PNAS
www.pnas.org

Supplementary Information for

A new family of neural wiring receptors in bilaterians defined by phylogenetic, biochemical and structural evidence

Shouqiang Cheng, Yeonwoo Park, Justyna D. Kurleto, Mili Jeon, Kai Zinn, Joseph W. Thornton, Engin Özkan

Corresponding author: Engin Özkan
Email: eozkan@uchicago.edu

This PDF file includes:

Supplementary Methods

Phylogenetics

Putative Dpr and DIP homologs were identified using BLAST with *D. melanogaster* proteins as queries (1). To exclude distant IgSF homologs, the BLAST hits were used as queries to search over the *D. melanogaster* proteome (reciprocal BLAST), and only those with a Dpr or DIP as the top hit were retained. For identifying IgLON homologs in protostomes, human IgLONs were used as queries and reciprocal BLAST was performed on the human proteome. Amalgam, CG34353, CG7166, DIP, Dpr, Klingon, Lachesin, and Wrapper were identified as IgLON homologs in *D. melanogaster*. Only some of the 21 Dprs appeared as IgLON homologs, presumably due to their fast rate of evolution. Amalgam was excluded from analysis because it could not be reliably aligned to other proteins. Nectin, Necl, Kirrel, and Nephrin were the only other IgSF subfamilies that could be reliably aligned to the Wirin family.

Sequences for insect, bony fish, and tetrapod proteins were obtained from the NCBI protein database. To acquire sequences from other organisms, whose proteomes were generally poorly represented in the NCBI protein database, transcripts were assembled from RNA-seq reads in the NCBI SRA database (2). To selectively assemble specific transcripts rather than the entire transcriptome, RNA-seq reads with similarity to reference proteins were extracted using TBLASTN. Transcripts were then *de novo* assembled using Velvet/Oases (3, 4), and coding sequences were identified using TransDecoder (http://transdecoder.github.io) (5).

Only the Ig domains were included in the alignment because the rest of the proteins could not be aligned across paralogs. The number of Ig domains varied across proteins: two for Dpr, three for DIP, IgLON, Klingon, Lachesin, Nectin, and Necl, five for Kirrel, and eight to ten for Nephrin. The first one and a half Ig domains of Dpr aligned with the corresponding part of the 3-Ig proteins. The rest of Dpr aligned with the second half of the third Ig domain of the 3-Ig proteins. The first two Ig domains of the 3-Ig proteins aligned with the corresponding part of Kirrel and Nephrin. The alignment of the third Ig domain to the rest of Kirrel and Nephrin was ambiguous. However, the phylogeny was robust to the uncertainty in alignment. When alignments were generated using different gap penalties, the resulting phylogenies were topologically identical to that in Fig. 1. Furthermore, using only the first two Ig domains resulted in a phylogeny that is also topologically identical to that in Fig. 1.

Based on preliminary alignments and phylogenies, proteins with exceptionally long branches or nonsensical species placements were discarded. The remaining proteins were classified into paralog groups. Proteins in each paralog group were aligned separately, removing paralog-specific insertions and alignment-ambiguous regions. The full alignment was assembled from paralog alignments through sequential profile alignment. All alignments were generated using MUSCLE with default settings (6). Preliminary phylogenies were inferred using FastTree2 (7). The ML phylogenies were inferred using RAxML v8.2.12 (8). The best-fit model of evolution for the full phylogeny was WAG + G + I + X (X: ML estimation of equilibrium amino acid frequencies). The best-fit model for DIP, IgLON, and Klingon individually, however, were LG + G + I + X. Therefore, paralog relationships within each family were inferred in separate analyses. Approximate likelihood ratio statistics were calculated using PhyML v.3.0 under the topology and equilibrium amino acid frequencies inferred using RAxML (9).

The unreduced phylogenies are attached to the manuscript as datasets. The provided datasets are for the following phylogenies:

Dataset 1	supporting Figure 1	Unreduced phylogeny file for the Wirin family
Dataset 2	supporting Figure 2A	Unreduced phylogeny file for the IgLON subfamily

Dataset 3	supporting Figure S1	Unreduced phylogeny file for the Dpr subfamily
Dataset 4	supporting Figure S2	Unreduced phylogeny file for the DIP subfamily
Dataset 5	supporting Figure S3A	Unreduced phylogeny file for the Klingon subfamily
Dataset 6	supporting Figure S3B	Unreduced phylogeny file for the Lachesin subfamily

The Extracellular Interactome Assay (ECIA)

Interactions between ectodomains of Dprs, DIPs and homologs were tested using ECIA (10) with minor modifications: The promoters in the bait and prey expression vectors have been replaced with the constitutively active Actin 5C promoter from *D. melanogaster* in lieu of the inducible metallothionein promoter. The transfection agent was also changed to TransIT-Insect (Mirus), which was used according to the manufacturer's recommended protocol. Mouse IgLON cDNAs were used in the binding experiments.

Throughout the manuscript, the outcome of the assay is reported in absorbance values at 650 nm, as the AP substrate KPL BluePhos (Seracare, catalog no. 5120-0061) is turned over to a blue product.

Homology Modeling

Dpr and DIP orthologs were modeled based on the Dpr6-DIP-α structure (PDB: 5EO9) using MODELLER (11). For the NEGR1-NTM IG1-IG1 complex, Dpr6 was used to model NEGR1, DIP-α and was used to model NTM. Further side chain rotamer optimization was performed using SCWRL4 (12) and manual inspection of alternate rotamers in PyMOL (13).

Expression, Purification and Crystallization of RIG-5 IG1 and ZIG-8 IG1

The N-terminal domain of *C. elegans* RIG-5 was cloned into pAcGP67A with a C-terminal hexahistidine tag, and co-transfected into Sf9 cells with linearized baculoviral DNA (Expression Systems) using the TransIT-Insect transfection reagent (Mirus). Amplified virus was used to infect High Five cells. Media was collected 60 hours post-infection. RIG-5 was first purified using Ni-NTA agarose resin, followed by size-exclusion chromatography on a Superdex 75 10/300 column (GE Healthcare) in HBS (10 mM HEPES, pH 7.2 and 150 mM NaCl). RIG-5 was crystallized with the sitting-drop vapor diffusion method, using a Mosquito crystallization robot (TTP Labtech) with 100 nl protein + 100 nl crystallant drops against a 50-µl crystallant reservoir.

The N-terminal domain of *C. elegans* ZIG-8 was expressed and purified as above. Successful expression required the extension of the IG1 construct to include all of the N-terminal sequence between the domain and the signal peptide. We now recognize that this part constitute the additional N-terminal helix, which packs onto and is disulfide linked to the IG1. ZIG-8-RIG-5 IG1-IG1 complex was purified on a Superdex75 10/300 column.

RIG-5 IG1 sample was concentrated to 17 mg/ml, and crystallized by the sitting-drop vapor diffusion method in 0.1 M HEPES pH 7.0 and 1 M sodium citrate. Crystals were cryoprotected in 0.1 M HEPES pH 7.0, 1.2 M sodium citrate, 10% glycerol and vitrified in liquid nitrogen. Diffraction data were collected at SSRL beamline 9-2.

ZIG-8-RIG-5 was crystallized at 17.5 mg/ml in two different conditions. Crystal form 1 (tetragonal) was crystallized in 0.2 M disodium hydrogen phosphate, 20% PEG 3350, and cryoprotected with 0.2 M NaCl, 0.1 M HEPES pH 7.2, 22%PEG 3350, 24% Glycerol. Crystal form II (monoclinic) was crystallized in 0.2 M lithium sulfate, 0.1 M sodium cacodylate, pH 6.5, 30% PEG 400. This condition did not require cryoprotection. Diffraction data were collected at APS beamline 24-ID-C.

RIG-5 homodimer structure was solved by molecular replacement (MR) using a DIP-η IG1 monomer (PDB ID: 6NRX) (14) as the model with *PHASER* (15). ZIG-8-RIG-5 complex was solved by MR using the RIG-5 IG1 homodimer structure with *PHASER*. The models were refined with *phenix.refine* (16) and real-space model building was performed in *Coot* (17). Model validation was performed using *Molprobity* (18) within the *PHENIX* suite (19).

Surface Plasmon Resonance (SPR)

ZIG-8 and RIG-5 full-length ectodomain constructs were expressed and purified as above. 2430 RUs of ZIG-8 was coupled on a Biacore CM5 chip (GE Healthcare) using NHS/EDC chemistry and random amine coupling. RIG-5 ectodomain was run over the chip as analyte in HBS with 0.05% Tween-20 as surfactant and 0.1% BSA to remove non-specific binding.

Analytical Ultracentrifugation

Full ectodomain ZIG-8 and RIG-5 were used in sedimentation velocity experiments in a Beckman analytical ultracentrifuge with an An50-Ti rotor at 20ºC. Protein samples were placed in charcoal-filled Epon centerpieces sandwiched between sapphire windows. Centrifugation was done at 50,000 rpm.

The AUC data were analyzed using the *c*(*s*) methodology in SEDFIT (20, 21). Partial-specific volume, density, and viscosity were calculated using SEDNTERP (22). Partial specific volumes used for RIG-5 and ZIG-8 were 0.7192 and 0.7227 cm^3/g, respectively. Figures were rendered in GUSSI (23). For K_D analyses, GUSSI was used to integrate the *c*(*s*) distributions, which were assembled into isotherm files that were imported into SEDPHAT (24), where a monomer-dimer model was imposed, with a fixed *s*-value for the dimer (4.2 S for RIG-5, 3.65 S for ZIG-8).

Signal peptide and transmembrane helix predictions; Sequence numbering for RIG-5

There is ambiguity with regards to the N-terminal end of RIG-5 covering the signal peptide. The C36F7.4f.1 transcript on the Wormbase database (25) only allows for a weak prediction of a signal peptide with Phobius or SignalP (26, 27), while the C36F7.4g.1 transcript yields a strongly predicted signal peptide ("MYLFALLCGVLLVFKQACSRG") if the second methionine in the transcript is used as the start methionine and therefore sixty amino acids are removed from the transcript. We used a numbering scheme throughout the manuscript that uses the C36F7.4g.1 sequence with 60 amino acids removed from the N terminus. It should be noted that the mature proteins (i.e. after the signal peptides are processed) for both transcripts have identical sequences. In the manuscript, we do not make a call about which transcript(s) are actually expressed in worms.

Definitions for protein families within the IgSF

Nectins and Necls are two related families within the IgSF found in vertebrates. In humans, the family include nine members, Nectins 1 to 5 and Necls 1 to 4. It should be noted that Necl5 is more closely related to Nectins than Necls.

Here, we define Kirrels as orthologs of the *C. elegans* protein SYG-1. Kirrels are found across bilaterians, including SYG-1, Rst and Kirre (Duf) in *Drosophila*, and Kirrel1 (Neph1), Kirrel2 (Neph3) and Kirrel3 (Neph2) in vertebrates.

Nephrins are heterophilic binding partners of Kirrels, and are orthologs of the *C. elegans* protein SYG-2. Nephrins are found across bilaterians, including SYG-2, SNS and Hibris in *Drosophila*, and Nephrin in vertebrates.

Constructs used in ECIA experiments for Wirins

For ECIA experiments, expression constructs included the following residues.

ZIG-8 full-length ectodomain: Ala22 to Ser249 (Wormbase CDS Y39E4B.8).

ZIG-8 IG1: Ala22 to Pro137.

RIG-5 full-length ectodomain: Arg20 to Arg397 (Wormbase CDS C36F7.4e minus the N-terminal 60 amino acids).

RIG-5 IG1: Arg20 to Pro130.

Mouse OBCAM full-length ectodomain: Thr30 to Asn14 (NCBI Accession NP_808574.2)

Mouse NTM full-length ectodomain: Gly34 to Asn321 (NCBI Accession NP_758494.2)

Mouse NEGR1 full-length ectodomain: Val32 to Gly318 (NCBI Accession NP_ 001034183.1).

Table S1. Data and refinement statistics for x-ray crystallography of the RIG-5 homodimer and two crystal forms for ZIG-8–RIG-5 IG1-IG1 heterodimers.

	RIG-5 IG1 homodimer	**ZIG-8–RIG-5 IG1–IG1 crystal form 1**	**ZIG-8–RIG-5 IG1–IG1 crystal form 2**
PDB ID	6ON6	6ON9	6ONB
Data Collection			
Space Group	$P6_122$	$P4_12_12$	$C2$
Cell Dimensions			
a, *b*, *c* (Å)	46.61, 46.61, 196.85	80.08, 80.08,166.62	87.46, 79.10, 89.98
α, *β*, *γ* (°)	90, 90, 120	90, 90, 90	90, 92.22, 90
Resolution (Å)	50-1.42 (1.51-1.42)*	100-2.00 (2.12-2.00)	100-1.70 (1.80-1.70)
R_{sym} (%)	2.7 (20.2)	17.8 (169.3)	4.7 (58.7)
$\langle I \rangle / \langle \sigma(I) \rangle$	55.5 (9.5)	7.94 (1.31)	14.15 (1.62)
$CC_{1/2}$ (%)	100.0 (99.6)	99.5 (67.2)	99.9 (81.2)
Completeness (%)	99.6 (97.7)	99.5 (97.6)	96.8 (93.8)
Redundancy	17.7 (15.2)	13.2 (12.4)	3.6 (3.4)
Refinement			
Resolution (Å)	50-1.42 (1.48-1.42)*	100-2.00 (2.05-2.00)	100-1.70 (1.72-1.70)
Reflections	24787	37511	65514
R_{cryst} (%)	19.70 (17.63)	17.42 (28.65)	17.39 (42.52)
R_{free} (%)†	22.61 (24.18)	20.76 (30.81)	20.73 (43.50)
Number of atoms			
Protein	841	1772	3562
Ligand/Glycans	20	30	67
Water	94	180	404
Average B-factors (Å^2)			
All	34.08	52.05	38.41
Protein	33.12	51.21	37.40
Ligand/Glycans	44.73	77.53	58.25
Solvent	40.47	56.16	43.98
R.m.s. deviations from ideality			
Bond Lengths (Å)	0.013	0.018	0.006
Bond Angles (°)	1.302	1.367	0.821
Ramachandran Statistics			
Favored (%)	96.23	95.50	96.15
Outliers (%)	0.0	0.0	0.0
Rotamer outliers (%)	0.0	1.02	0.25
All-atom Clashscore‡	3.51	2.82	2.10
Coordinate error (Å)§	0.15	0.20	0.24

* The values in parentheses are for reflections in the highest resolution bin.
† 1240, 1961, and 3179 reflections (5%) were not used during refinement for cross validation.
‡ As reported by *Molprobity*.
§ Maximum-likelihood-based error estimate by *phenix.refine* version 1.15.
None of the models contained Cβ outliers.

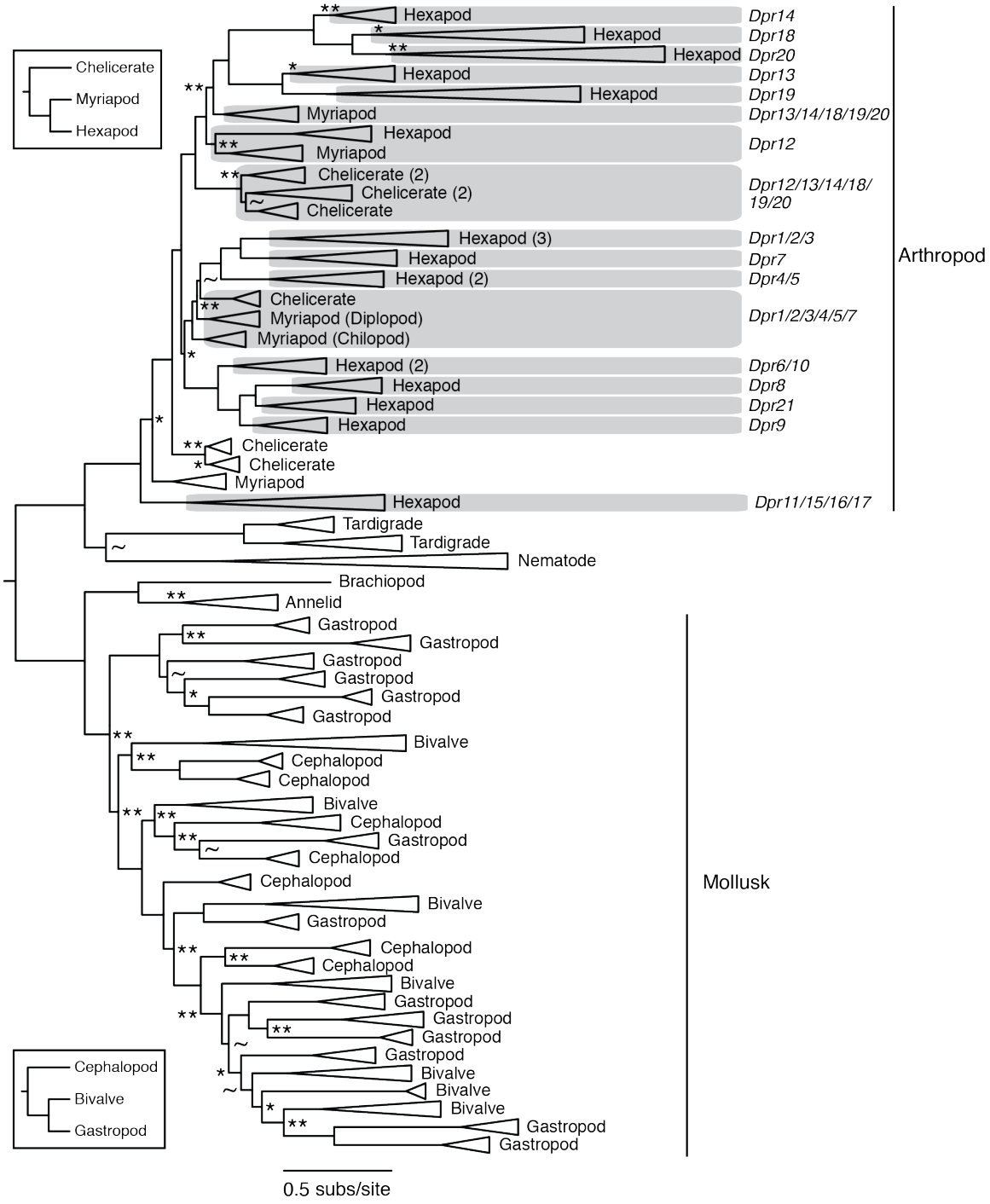

Fig. S1. The ML phylogeny of the Dpr family. Approximate likelihood ratio statistics (aLRS) are shown as branch supports: ** < 9.2 (= 2 ln100), * < 4.6 (= 2 ln10), ~ < 2.2 (= 2 ln3). Unmarked branches have aLRS > 9.2. The insets show the arthropod and mollusk phylogenies. The arthropod paralogs are labeled following the *D. melanogaster* Dpr nomenclature. The numbers next to some clades show the maximum number of paralogs in a clade when there are gene duplications in its subclades.

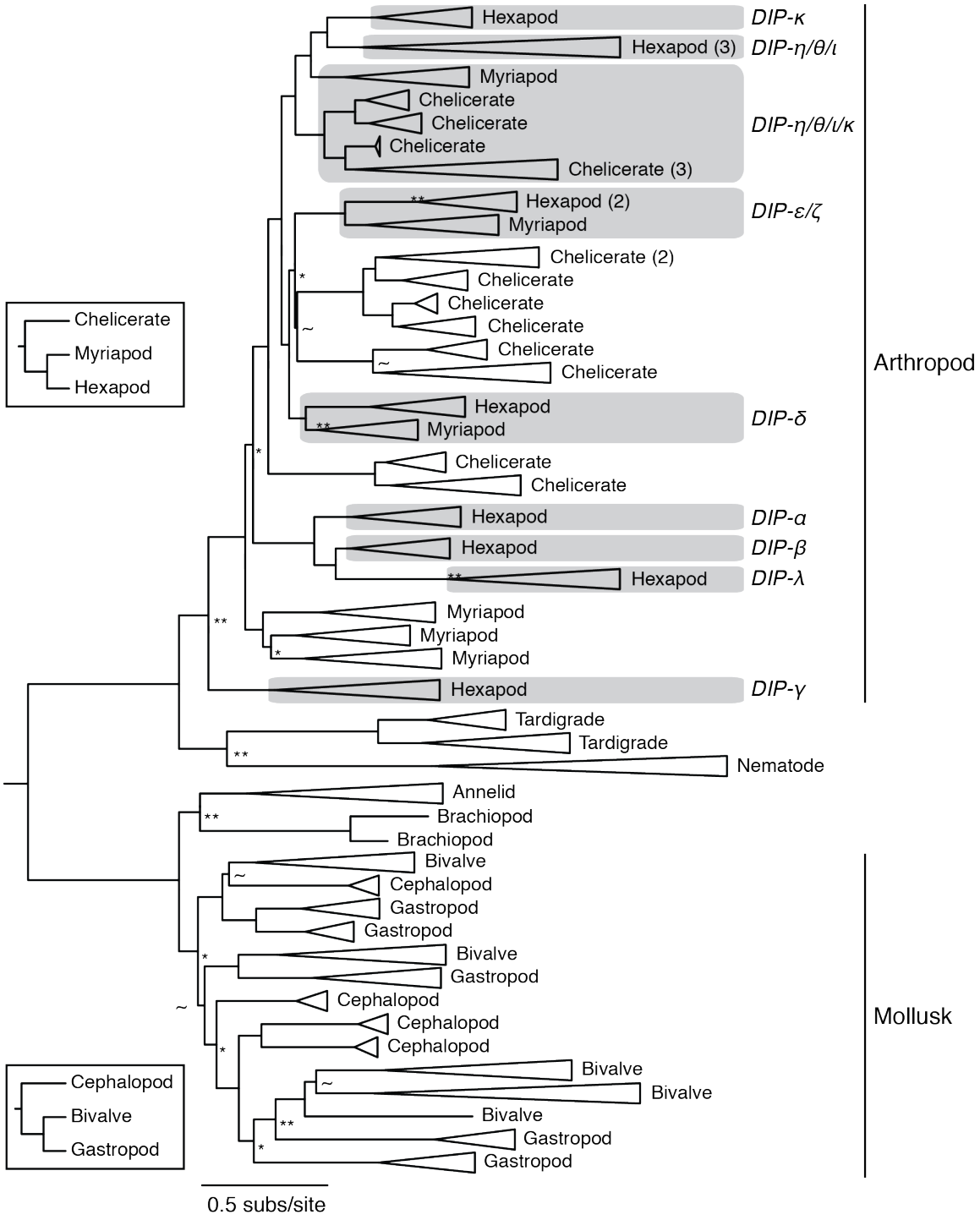

Fig. S2. The ML phylogeny of the DIP family shown as in Fig. S1.

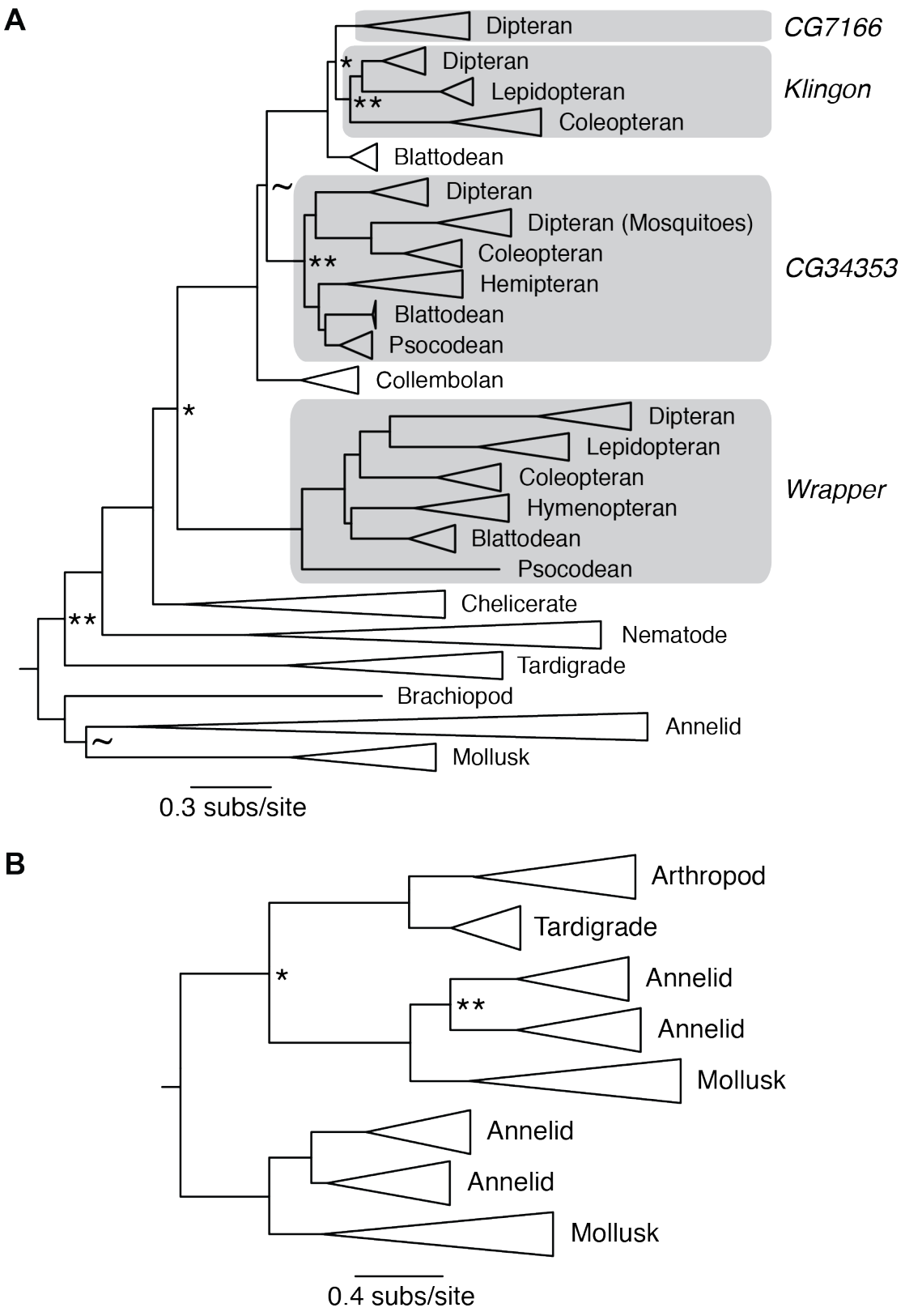

Fig. S3. A. The ML phylogeny of the Klingon family shown as in Fig. S1.
B. The ML phylogeny of the Lachesin family shown as in Fig. S1.

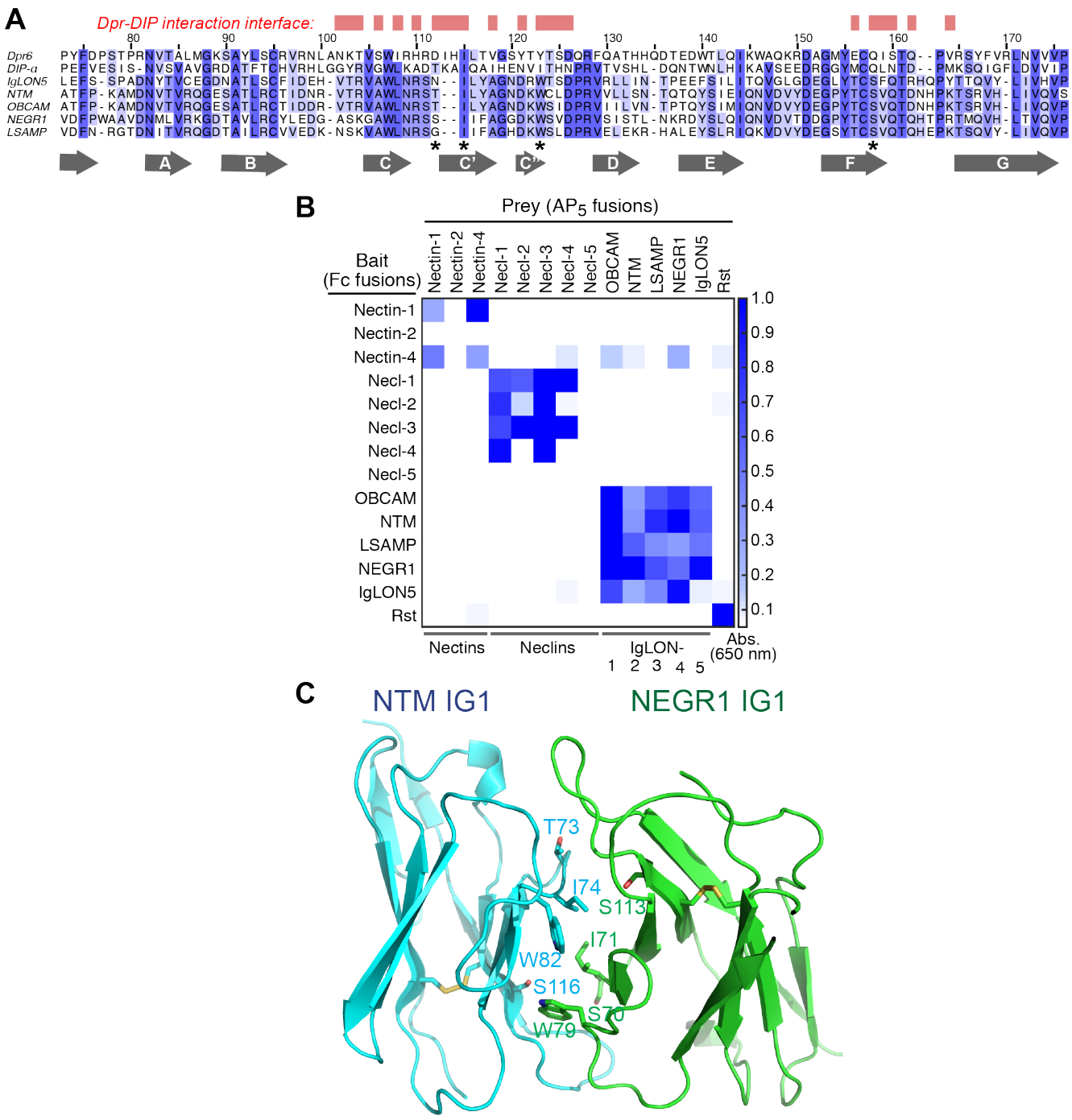

Fig. S4. Homology modeling of the mouse NTM-NEGR1 complex.
A. Sequence alignment and secondary structural elements of Dpr6, DIP-α, and the five mouse IgLONs. Amino acids mutated in Fig. 2D are labeled with and asterisk. The sequence alignment (also in Fig.2B) visually demonstrates that IgLONs are co-orthologous to Dprs and DIPs, and are not only DIP orthologs.
B. Binding experiments for mouse IgLONs, Nectins and Nectin-like proteins using ECIA. The homodimerization of Rst, the fly ortholog of mammalian Kirrels, serves as a positive control.
C. Homology model of the NTM IG1-NEGR1 IG1 heterodimer based on the structure of the Dpr6-DIP-α complex. Side chains of amino acids mutated in Fig. 2C are shown as sticks.

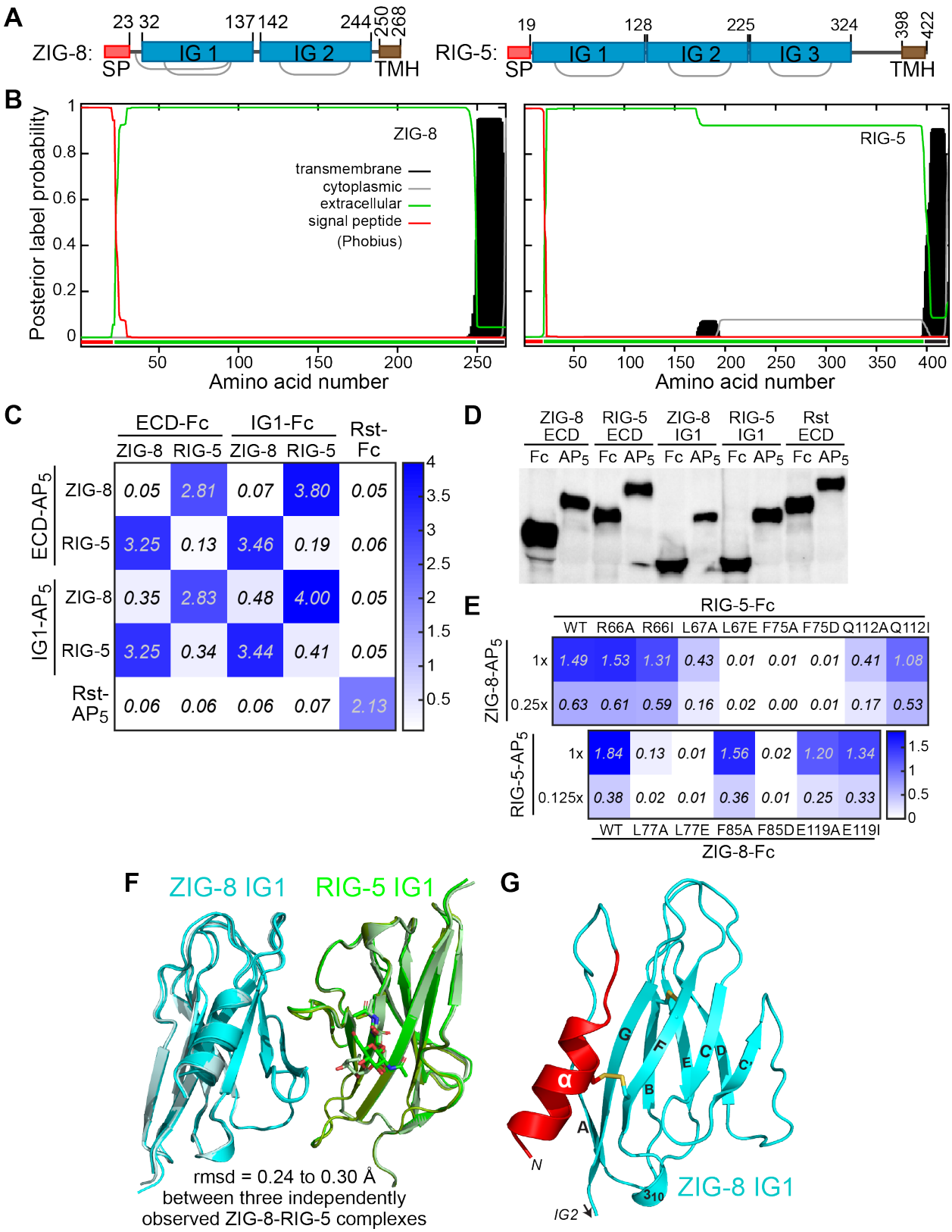

Fig. S5. A. The domain structure of ZIG-8 and RIG-5. The gray arches represent predicted disulfides bonds. SP: Signal Peptide. TMH: Transmembrane helix. IG: Immunoglobulin domain.

B. Signal peptide and transmembrane helix predictions by the *Phobius* server (26).
C. Binding experiments for ZIG-8 and RIG-5, performed using ECIA. The heterophilic interaction between ZIG-8 and RIG-5 is observed between both ectodomains (C) and IG1 domains only (D). RIG-5 and ZIG-8 form weaker IG1-IG1 homodimers, similar to DIPs in *Drosophila*. Fc (bait) and AP_5 (prey) concentrations were normalized by dilutions.
D. Western blots of ZIG-8 and RIG-5 ECIA constructs used in Fig. S5C. Both bait and prey constructs were detected by an iFluor 488-coupled anti-His tag antibody (Genscript, A01800). The bands were quantified for normalization of protein amounts used in the ECIA experiment.
E. Mutations at the observed ZIG-8 and RIG-5 interface affect heterophilic binding. This panel includes a more extensive set of mutations than, and including, those in Fig. 3C. To effectively compare wild-type to mutants, protein concentrations within each mutant bait series (rows) were normalized. Each prey was tested at two concentrations (1x and 0.25 or 0.125x) to ensure that binding affinities are compared at non-saturating concentrations.
F. The structure of the ZIG-8 IG1-RIG-5 IG1 complex observed in crystal form #1 (tetragonal form) and the two NCS copies in crystal form #2 (monoclinic). The three structure models are within 0.3 Å rmsd of each other. The sticks on RIG-5 represent N-linked glycosylation.
G. The N-terminal α-helical addition to the IG1 domain in ZIG-8 (drawn in red).

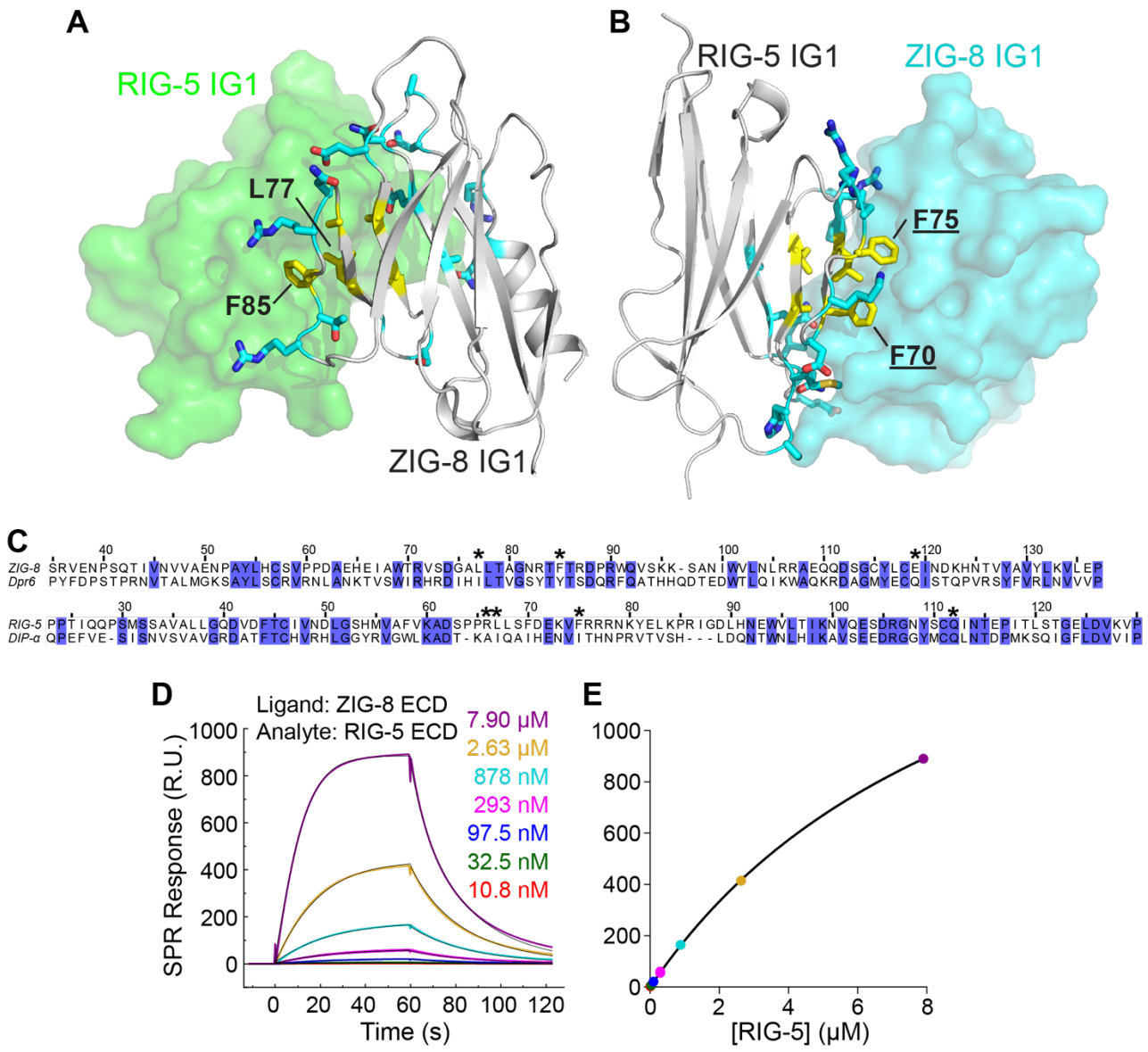

Fig. S6. The heterodimeric ZIG-8-RIG-5 complex.
A,B. ZIG-8 and RIG-5 interface amino acids L77, F85 (ZIG-8) and F70, F75 (RIG-5) are at the hydrophobic core (yellow residues) of the ZIG-8-RIG-5 interface, and fill in surface cavities in their corresponding binding partners.
C. Alignment of the entire IG domains for ZIG-8 and RIG-5 with Dpr6 and DIP-α. The amino acids mutated in Fig. 3C are labeled with an asterisk.
D. SPR data for the binding of RIG-5 ECD on a ZIG-8 ECD-coupled SPR chip. Thin black lines represent kinetic model fits to the binding data, collected at concentrations ranging from 10.8 nM to 7.9 μM. The calculated parameters are, K_D = 7.9 μM, k_{on} = 7.59×10^4 ± 2.1×10^3 M^{-1}s^{-1}, k_{off} = 0.60 ± 0.02 s^{-1}, with a mass transfer constant, t_c = 2.9×10^6 RU M^{-1} s^{-1}, indicating that the mass transport limitation is dominant, and the kinetic parameters might be suspect.
E. Binding isotherm for equilibrium fits to the SPR data in (D). The calculated K_D is 10.3 ± 0.3 μM and maximal response is 2047 RU.

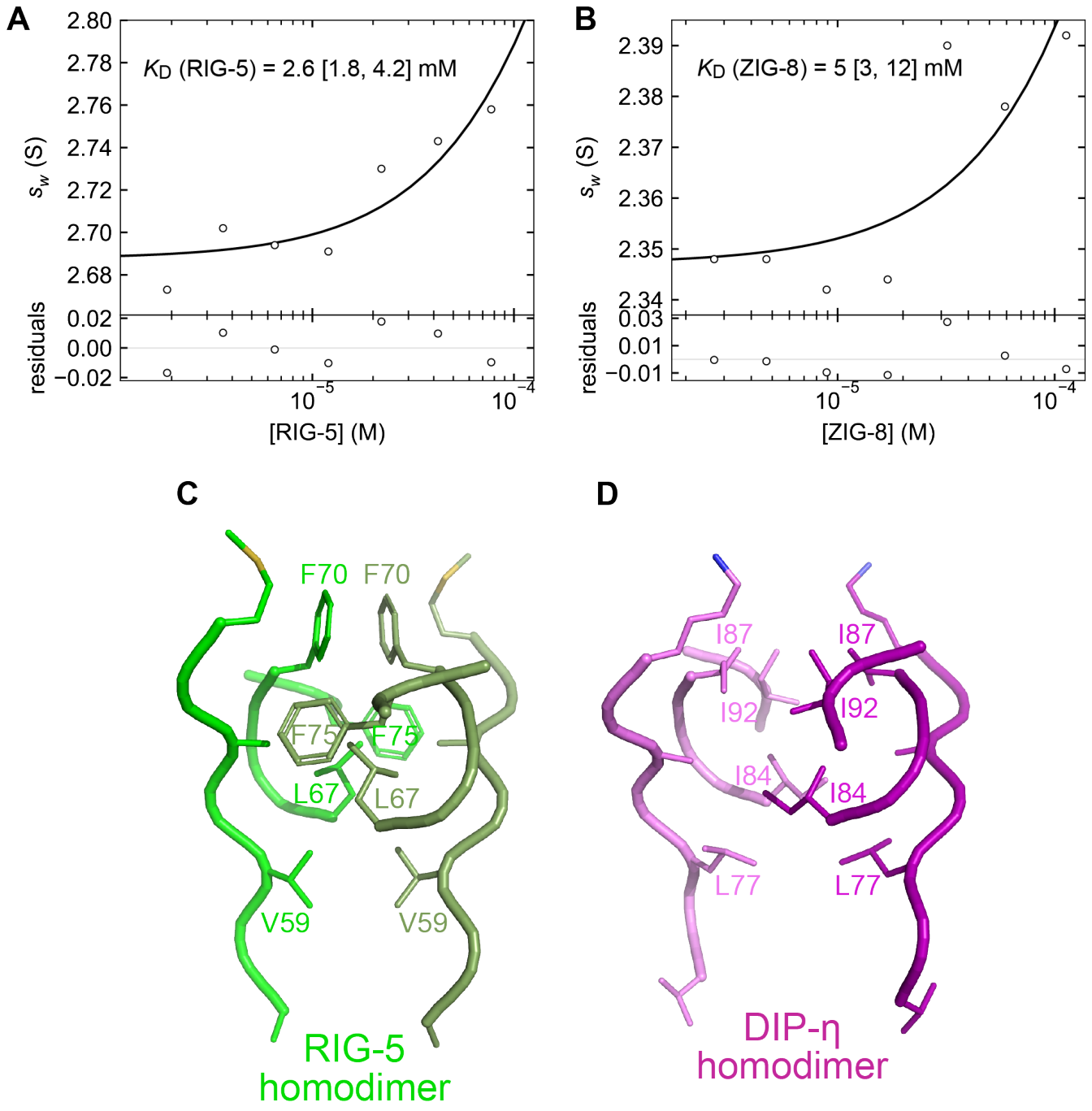

Fig. S7. The homodimeric ZIG-8 and RIG-5 complexes.
A, B. s_w data for RIG-5 ectodomain (A) and ZIG-8 ectodomain (B) are shown as circles, obtained by integration of the $c(s)$ distributions. The line is the fit to the data assuming a monomer-dimer model. The values between square brackets correspond to 68.3% confidence intervals.
C, D. A view of the hydrophobic core of RIG-5 (C) and DIP-η (D) homodimers. Despite zero sequence identity at the core, the two structures take similar shapes and adapt nearly identical interaction geometries.

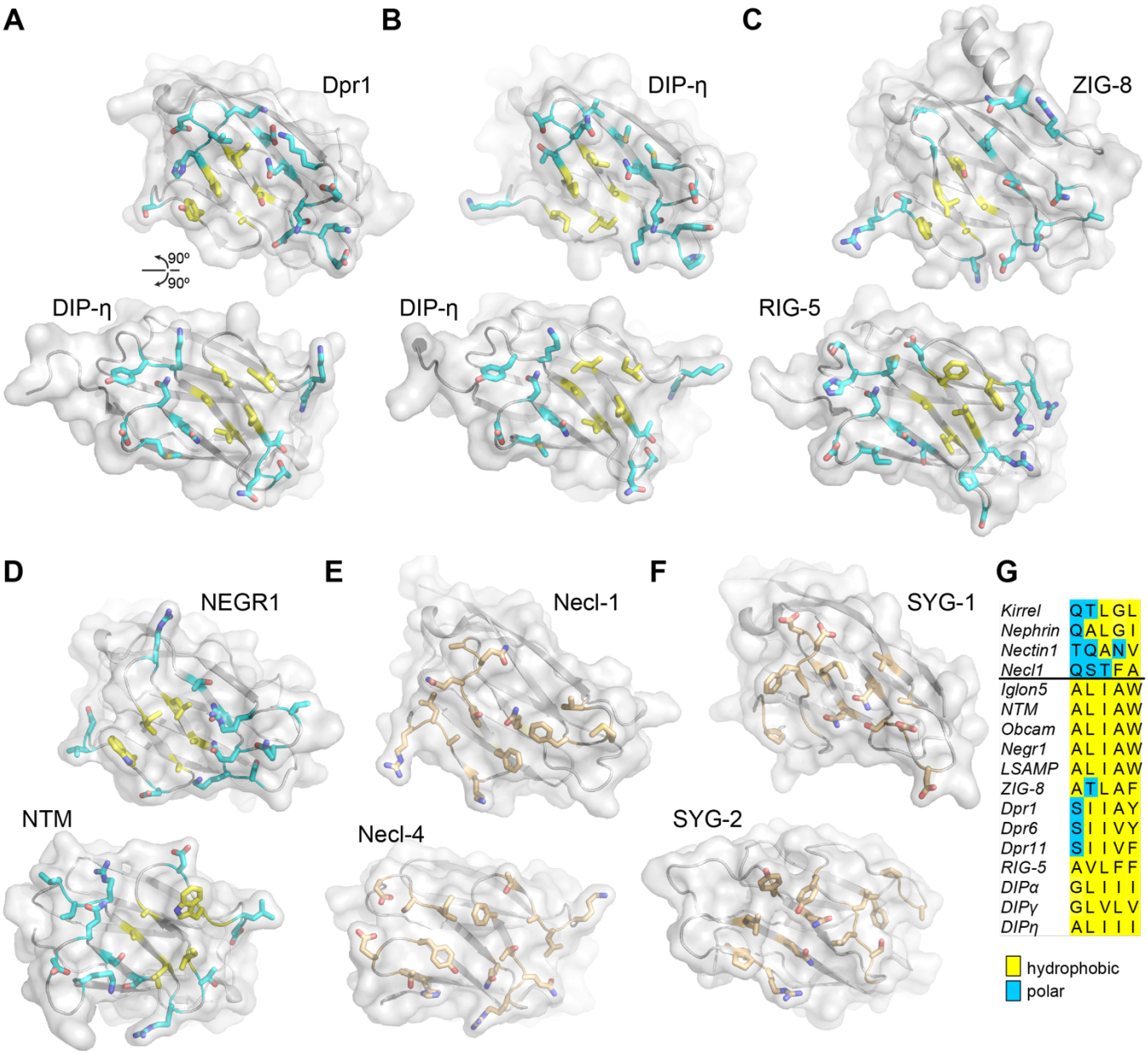

Fig. S8. Comparison of interfaces of IG1-IG1 complexes from Wirins and related families. **A-F.** Amino acids at the interface are shown with side chains as sticks. In Wirins, yellow-colored amino acids are at the hydrophobic core, as previously defined in Cheng *et al*. (14); cyan represents periphery. Since there is no clear hydrophobic core at the interface for the Necl and SYG complexes, and all of their interface amino acids were colored light orange. The NEGR-1-NTM structure is a homology model (Fig. S4B), while all others are crystallographically determined. The structures were aligned so that the subunits depicted at the bottom are superimposed on to each other. The PDB IDs of the structures shown are 6NRW (A), 6NRX (B), 6ONB (C), 5ZO2 (E), and 4OFY (F).
G. Amino acid positions at the "hydrophobic core" is compared.

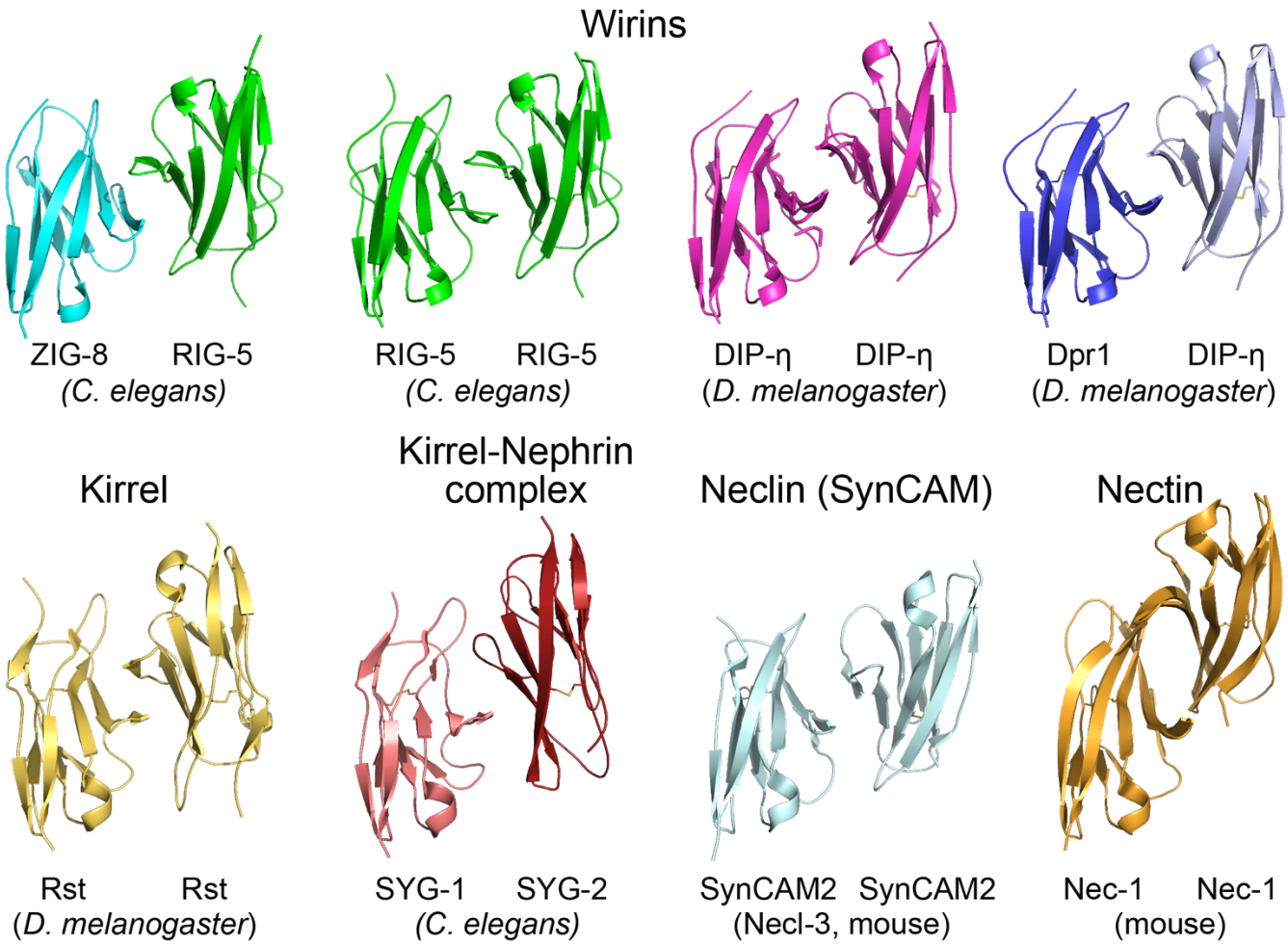

Fig. S9. Wirins and the four related IgSF protein families.
Comparison of IG1-IG1 complexes from Wirins and related families. Family names are noted above the structures. The structures were aligned so that the subunits depicted on the left are superimposed on to each other. The PDB IDs of the structures shown include 6ONB, 6ON6, 6NRX, 6NRW, 4OF8, 4OFY, 3M45, and 5B21.

Supplementary References

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410.
2. Shumway M, Cochrane G, Sugawara H (2010) Archiving next generation sequencing data. *Nucleic Acids Res* 38(Database issue):D870-871.
3. Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18(5):821–829.
4. Schulz MH, Zerbino DR, Vingron M, Birney E (2012) Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28(8):1086–1092.
5. Haas BJ, et al. (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 8(8):1494–1512.
6. Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113.
7. Price MN, Dehal PS, Arkin AP (2010) FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5(3):e9490.
8. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.
9. Guindon S, et al. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59(3):307–321.
10. Özkan E, et al. (2013) An Extracellular Interactome of Immunoglobulin and LRR Proteins Reveals Receptor-Ligand Networks. *Cell* 154(1):228–239.
11. Eswar N, et al. (2006) Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics* Chapter 5:Unit 5.6.
12. Krivov GG, Shapovalov MV, Dunbrack RL (2009) Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* 77(4):778–795.
13. Schrödinger, LLC (2017) *The PyMOL Molecular Graphics System, Version 2.0* Available at: https://pymol.org/2/.
14. Cheng S, et al. (2019) Molecular basis of synaptic specificity by immunoglobulin superfamily receptors in Drosophila. *Elife* 8:e41028.
15. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658–674.
16. Afonine PV, et al. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4):352–367.
17. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.

18. Chen VB, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12–21.
19. Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
20. Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys J* 78(3):1606–1619.
21. Zhao H, Brautigam CA, Ghirlando R, Schuck P (2013) Overview of current methods in sedimentation velocity and sedimentation equilibrium analytical ultracentrifugation. *Curr Protoc Protein Sci* Chapter 20:Unit20.12.
22. Laue TM, Shah BD, Ridgeway RM, Pelletier SL (1992) Computer-aided interpretation of analytical sedimentation data for proteins. *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, eds Harding SE, Rowe AJ, Horton JC (The Royal Society of Chemistry, Cambridge, UK), pp 90–125.
23. Brautigam CA (2015) Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. *Meth Enzymol* 562:109–133.
24. Brautigam CA (2011) Using Lamm-Equation modeling of sedimentation velocity data to determine the kinetic and thermodynamic properties of macromolecular interactions. *Methods* 54(1):4–15.
25. Harris TW, et al. (2014) WormBase 2014: new views of curated biology. *Nucleic Acids Res* 42(Database issue):D789-793.
26. Käll L, Krogh A, Sonnhammer ELL (2007) Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Res* 35(Web Server issue):W429-432.
27. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8(10):785–786.