).
49. McGibbon, R. T. et al. MDTraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.* **109**, 1528–1532 (2015).
50. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **14**, 33–38 (1996).
51. Xu, T. H. et al. Alzheimer's disease-associated mutations increase amyloid precursor protein resistance to γ-secretase cleavage and the Aβ42/Aβ40 ratio. *Cell Discov.* **2**, 16026 (2016).
52. Yan, Y. et al. Dimerization of the transmembrane domain of amyloid precursor protein is determined by residues around the γ-secretase cleavage sites. *J. Biol. Chem.* **292**, 15826–15837 (2017).
53. Chen, X. et al. A ligand-observed mass spectrometry approach integrated into the fragment based lead discovery pipeline. *Sci. Rep.* **5**, 8361 (2015).



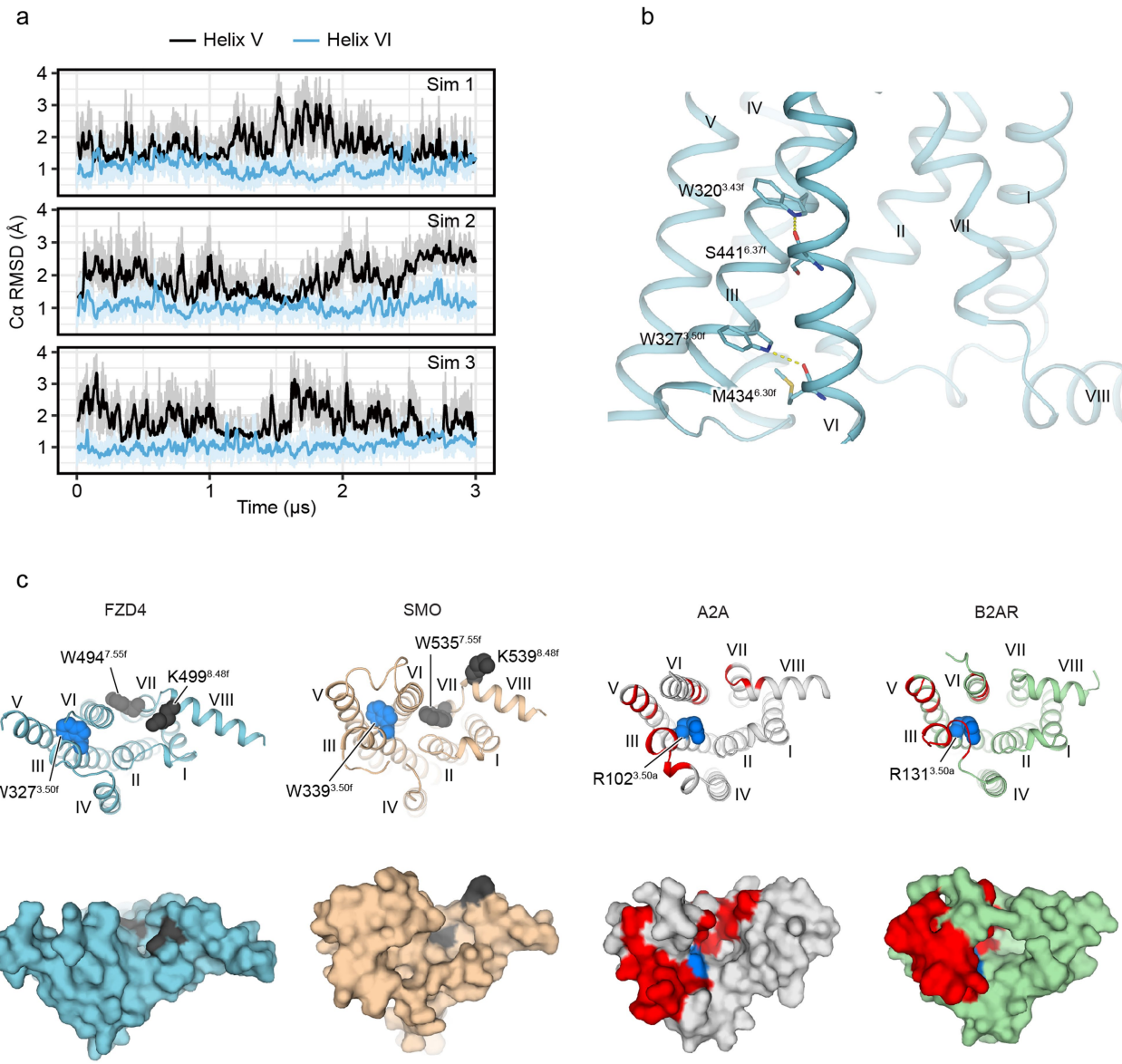
Extended Data Fig. 1 | Sequence alignment between FZD4 and nine other FZDs as well as SMO. Colours represent the similarity of residues: red background, identical; red text, strongly similar. The alignment was generated using MAFFT (https://www.ebi.ac.uk/Tools/msa/mafft/) and the graphic was prepared on the ESPript 3.0 server (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi).



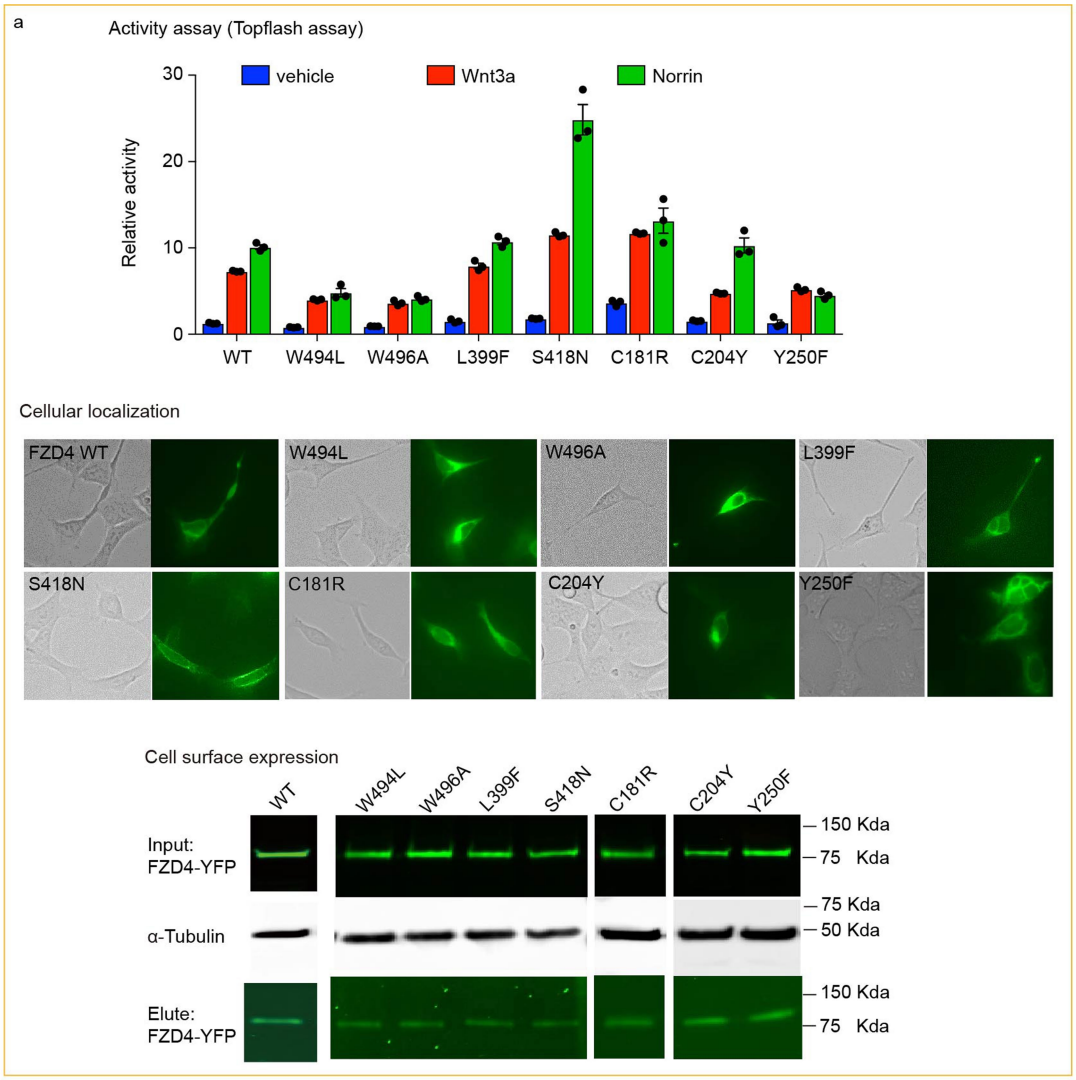
LR (Surface Expression)	Construct 941(1)	Construct 11269(1)
% of Total	78.35%	63.25%
X-Mean	221.39	140.96

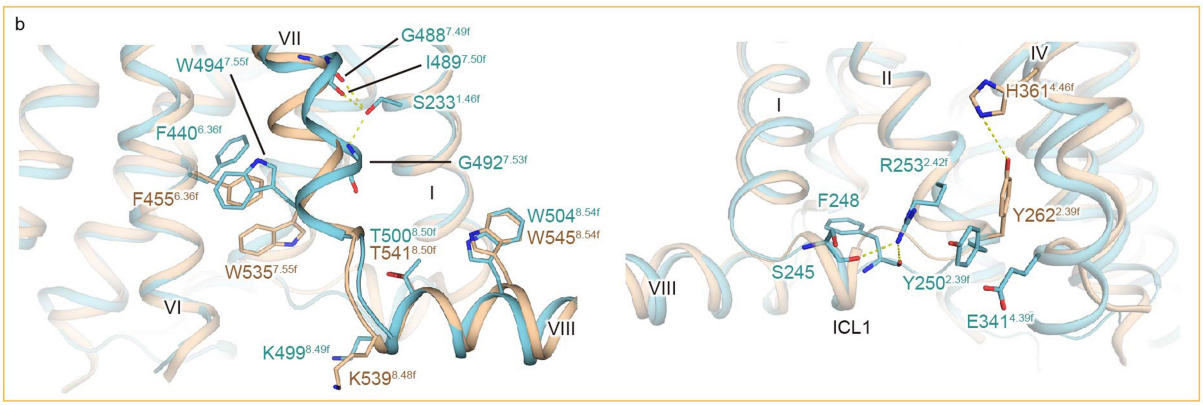
(1) Construct 941: N-BRIL-FZD4 (residues 178-537)
(2) Construct 11269: Crystallization construct

Extended Data Fig. 2 | Crystallization of ΔCRD-FZD4 and structure determination. a, Schematic of the ΔCRD-FZD4 construct. To obtain crystals that would diffract well, we truncated the N-terminal CRD region (residues 1–177) and C-terminal flexible region (residues 517–537) and introduced four single mutations (M309L, C450I, C507F and S508Y, coloured in red) that are designed based on sequence conservation analysis across ten human FZDs. Residues that are involved in the X.50f numbering system are coloured in green. The cysteines that form endogenous disulfide bonds are indicated in orange. **b**, Fluorescence-activated cell sorting staining data, to monitor the surface expression of the construct used in this study. The experiment was repeated twice with similar results. **c**, Crystals of ΔCRD-FZD4 in the apo state. **d**, Crystal packing of ΔCRD-FZD4.

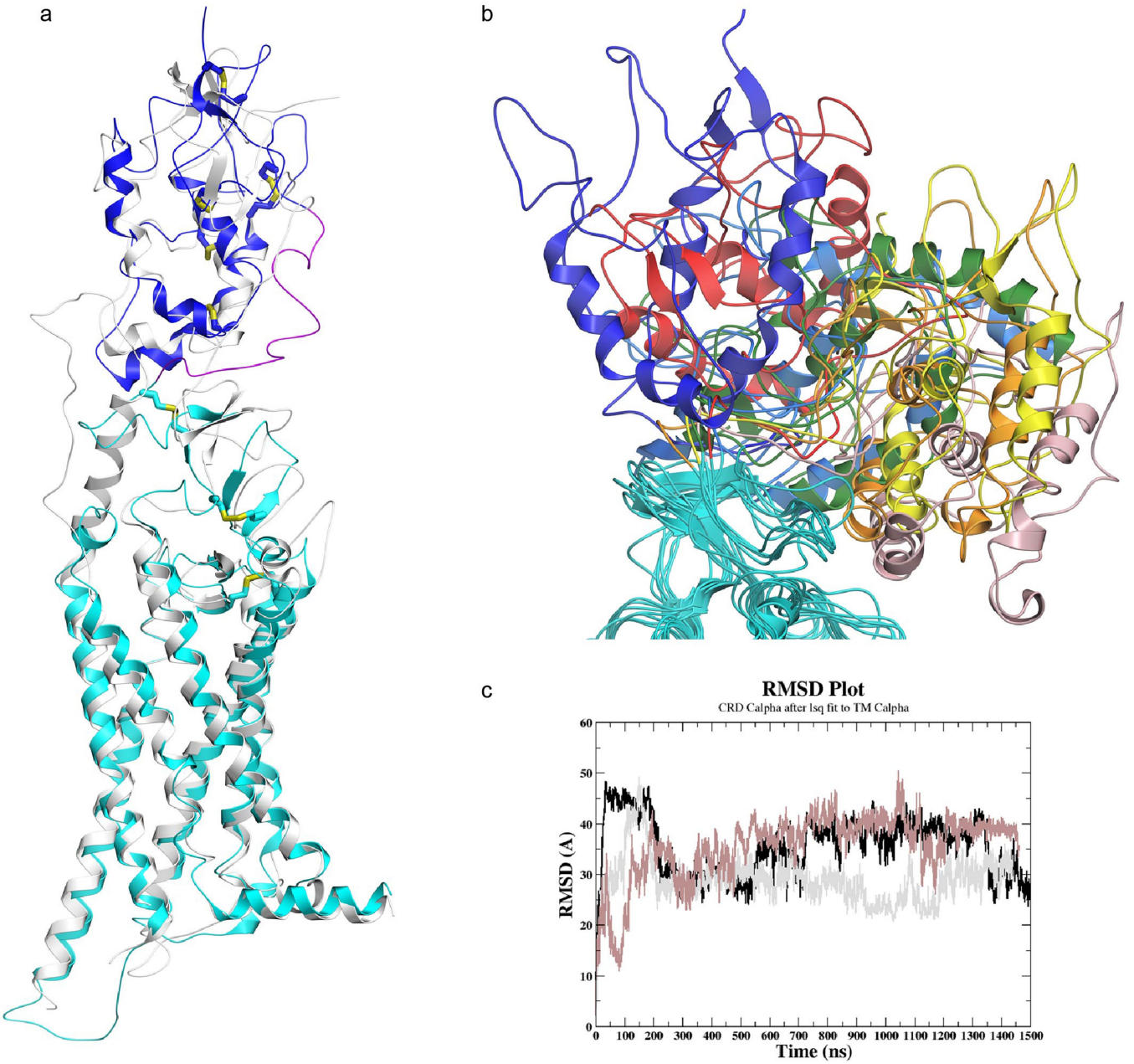


Extended Data Fig. 3 | Dynamics of helix V and VI in FZD4 and a comparison of the intracellular side of FZD4 with other GPCRs. **a**, Cα r.m.s.d. for residues G409 to S418 (helix V) and R432 to S441 (helix VI) plotted as a rolling average over a 10-ns window. **b**, Unique hydrogen bonds between helix III and helix VI ($W320^{3.43f}$–$S411^{6.37f}$ and $W327^{3.50f}$–$M434^{6.30f}$) maintain helix VI in an inward conformation. **c**, The intracellular cavity of FZD4 is close to $W^{7.55f}$ and the KTXXXW motif (coloured in dark grey), a region that is key for downstream signalling. Compared with class-A GPCRs (adenosine A_{2A} and β2 adrenergic receptors), FZD4 leaves no cavity among helices III,VI and VII, whereas SMO has a side-pocket at this position. $R^{3.50}$ (labelled with a blue sphere) is a key residue in the activation of class-A GPCRs and is exposed to the intracellular surface. The residue in the same position in class-F, $W^{3.50f}$, points in a different direction and is not exposed. The G-protein-binding sites of class-A GPCRs are labelled in red.

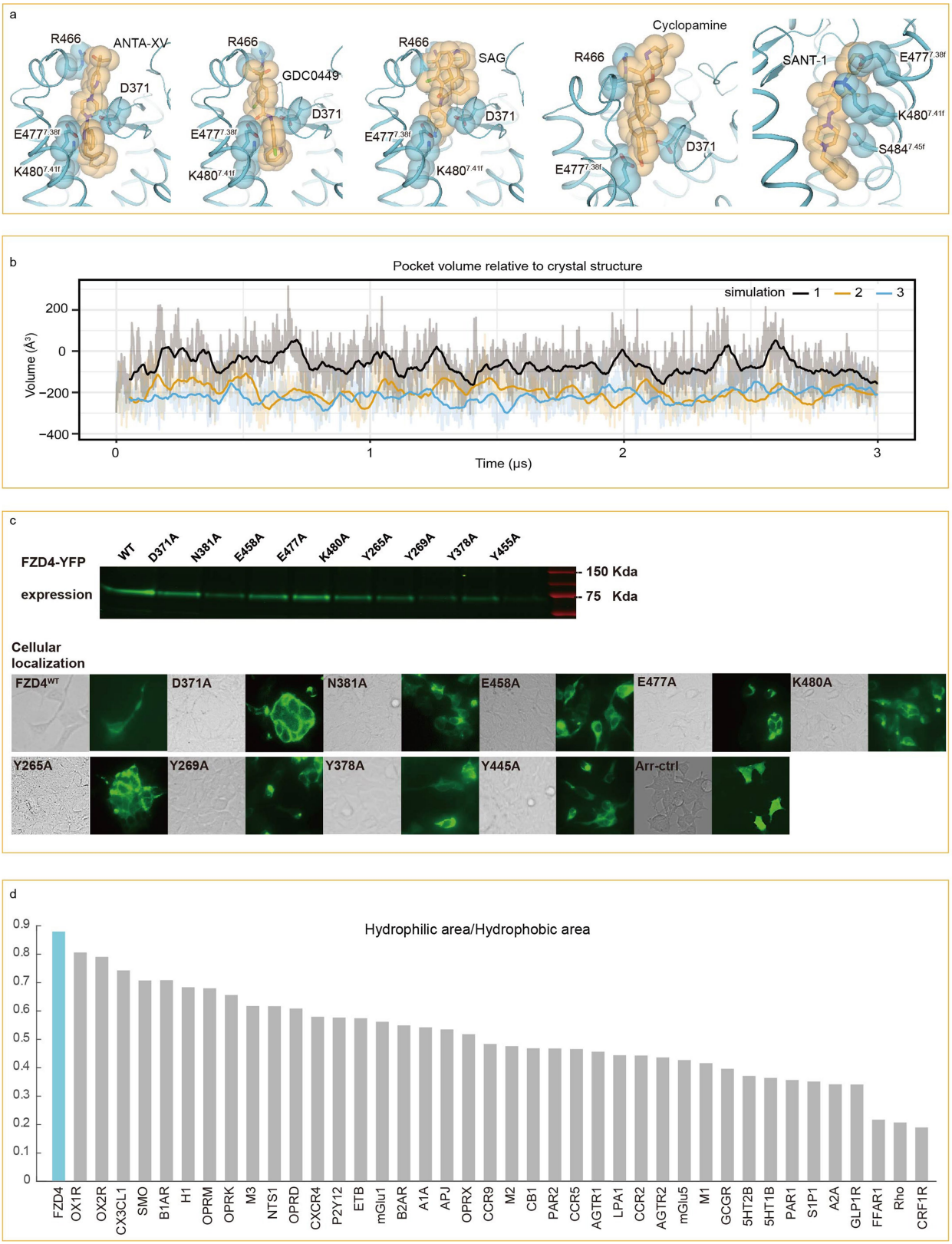




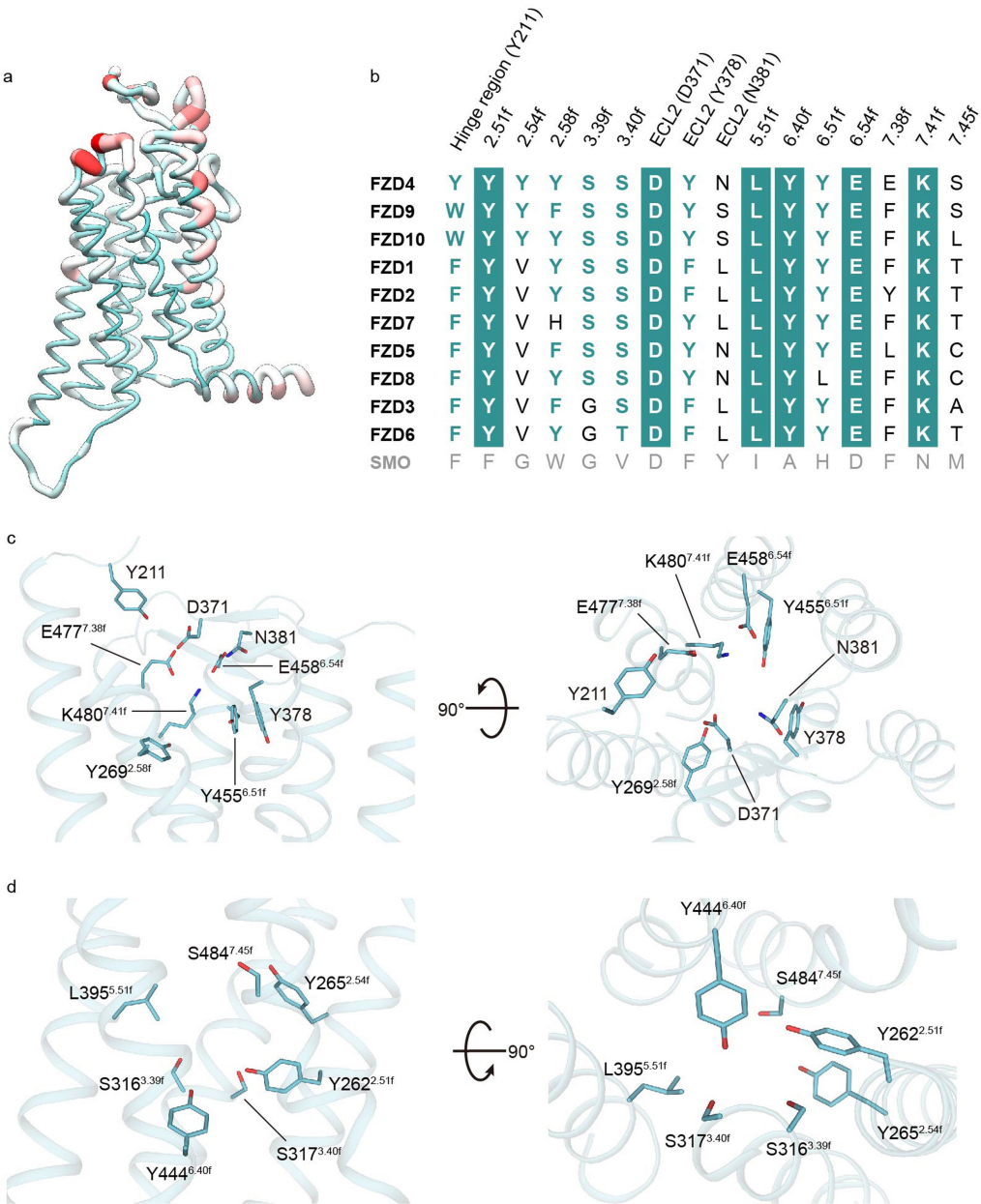
Extended Data Fig. 4 | Activity analysis of disease-related and signalling-related mutations of ΔCRD-FZD4, and the conformational change of family-conserved amino acids of ΔCRD-FZD4. a, TOPflash analysis, cellular localization and cell surface expression analysis of disease-related and signalling-related mutations. Each data point represents mean ± s.e.m., repeated in triplicate. The experiment was repeated using two independent methods with similar results. All the mutations affected FZD4 downstream signalling in some way—some decreased the WNT3A–Norrin signal substantially, some showed gain-of-function activity—which suggests complex mechanisms underlying these mutation-caused diseases. **b**, Conformational rearrangement on family-conserved residues W494$^{7.55f}$ and Y250$^{2.39f}$ was observed when comparing FZD4 (cyan) with SMO (gold). It is noteworthy that Y250$^{2.39f}$ in FZD4 points outwards from the 7TM bundle.



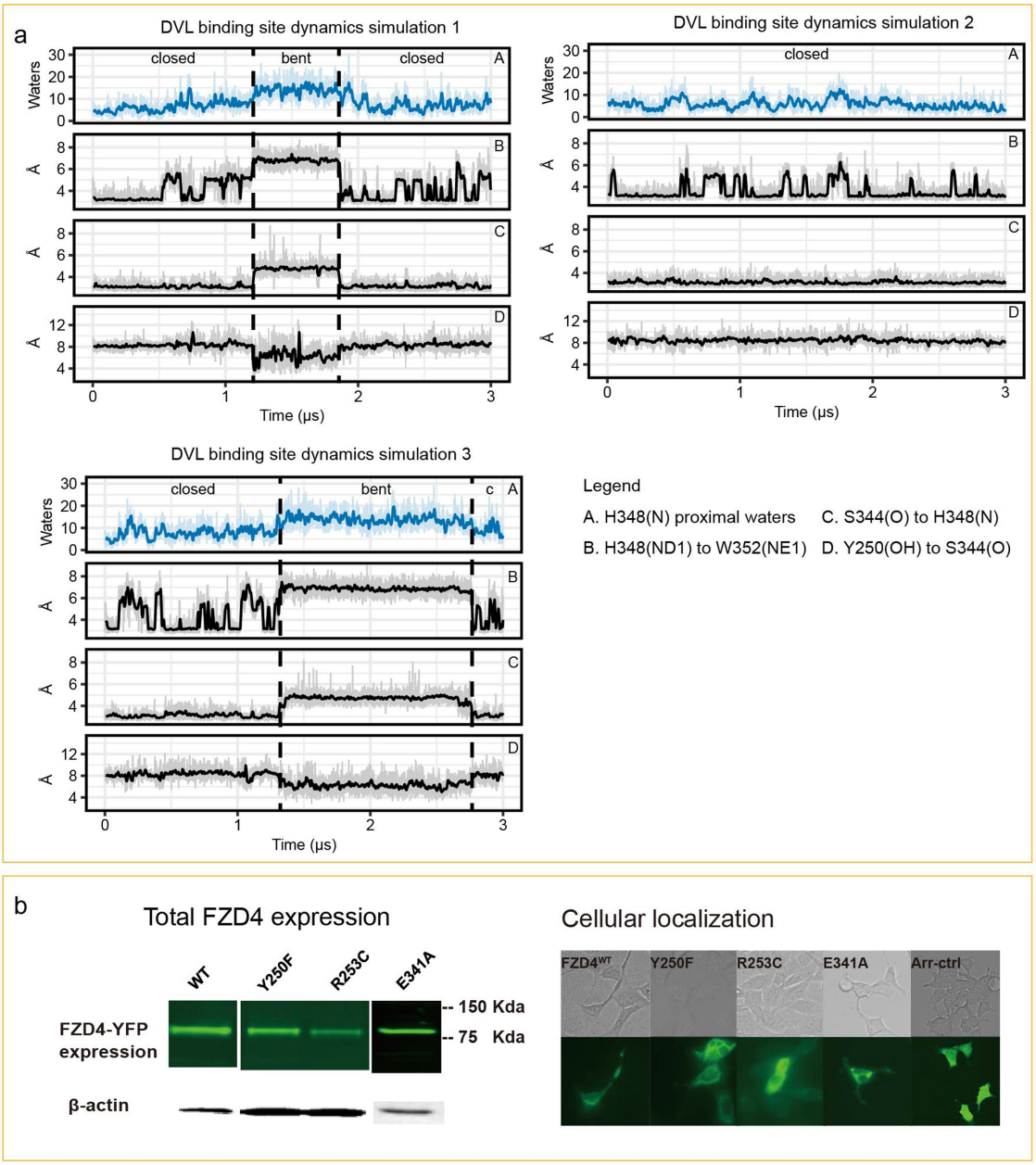
Extended Data Fig. 5 | The model of full-length FZD4 and molecular dynamics analysis. a, The model of full-length FZD4. Cyan, TMD (crystal structure); magenta, linker; blue, CRD (modelled from PDB ID: 5CM4). The human smoothened receptor crystal structure (grey; PDB ID: 5L7D) is overlaid onto the model. Disulfide bridges in the FZD4 model are shown in stick representation. **b**, Motions of the CRD domain observed during the molecular dynamics simulation. Snapshots of the positions of the CRD domain and linker during one of three independent molecular dynamics simulations are shown in cartoon representation, at 250 ns (light blue), 500 ns (green), 750 ns (pink), 1,000 ns (yellow), 1,250 ns (orange) and 1,500 ns (red). The initial full-length CRD and linker domain are shown in blue, and all TMDs are shown in cyan. **c**, Motions of the CRD domain observed during the molecular dynamics simulation. The r.m.s.d. plot of Cα atoms of the CRD and linker domain with respect to the initial full-length FZD4 model during the three molecular dynamics simulations is shown.



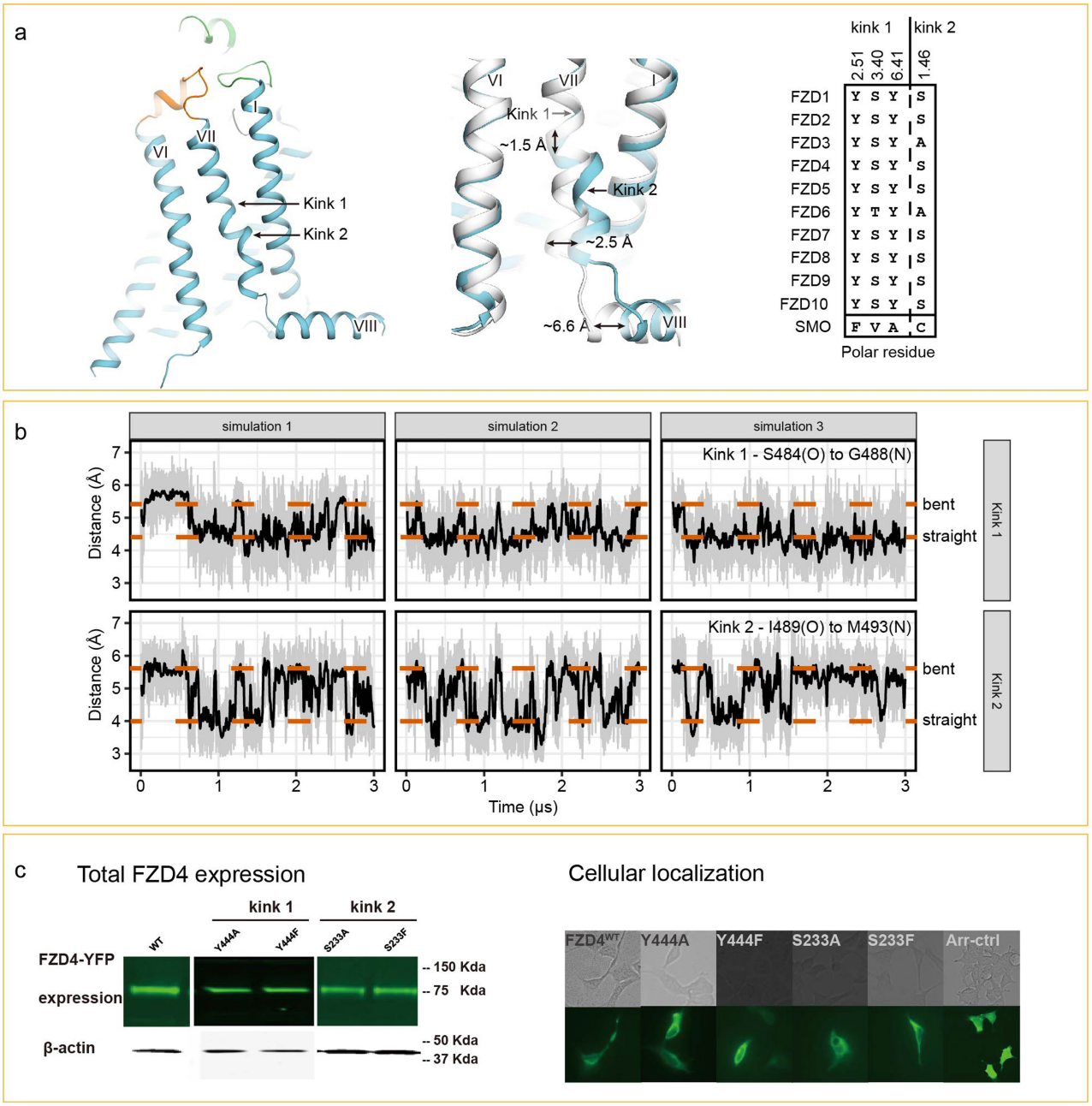
Extended Data Fig. 6 | Analysis of the TMD pocket of FZD4. **a**, Superimposition of other SMO ligands (ANTA-XV, GDC0449, SAG, Cyclopamine and SANT-1) in the pocket of FZD4. All of these SMO ligands collide with the FZD4 pocket. **b**, The volume of the TMD pocket during simulation. Volumes are displayed as rolling averages over a 10-ns window. **c**, Total expression and cellular-localization analysis for pocket mutations. Total FZD–YFP expression is determined by YFP fluorescence; cellular localization is determined by fluorescence microscopy (related to Fig. 3). A plasmid expressing C-terminally sfGFP-tagged arrestin protein (Arr-ctrl), which localizes largely to the cytoplasm, was used as negative control for cell membrane localization. The experiment was repeated using two independent methods with similar results. **d**, The FZD4 pocket has the highest hydrophilic/hydrophobic ratio of all GPCR structures that have been solved to date.



Extended Data Fig. 7 | Homology model and conservation analysis for ten human FZDs. a, Superposition of ten human FZD homology models. The green, thin regions represents high conservation, and the red, thick regions represent low conservation. **b**, Sequence alignment of pocket residues, with conserved amino acids across the FZD family highlighted in dark green. **c**, Conserved amino acids in the top of the FZD4 pocket. **d**, Conserved amino acids in the bottom of the FZD4 pocket.



Extended Data Fig. 8 | Molecular dynamics and mutation analysis of the Dishevelled-binding site. a, The Dishevelled-binding site plotted as a moving average over a 10-ns window. Water molecules less than 10 Å from H348(N) were defined as proximal for this analysis. **b**, Total expression and cellular localization analysis for mutations in the Dishevelled-binding site (related to Fig. 4). The experiment was repeated using two independent methods with similar results.



Extended Data Fig. 9 | Analysis of the two unusual kinks. a, Two kinks with conserved polar networks fluctuate between bent and straight conformations during simulation. **b**, Molecular dynamics traces of the kink 1 and kink 2 backbone distances plotted as a moving average over a 10-ns window. **c**, Total expression and cellular localization analysis of kink 1/kink 2 mutations (related to Fig. 4). The experiment was repeated using two independent methods with similar results.

Extended Data Table 1 | Crystallographic data table and affinity mass spectrometry analysis

a

	ΔCRD-FZD4
Data collection	
Space group	$C222_1$
Cell dimensions	
a, *b*, *c* (Å)	61.67, 154.69, 114.40
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	45.99-2.40 (2.49-2.40)[a]
R_{sym} or R_{merge}	8.54 (51.15)
I / σI	18.19 (3.52)
Completeness (%)	100 (100)
Redundancy	20.7 (13.2)
Refinement	
Resolution (Å)	30.00-2.40
No. reflections	21,728
R_{work} / R_{free} (%)	21.00/23.30
No. atoms	
Protein	2,996
Lipids and other	205
Wilson *B*-factors (Å^2)	76.8
B-factors (Å^2)	
ΔCRD-FZD4	90.4
Rubredoxin	111.3
Lipids and other	110.4
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	0.76

b

	SMO	FZD4 5068[a]	FZD4 11269[b]
LY2940680	99[c]	0[d]	0

a, Data collection and refinement statistics. Values in parentheses are for highest-resolution shell. A total of 36 crystals were used. **b**, Affinity mass spectrometry analysis of the LY2940680 interaction with SMO or FZD4.
[a]The FZD4 construct with N-terminal fragment (residues 1–177) replaced by BRIL.
[b]The FZD4 construct used for crystallization in this study.
[c]S/C = 99 indicates that the compound was only detected in the protein incubation sample while absent in the control, thus it specifically interacted with the corresponding receptor.
[d]S/C = 0 indicates that the compound was absent in the protein incubation sample, thus it did not interact with the corresponding receptor.

nature research

Corresponding author(s): Fei Xu

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   Due to radiation damage, X-ray diffraction data collection of the protein crystals was limited to 5-10 degree per crystal. To collect a complete data set for structure determination, diffraction data from multiple crystals were integrated and scaled using XDS. By calculating completeness of the data set, diffraction data from 36 ΔCRD-FZD4 crystals were used to ensure the completeness was 100%. For the super-TOPflash assay, cell surface biotinylation labeling and fluorescent imaging analyses, 2-3 independent experiments were performed in technical duplicate/ triplicate to ensure each data point was repeatable.

2. Data exclusions

   Describe any data exclusions.

   No data were excluded from the analyses.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   All attempts at replication were successful.

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   Randomization is not relevant to this study, as protein and crystal samples are not required to be allocated into experimental groups in protein structural studies, and no animals or human research participants are involved in this study.

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   Blinding is not relevant to this study, as protein and crystal samples are not required to be allocated into experimental groups in protein structural studies, and no animals or human research participants are involved in this study.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

XDS, Phaser, Phexin, Buster, COOT, Pymol, UCSF Chimera, GROMACS, Schrödinger Suite, ConSurf server, Cavity, MassHunter; MDTraj 1.7.2, VMD 1.9.2, R statistical package, GraphPad Prism 7.0, Image Lab 6.0.0., Compomug, ICM-Pro

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

All unique materials used are readily available from the authors or from standard commercial sources.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

Primary antibodies:
Anti-beta-actin: Abcam ab6276; Anti-alpha-tubulin: Abcam, Ab7291.
Secondary antibodies:
Anti-mouse IgG, HRP-linked Antibody: Cell Signaling Technology, 7076S
1:5000 dilution against 2% BSA in TBST buffer for the primary antibody and 1:2000 for the secondary antibody.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

sf9 cell line was obtained from Invitrogen; HEK293 cells were obtained from American Type Culture Collection (ATCC).

b. Describe the method of cell line authentication used.

The HEK293 cell line was authenticated using a PCR based multiplex assay based on the use of short tandem repeats (STR). The Sf9 cell line was authenticated through morphology check by microscope and growth curve analysis.

c. Report whether the cell lines were tested for mycoplasma contamination.

All cell lines tested are negative for mycoplasma contamination

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

No commonly misidentified cell lines were used

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	No animals were used.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	The study did not involve human research participants.