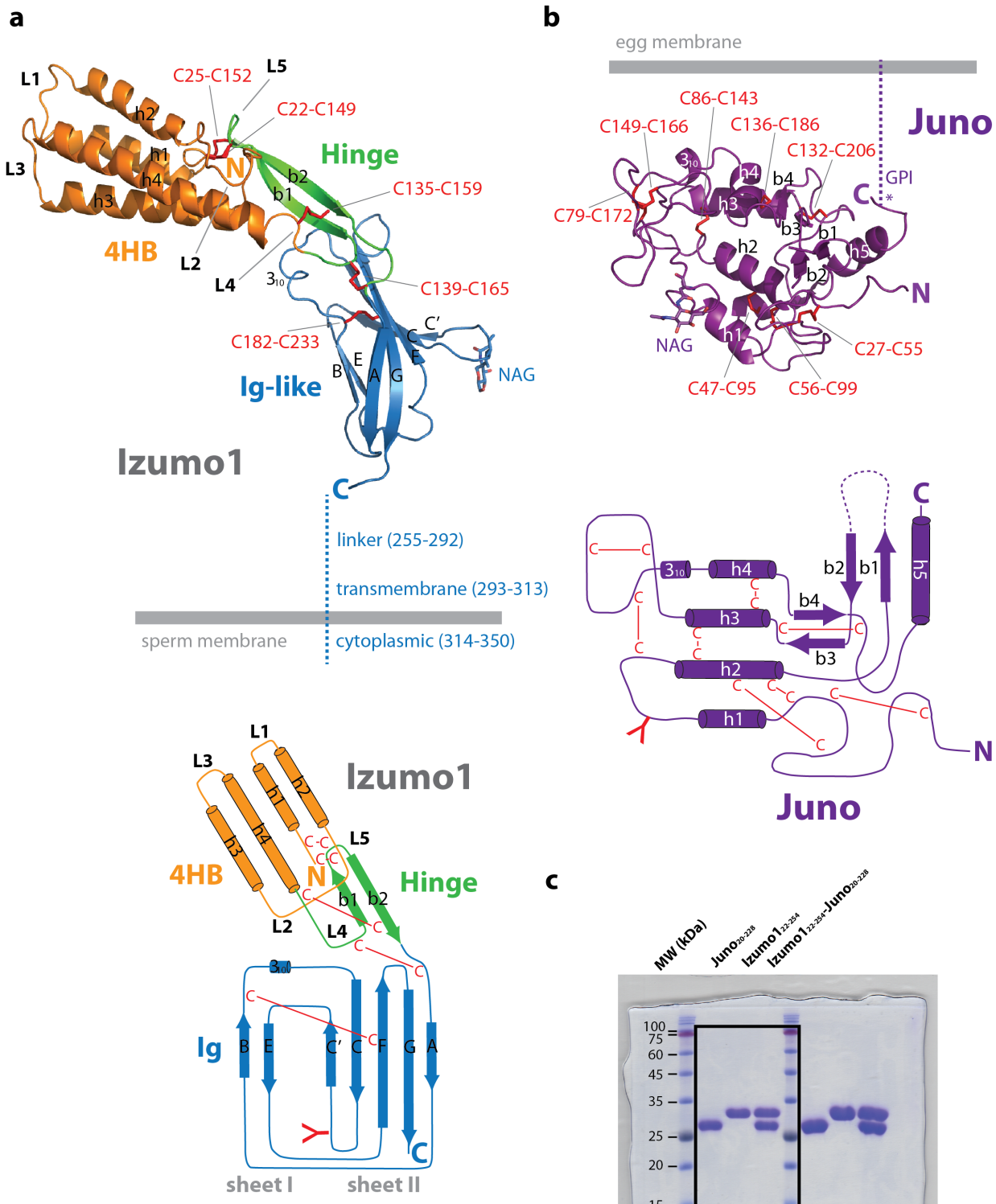
SUPPLEMENTARY INFORMATION

Molecular architecture of the human sperm Izumo1 and egg Juno fertilization complex

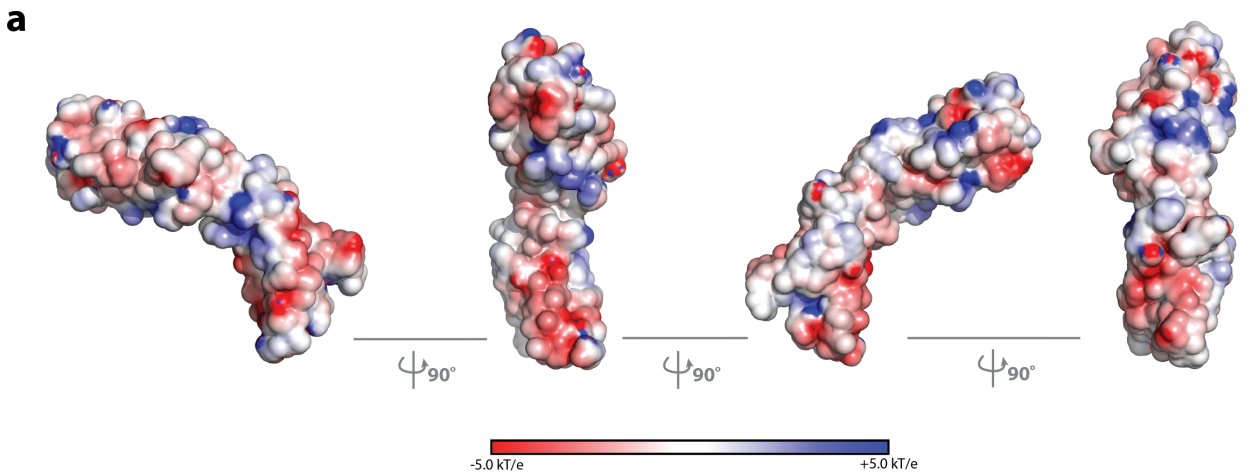
Halil Aydin[1], Azmiri Sultana[1], Sheng Li[2], Annoj Thavalingam[1] & Jeffrey E. Lee[1]

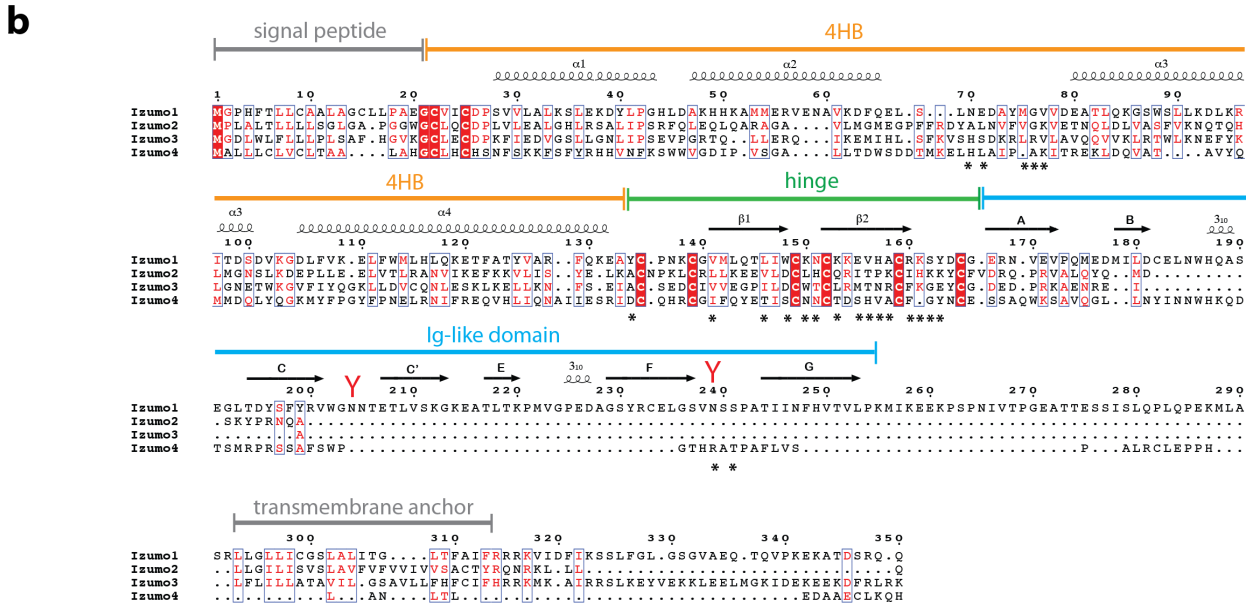
[1] Department of Laboratory Medicine and Pathobiology, Faculty of Medicine, University of Toronto, Toronto, ON M5S 1A8,
[2] Department of Medicine, University of California, San Diego, La Jolla, CA 92093, USA

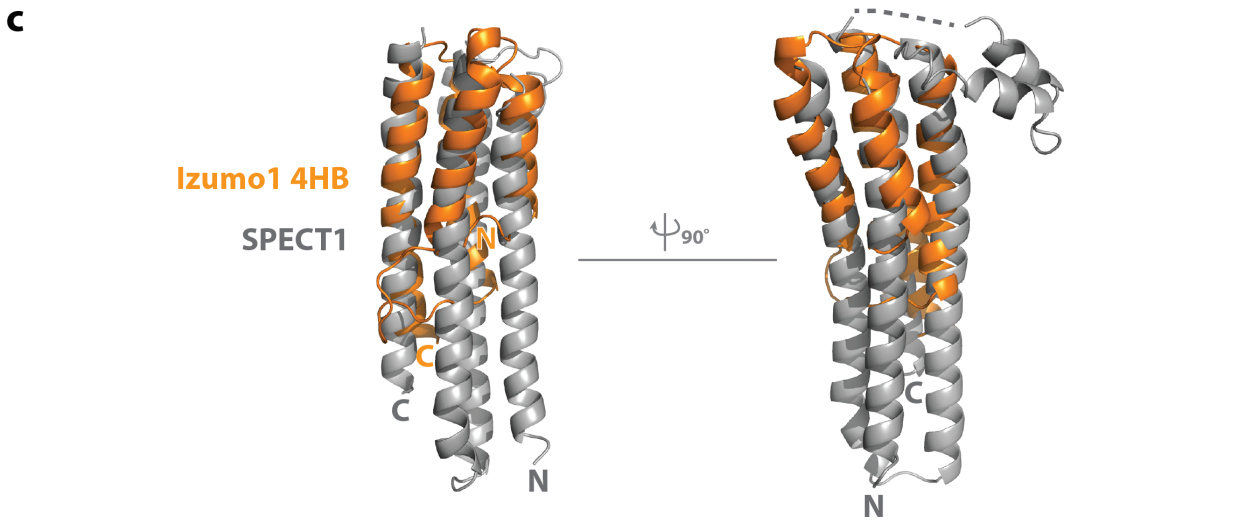


Extended Data Figure 3b

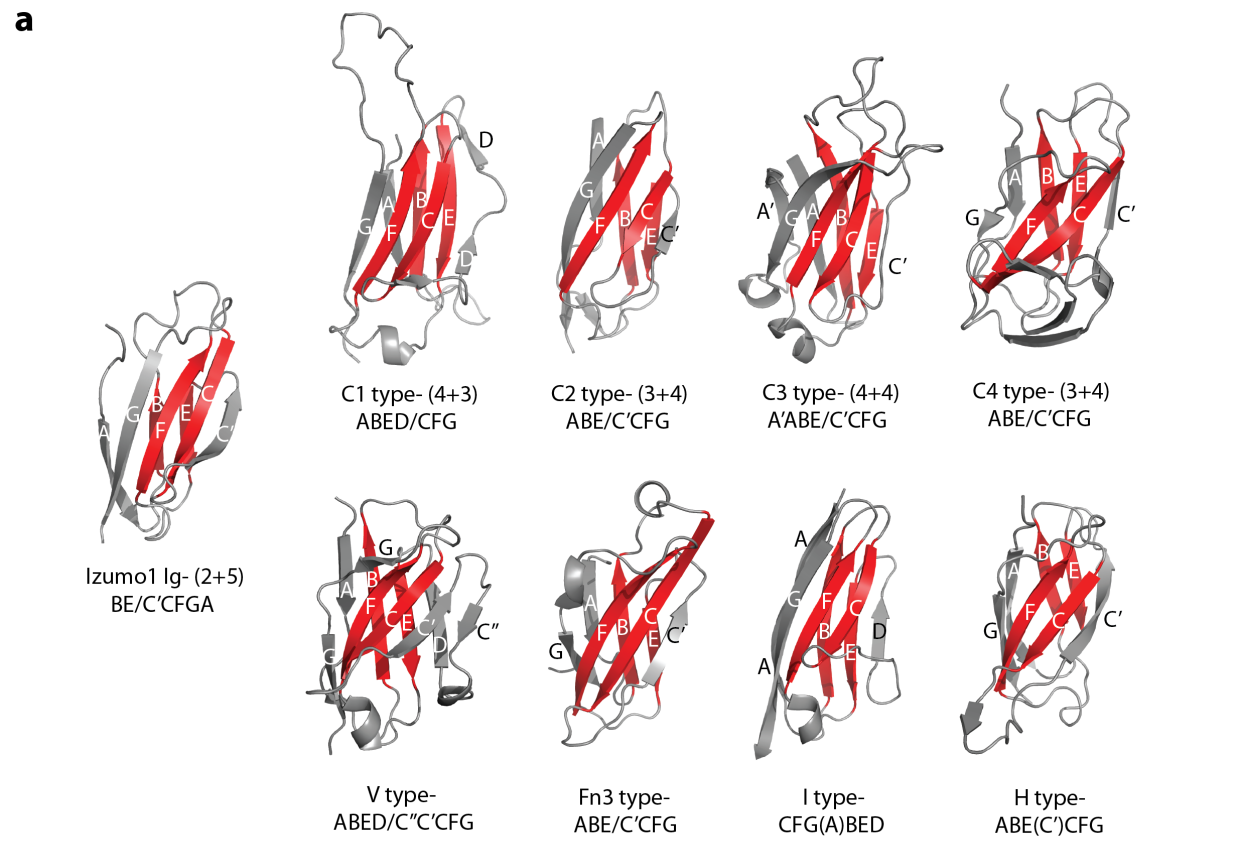
Supplementary Figure 1 | Overall structural architecture of Izumo1 and Juno. Ribbon and topology schematic of **(a)** Izumo1$_{22-254}$ and **(b)** Juno$_{20-228}$. β-strands and α-helices are depicted in the topology diagram as arrows and cylinders, respectively. Disulphide bonds are shown as red sticks. The four-helix bundle (4HB) motif is a very common natural structural fold that occurs in proteins of diverse functions[1]. Commonly, helices in 4HB are tilted +25º or -35º to maximize interhelical contacts and hydrophobic residues pack in a knobs-in-holes fashion to stabilize the core[1]. Atypically, the Izumo1 4HB adopts a rare parallel/antiparallel arrangement of non-tilted helices with no knobs-in-holes packing, suggesting that the 4HB region may be more conformationally labile. **(c)** Source data: Uncropped scanned Coomassie-stained SDS-PAGE gel from Extended Data Figure 3b.

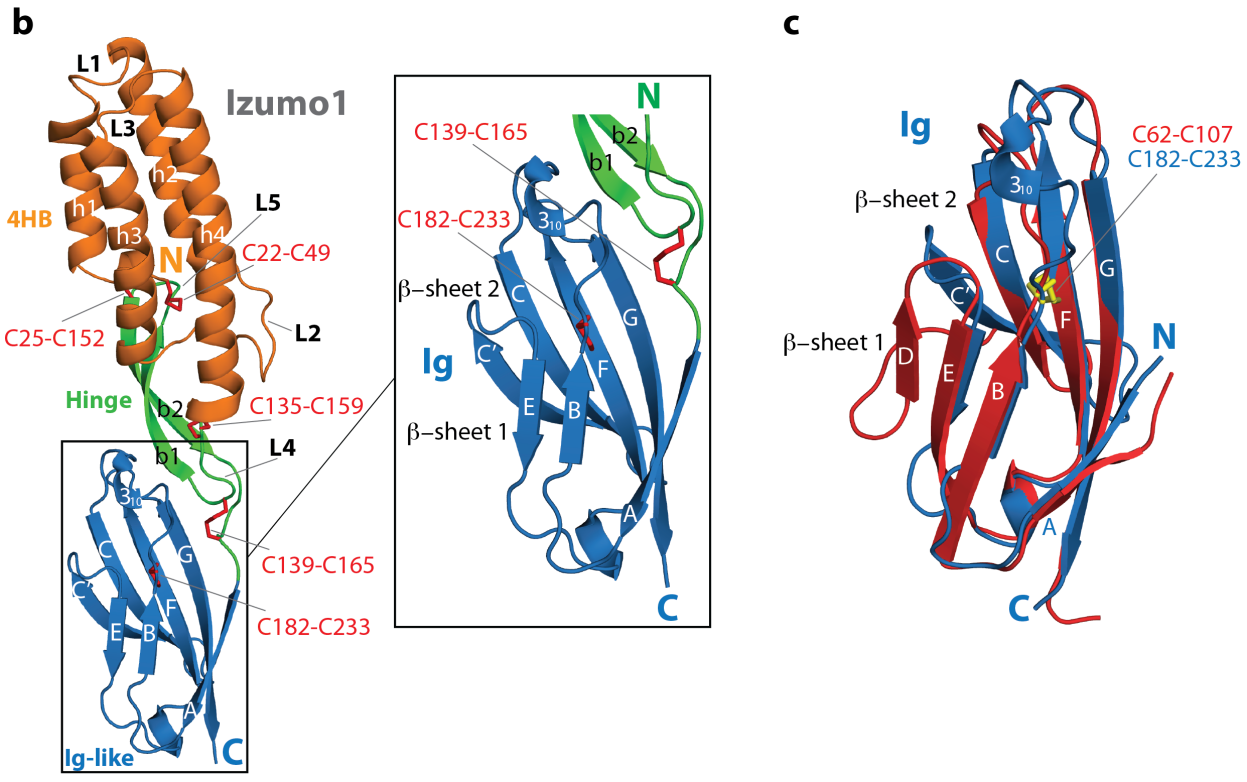

a
90°
90°
90°
-5.0 kT/e
+5.0 kT/e


b
signal peptide
4HB
4HB
hinge
Ig-like domain
transmembrane anchor
Izumo1
Izumo2
Izumo3
Izumo4

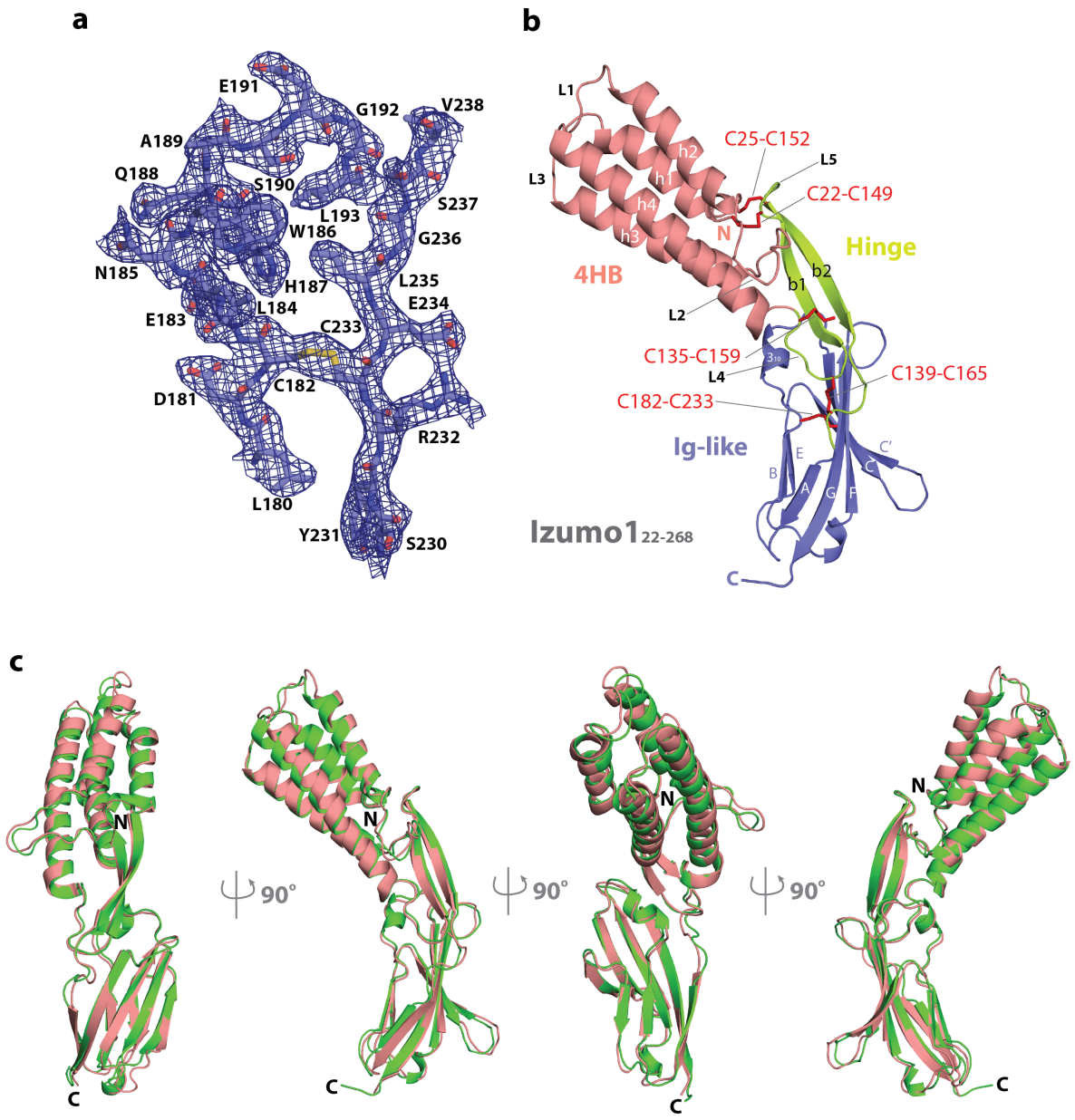

c
Izumo1 4HB
SPECT1
90°
N
C
C
N
C
N

Supplementary Figure 2 | Structural properties of Izumo1. (a) Electrostatic potential from -5 to +5 kT/e mapped onto the molecular surface of $Izumo1_{22-254}$ is shown in four orientations, 90° apart. **(b)** Multiple sequence alignment of human Izumo1, Izumo2, Izumo3 and Izumo4. The eight cysteines observed in Izumo1 are highly conserved in the human Izumo family. However, there is very poor overall sequence conservation in the family (<3% sequence identity). No residues involved in binding Juno are conserved (highlighted with asterisks). Moreover, Izumo2, Izumo3 and Izumo4 do not appear to contain an intact Izumo1 Ig-like domain. Izumo4 also seems to lack a transmembrane anchor. **(c)** A 3D structural database search using DaliLite[2] (v.3) revealed similarity between Izumo1 4HB and SPECT1[3], a *Plasmodium* protein involved in host cell traversal. Overall, 97 positions between SPECT1 (PDB ID: 4U5A) and Izumo1 4HB (residues 22-134) superimpose with 3.4-Å rmsd. Interestingly, SPECT1 is required for pore formation by highly motile sporozoites prior to invading hepatocytes; its mechanism is however poorly defined[4,5].

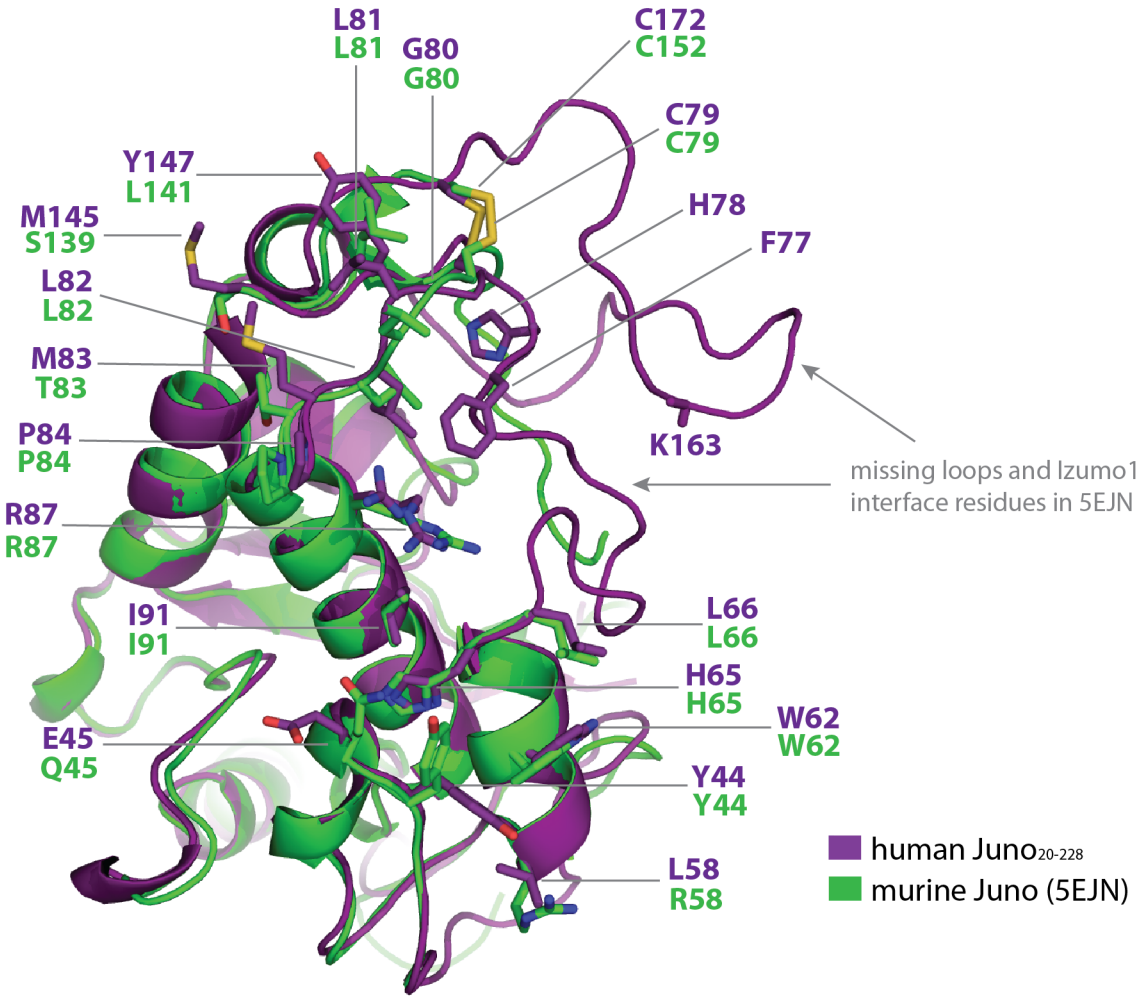

a
Izumo1 Ig- (2+5)
BE/C'CFGA
C1 type- (4+3)
ABED/CFG
C2 type- (3+4)
ABE/C'CFG
C3 type- (4+4)
A'ABE/C'CFG
C4 type- (3+4)
ABE/C'CFG
V type-
ABED/C"C'CFG
Fn3 type-
ABE/C'CFG
I type-
CFG(A)BED
H type-
ABE(C')CFG


b
L1
Izumo1
L3
h2
4HB
h1
h3
h4
L5
N
C22-C49
C25-C152
L2
Hinge
b2
C135-C159
b1
L4
C139-C165
C182-C233
Ig-like
C
N
C139-C165
b2
b1
C182-C233
β–sheet 2
Ig
β–sheet 1
C
c
Ig
C62-C107
C182-C233
β–sheet 2
β–sheet 1
N
C

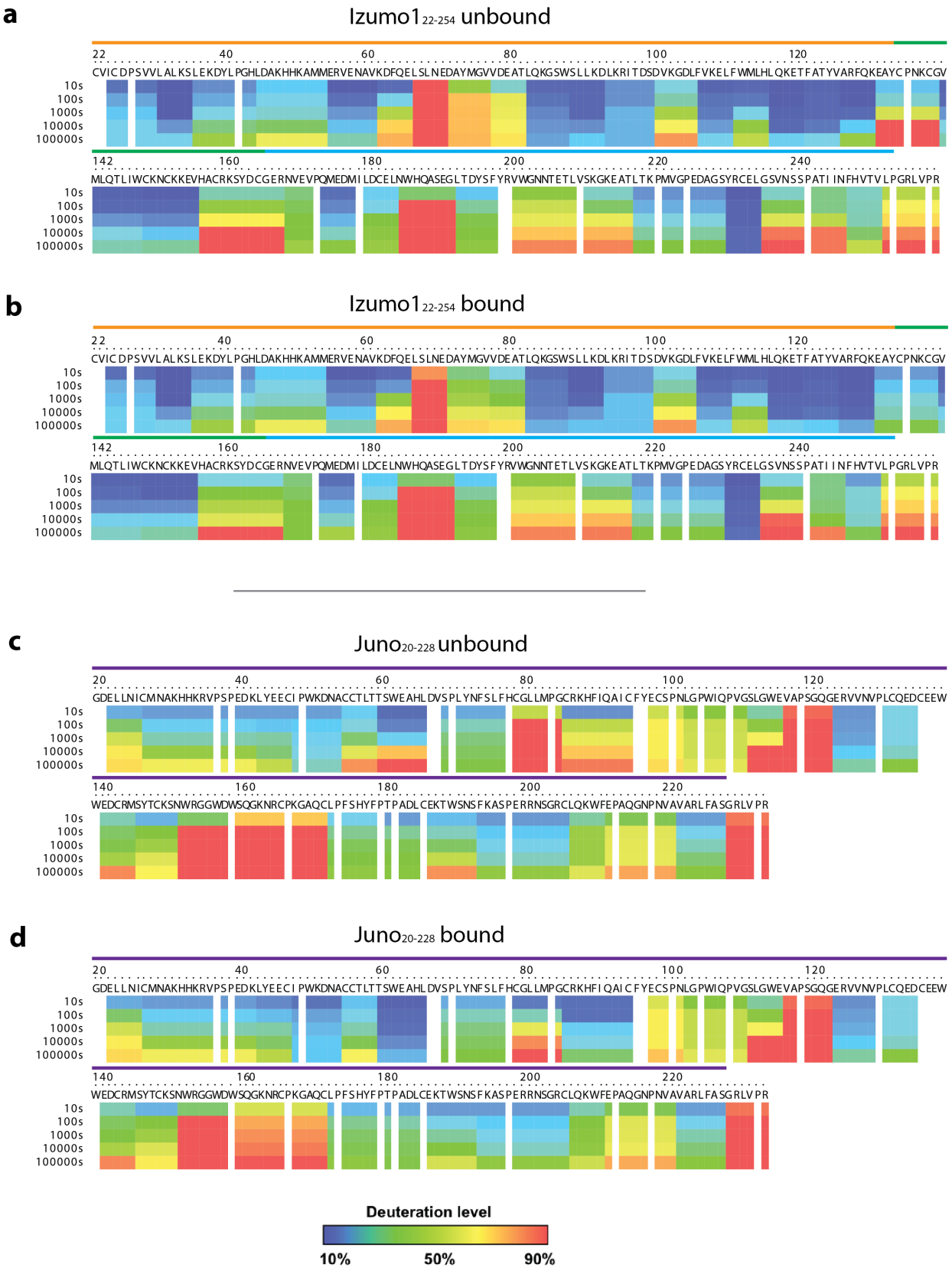
Supplementary Figure 3 | Comparison of the Izumo1 Ig-like domain and various Ig-superfamily (IgSF) subtypes. (a) The Izumo1 Ig-like domain adopts a 2+5 stranded topology; with sheet II (front sheet in this view) includes five strands (*a, g, f, c, c'*) and sheet I contains two strands (*b, e*). The 2+2 stranded common structural core (strands coloured in red and labeled *b, c, e, f*) is conserved in all IgSF domains. Due to a disulphide bond preceding the Izumo1 Ig-like domain, strand A is forced to hydrogen bond with strand G in sheet I. The Izumo1 Ig-like domain is the only known member to adopt a 2+5 arrangement, thus suggesting a new IgSF subtype. Selected examples illustrating the various IgSF subtypes are as follows: C1-type, T-cell receptor constant (PDB ID: 1BEC); C2-type, CD4 constant (PDB ID: 3CD4); C3-type, Cu/Zn superoxide dismutase (PDB ID: 2SOD); C4-type, neocarzinostatin (PDB ID: 1NCO); V-type, Ig-variable domain heavy chain (PDB ID: 7FAB); Fn3-type, neuroglian (PDB ID: 1CFB); I-type, telokine (PDB ID: 1TLK); and H-type, galactose oxidase (PDB ID: 1GOG). **(b)** The Izumo1 Ig-like domain adopts a novel 2+5 β-strand arrangement. A disulphide bond (shown as red sticks) between C139 and C165 stabilizes the interaction of the A-strand with β-sheet II. The Izumo1 Ig-like domain is coloured in blue and hinge region is shown in green. **(c)** Superimposition of Izumo1 Ig-like domain (shown in blue) and the C2-type Ig-like domain from murine FGFR2 (shown in red; PDB ID: 4HWU). A comparative analysis of the Izumo1 Ig-like domain to structures deposited in the PDB using the secondary structure matching (SSM) server revealed that its closest structural neighbour is the murine fibroblast growth factor receptor-2 (FGFR2), which adopts a C2-type Ig-like fold (rmsd 1.3 Å for 69 Cα atoms with SSM Q-score=0.57). The Izumo1 Ig-like domain adopts a C2-type Ig-like fold, similar to murine FGFR2 with the exception of the 2+5 β-strand arrangement.



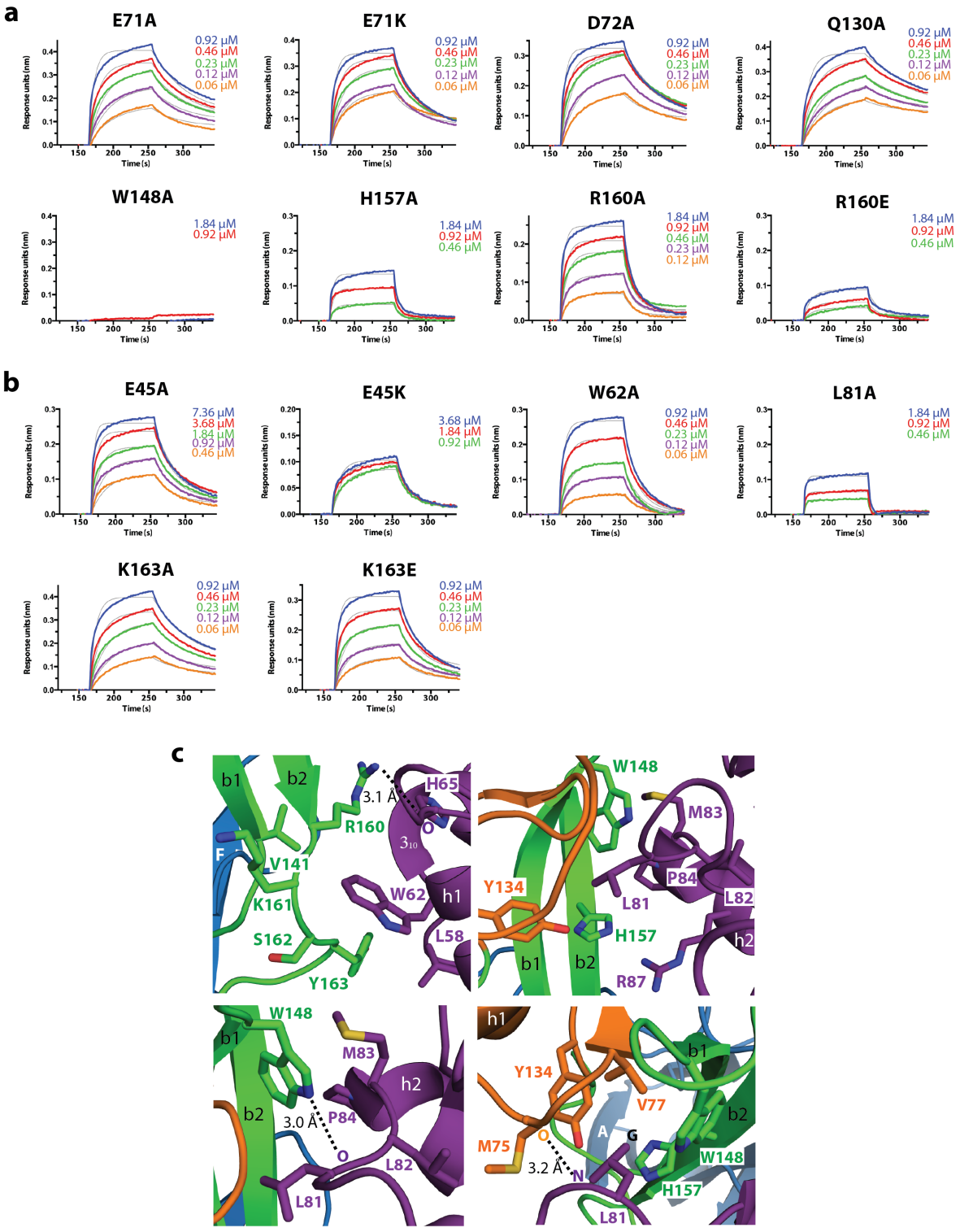
Supplementary Figure 4 | Crystal structure of Izumo1$_{22\text{-}268}$. (a) A representative 2.9-Å resolution σ_A-weighted composite omit $2|F_o|-|F_c|$ electron density map of Izumo1$_{22\text{-}268}$ contoured at 1σ and superimposed with its final refined model. **(b)** Ribbon diagram of Izumo1$_{22\text{-}268}$. The 4HB region, hinge region and Ig-like domain are shown in salmon, lemon and purple, respectively. Conserved disulphide linkages are depicted as red sticks. **(c)** Superimposition of the Izumo1$_{22\text{-}254}$ (shown in salmon) and Izumo1$_{22\text{-}268}$ (shown in green) crystal structures. Overall, the structures align with a rmsd of 1.0 Å over all Cα atoms. Residues 257-268 in Izumo1$_{22\text{-}268}$ are disordered suggesting that the linker region between the Ig-like domain and the transmembrane domain is highly flexible.



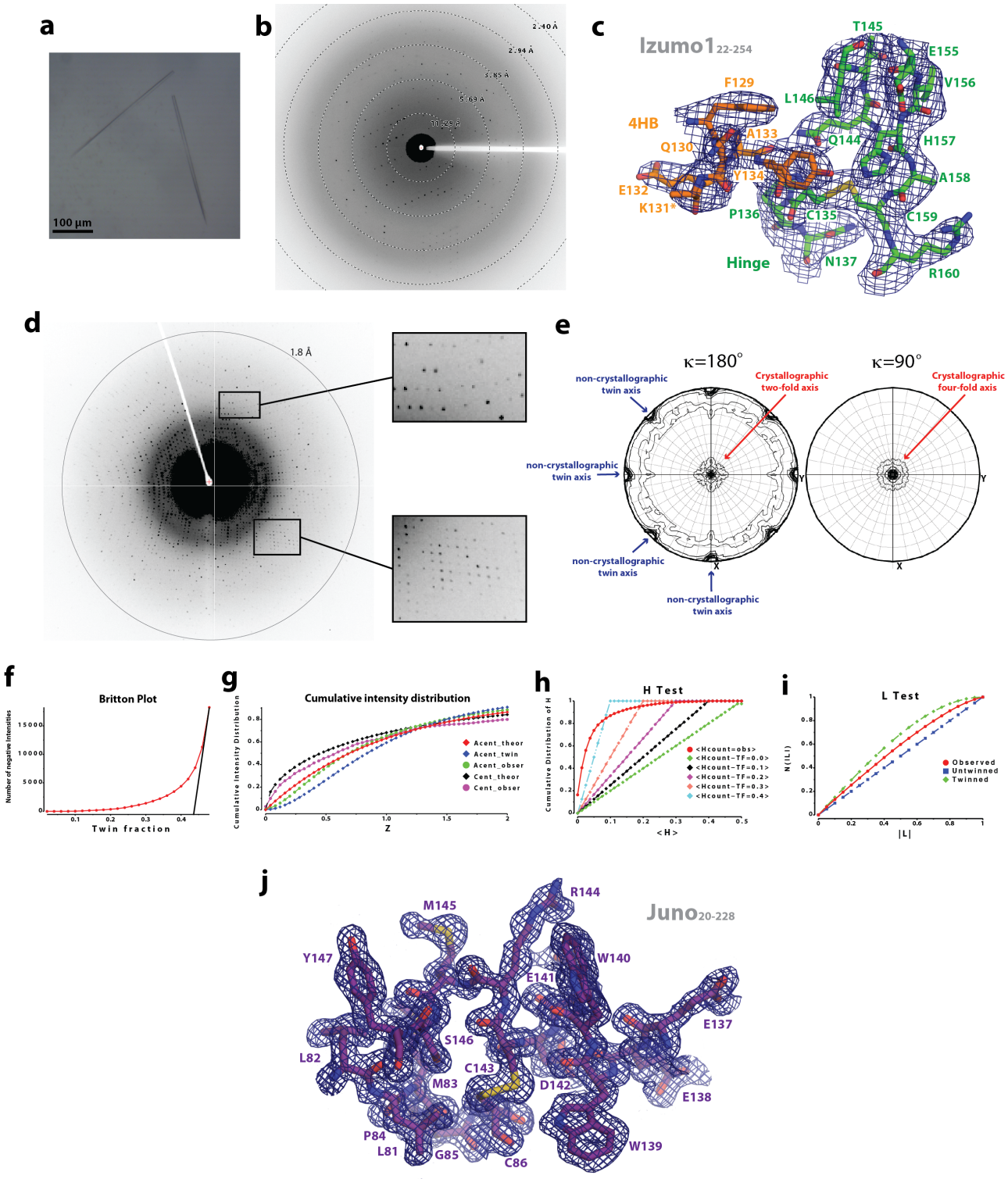
Supplementary Figure 5 | Structural comparison of human and murine Juno. During revisions for our manuscript, a murine Juno crystal structure (PDB code: 5EJN) was determined[6]. The reported murine Juno model consists of only a partial structure, as a number of residues (residues 34-40, 67-78, 112-117 and 152-165 in chain A) in the loop regions are missing due to flexibility and disorder in the crystal lattice. Superimposition of human Juno (PDB code: 5F4Q) and murine Juno (PDB code: 5EJN) reveals no major structural differences in the core regions (rmsd=1.44 Å over 201 residues). However, two missing loops in murine Juno (residues 67-78 and 152-165) play an important role in Izumo1 binding.


a
Izumo1 22-254 unbound
22 40 60 80 100 120
CVICDPSVVLALKSLEKDYLPGHLDAKHHKAMMERVENAVKDFQELSLNEDAYMGVVDEATLQKGSWSLLKDLKRITDSDVKGDLFVKELFWMLHLQKETFATYVARFQKEAYCPNKCGV
142 160 180 200 220 240
MLQTLIWCKNCKKEVHACRKSYDCGERNVEVPQMEDMILDCELNWHQASEGLTDYSFYRVWGNNTETLVSKGKEATLTKPMVGPEDAGSYRCELGSVNSSPATIINFHVTVLPGRLVPR
b
Izumo1 22-254 bound
22 40 60 80 100 120
CVICDPSVVLALKSLEKDYLPGHLDAKHHKAMMERVENAVKDFQELSLNEDAYMGVVDEATLQKGSWSLLKDLKRITDSDVKGDLFVKELFWMLHLQKETFATYVARFQKEAYCPNKCGV
142 160 180 200 220 240
MLQTLIWCKNCKKEVHACRKSYDCGERNVEVPQMEDMILDCELNWHQASEGLTDYSFYRVWGNNTETLVSKGKEATLTKPMVGPEDAGSYRCELGSVNSSPATIINFHVTVLPGRLVPR
c
Juno 20-228 unbound
20 40 60 80 100 120
GDELLNICMNAKHHKRVPSPEDKLYEECIPWKDNACCTLTTSWEAHLDVSPLYNFSLFHCGLMPGCRKHFIQAICFYECSPNLGPWIQPVGSLGWEVAPSGQGERVVNVPLCQEDCEEW
140 160 180 200 220
WEDCRMSYTCKSNWRGGWDWSQGKNRCPKGAQCLPFSHYFPTPADLCEKTWSNSFKASPERRNSGRCLQKWFEPAQGNPNVAVARLFASGRLVPR
d
Juno 20-228 bound
20 40 60 80 100 120
GDELLNICMNAKHHKRVPSPEDKLYEECIPWKDNACCTLTTSWEAHLDVSPLYNFSLFHCGLMPGCRKHFIQAICFYECSPNLGPWIQPVGSLGWEVAPSGQGERVVNVPLCQEDCEEW
140 160 180 200 220
WEDCRMSYTCKSNWRGGWDWSQGKNRCPKGAQCLPFSHYFPTPADLCEKTWSNSFKASPERRNSGRCLQKWFEPAQGNPNVAVARLFASGRLVPR
10s
100s
1000s
10000s
100000s
Deuteration level
10%
50%
90%

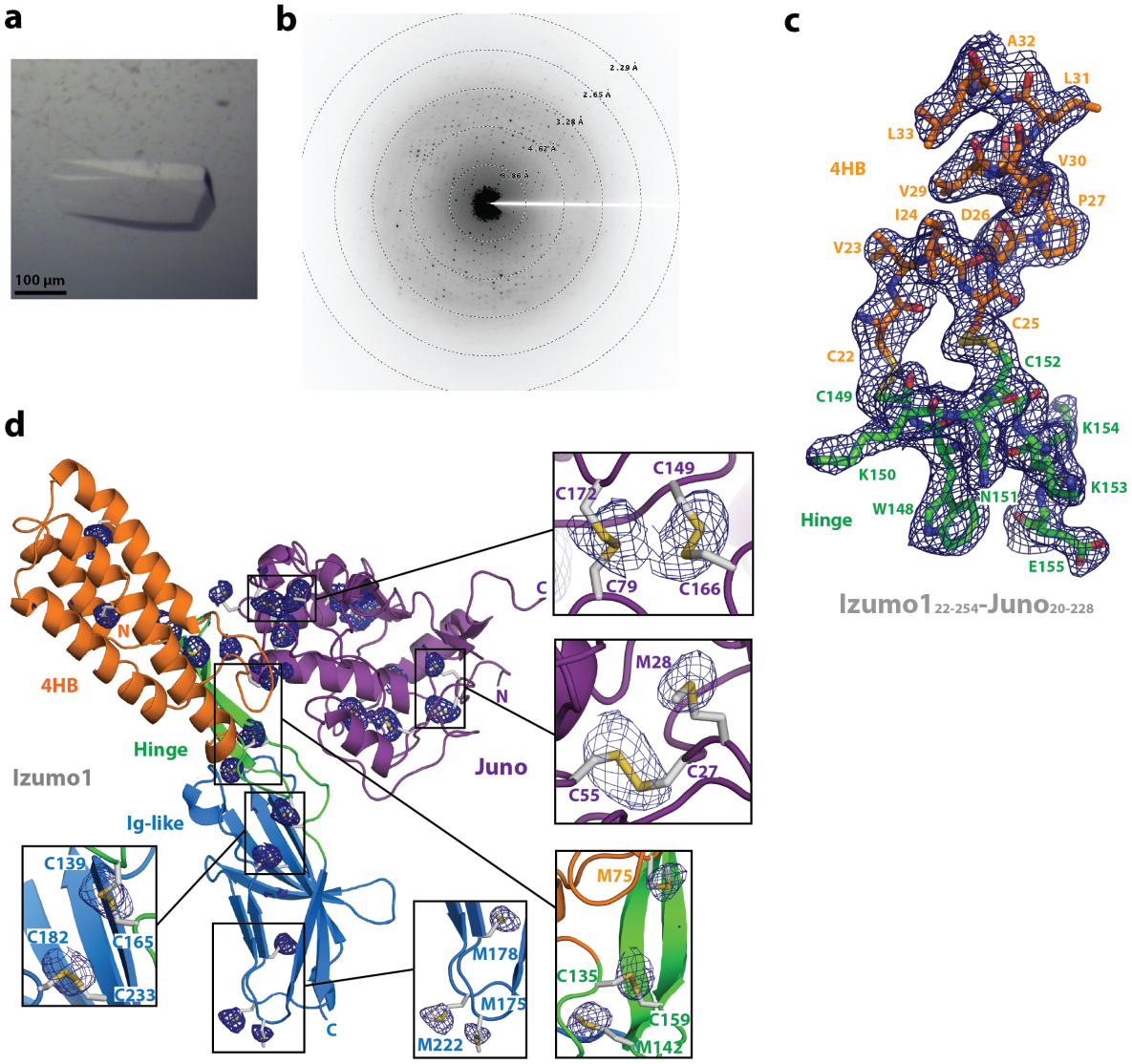
Supplementary Figure 6 | DXMS analysis. The DXMS exchange profiles of **(a)** unbound Izumo1$_{22\text{-}254}$, **(b)** Izumo1$_{22\text{-}254}$ bound to Juno, **(c)** unbound Juno$_{20\text{-}228}$, and **(d)** Juno$_{20\text{-}228}$ bound to Izumo1. Percentage deuteration exchange was determined over five different time points (10, 100, 1000, 10000 and 100000 seconds) for all observable Izumo1$_{22\text{-}254}$ and Juno$_{20\text{-}228}$ residues in the bound and unbound states. The lines above the residue numbers are colour coded with the regions observed in the crystal structures (4HB, hinge and Ig-like domains, and Juno).


a
E71A
0.92 μM
0.46 μM
0.23 μM
0.12 μM
0.06 μM
E71K
0.92 μM
0.46 μM
0.23 μM
0.12 μM
0.06 μM
D72A
0.92 μM
0.46 μM
0.23 μM
0.12 μM
0.06 μM
Q130A
0.92 μM
0.46 μM
0.23 μM
0.12 μM
0.06 μM
W148A
1.84 μM
0.92 μM
H157A
1.84 μM
0.92 μM
0.46 μM
R160A
1.84 μM
0.92 μM
0.46 μM
0.23 μM
0.12 μM
R160E
1.84 μM
0.92 μM
0.46 μM
b
E45A
7.36 μM
3.68 μM
1.84 μM
0.92 μM
0.46 μM
E45K
3.68 μM
1.84 μM
0.92 μM
W62A
0.92 μM
0.46 μM
0.23 μM
0.12 μM
0.06 μM
L81A
1.84 μM
0.92 μM
0.46 μM
K163A
0.92 μM
0.46 μM
0.23 μM
0.12 μM
0.06 μM
K163E
0.92 μM
0.46 μM
0.23 μM
0.12 μM
0.06 μM
Response units (nm)
Time (s)
c
b1
b2
H65
3.1 Å
R160
O
3₁₀
V141
W62
h1
K161
S162
L58
Y163
W148
M83
P84
L82
Y134
L81
H157
h2
b1
b2
R87
W148
b1
M83
h2
b2
3.0 Å
P84
O
L82
L81
h1
b1
Y134
V77
b2
M75
O
A
G
3.2 Å
N
W148
H157
L81

Supplementary Figure 7 | Biolayer interferometry binding affinity and kinetic analysis of Izumo1-Juno interface residues. (a) Biotin-labeled wild-type $Juno_{20-228}$ was coupled to streptavidin-coated biosensors and monitored for binding to various mutant $Izumo1_{22-254}$ over multiple concentrations. **(b)** Various biotin-labeled mutant $Juno_{20-228}$ was coupled to streptavidin-coated biosensors and monitored for binding to wild-type $Izumo1_{22-254}$ at multiple concentrations. The data were analyzed based on a 1:1 binding model using the BLItz Pro software, with the fitted curves shown as grey lines. **(c)** Specific conserved interactions at the Izumo1-Juno binding interface. Residues involved in Izumo1-Juno binding are shown in stick representation. Key binding interface residues $W62_{Juno}$, $L81_{Juno}$, $W148_{Izumo1}$ and $H157_{Izumo1}$ are conserved in most mammals. Dashed lines indicate hydrogen bonds.


a
100 µm
b
2.40 Å
2.94 Å
3.85 Å
5.69 Å
c
Izumo1 22-254
T145
E155
V156
F129
L146
4HB
A133
Q130
Q144
H157
Y134
A158
E132
K131*
P136
C135
C159
Hinge
N137
R160
d
1.8 Å
e
κ=180°
κ=90°
non-crystallographic twin axis
Crystallographic two-fold axis
Crystallographic four-fold axis
non-crystallographic twin axis
non-crystallographic twin axis
non-crystallographic twin axis
f
Britton Plot
Number of negative intensities
Twin fraction
g
Cumulative intensity distribution
Cumulative Intensity Distribution
Z
Acent_theor
Acent_twin
Acent_obser
Cent_theor
Cent_obser
h
H Test
Cumulative Distribution of H
<H>
<Hcount=obs>
<Hcount-TF=0.0>
<Hcount-TF=0.1>
<Hcount-TF=0.2>
<Hcount-TF=0.3>
<Hcount-TF=0.4>
i
L Test
N(|L|)
|L|
Observed
Untwinned
Twinned
j
Juno 20-228
M145
R144
Y147
W140
E141
E137
S146
L82
C143
M83
D142
E138
P84
L81
G85
C86
W139

Supplementary Figure 8 | Izumo1$_{22-254}$ and Juno$_{20-228}$ structure determination. (a) Crystals of the unbound Izumo1$_{22-254}$ grown in 0.07 M sodium acetate (pH 4.6), 5.6% (w/v) PEG 4000 and 30% (v/v) glycerol. **(b)** A representative X-ray diffraction image of Izumo1$_{22-254}$ illustrates the anisotropic diffraction of the needle-shaped crystals. **(c)** Representative 3.1-Å resolution σ_A-weighted composite omit $2|F_o|-|F_c|$ electron density map of Izumo1$_{22-254}$ contoured at 1σ and superimposed with its final refined model. **(d)** A representative example of an X-ray diffraction image of the merohedrally twinned Juno$_{20-228}$. The inset windows show the medium-resolution and high-resolution parts of the image. **(e)** The self-rotation plot at $\kappa = 180°$ reveals the presence of non-crystallographic symmetry. The plot was generated from the native data reduced to space group $P4_1$. The self-rotation function shows a clear two-fold axis at the centre (red arrow) and four equivalent peaks on the perimeter generated by non-crystallographic twin-related two-fold axes (blue arrow). The self-rotation function at $\kappa = 90°$ shows the presence of the true crystallographic four-fold symmetry. **(f)** The Britton plot was generated after detwinning by plotting the number of negative intensities as a function of estimated twin fraction α[7]. An overestimation of α will result in increased number of negative intensities and the exact value of α can be extrapolated from this increase (black line). **(g)** Cumulative intensity distribution N(z) plots for Juno$_{20-228}$[8]. *I* is the intensity of centric and acentric reflections and $Z = I/[I]$. **(h)** The cumulative distribution of *H* determined using the twin operator *k, h, -l* for merohedral twinning[9]. The Yeates *H*-plots indicate the twinning fraction of Juno$_{20-228}$ is larger than 0.4. Dashed lines indicate the theoretically determined twinning slopes for five different twinning fractions. **(i)** Twinning analysis using the *L*-test[8]. The curve for the observed data (red) falls between the curves for theoretically perfectly twinned data (green) and theoretically untwinned data (blue). **(j)** Representative 1.8-Å resolution σ_A-weighted $2|F_o|-|F_c|$ electron density map of Juno$_{20-228}$ contoured at 1σ and superimposed with its final refined model.



Supplementary Figure 9 | Izumo1$_{22-254}$-Juno$_{20-228}$ complex structure determination. (a) Izumo1$_{22-254}$-Juno$_{20-228}$ complex crystals were grown in 0.1 M MES (pH 6.5), 20% (w/v) PEG 4000 and 0.6 M NaCl. **(b)** X-ray diffraction image of the Izumo1$_{22-254}$-Juno$_{20-228}$ complex. **(c)** Representative 2.4-Å resolution σ_A-weighted composite omit $2|F_o|-|F_c|$ electron density map of Izumo1$_{22-254}$-Juno$_{20-228}$ complex contoured at 1σ and superimposed with its final refined model. **(d)** Overview of the anomalous difference Fourier electron density map for sulphur atoms in the asymmetric unit of the Izumo1$_{22-254}$-Juno$_{20-228}$ complex. Twenty-four datasets collected at 1.7712 Å with R_{meas} <10% were merged and scaled using XSCALE[10]. The anomalous Fourier difference map was calculated using PHENIX[11] and contoured at 3σ (blue mesh). Twenty-four

strong electron density features were observed and this represent the entire S-atom substructure of Izumo1$_{22\text{-}254}$ and Juno$_{20\text{-}228}$. Note that the two sulphur atoms in a disulphide linkage display a single anomalous peak. Overall, the anomalous sulphur data confirm that the tracing and model building of Izumo1 and Juno were correct. The inset windows show magnified images of selected cysteine and methionine residues with corresponding anomalous signals.

Supplementary Table 1. Data collection and refinement statistics

	Izumo1$_{22-254}$	**Izumo1$_{22-268}$**	**Juno$_{20-228}$**	**Izumo1$_{22-254}$-Juno$_{20-228}$**	**Izumo1$_{22-254}$- Juno$_{20-228}$ merged sulphur anomalous data**
Data collection					
X-ray wavelength	0.9794	0.97949	1.5406	1.77120	1.77120
Space group	*H*3	*H*3	$P4_1$	*C*2	*C*2
Cell dimensions					
a, *b*, *c* (Å)	126.0 126.0 61.0	121.3 121.3 53.4	73.2 73.2 163.1	148.0 55.8 81.6	148.1 55.9 81.8
α, β, γ (°)	90 90 120	90 90 120	90 90 90	90 103.6 90	90 103.7 90
Resolution (Å)*	40.6 - 3.1 (3.3 - 3.1)	47.6 - 2.9 (3.1 - 2.9)	43.7 - 1.8 (1.9 - 1.8)	48.0 - 2.4 (2.5 - 2.4)	48.0 - 2.8 (2.9 - 2.8)
Total number of reflections	18,861	37,350	640,808	184,910	2,819,240
No. of unique reflections	6,483	6,484	79,079	25,587	16,254
R_{meas} (%)	11.5 (47.4)	12.0 (52.6)	11.7 (73.5)	7.6 (67.9)	9.9 (31.1)[#]
$\langle I/sI \rangle$	9.0 (3.1)	12.9 (3.7)	17.5 (3.0)	14.4 (2.8)	50.9 (21.8)
$CC_{1/2}$ (%)	99.3 (77.4)	99.6 (88.9)	99.9 (89.8)	99.8 (95.4)	100.0 (99.9)
Completeness (%)	97.4 (91.3)	100.0 (100.0)	100.0 (99.8)	100.0 (99.8)	99.9 (100.0)
Redundancy	2.9 (2.7)	5.8 (5.8)	8.1 (7.6)	7.2 (7.2)	
Anomalous multiplicity	-	-	-	-	89.6 (89.1)
CC_{anom}	-	-	-	-	0.51 (0.19)
Refinement					
Resolution (Å)	40.6 - 3.1 (3.2 - 3.1)	47.6 - 2.9 (3.0 - 2.9)	43.7 - 1.8 (1.9 - 1.8)	48.0 - 2.4 (2.5 - 2.4)	
R_{work}/ R_{free} (%)	21.9/ 25.0	18.8/ 23.2	19.6/ 22.7	18.0/ 22.2	
No. atoms					
Protein	1733	1822	6320	3478	
Ligand/ion	14	14	78	21	
Water	4	13	1198	25	
B-factors					
Protein	79.8	48.9	15.4	62.5	
Ligand/ion	107.5	79.9	36.6	107.6	
Water	58.3	43.9	21.1	52.4	
R.m.s deviations					
Bond lengths (Å)	0.004	0.009	0.008	0.008	
Bond angles (°)	1.0	1.3	1.1	1.2	
Ramachandran					
favoured/allowed (%)[$]	100	99.5	100	99.2	
generously allowed (%)[$]	0	0.5	0	0.8	
disallowed (%)[$]	0	0	0	0	
Molprobity validation					
Overall score	1.96	2.71	1.45	2.71	
Clashscore, all atoms	10.1	14.2	4.1	13.0	

*Highest resolution shell is shown in parenthesis.
[#]The sulphur anomalous data represent merged data from 24 individual datasets that process with R_{meas}<10%
[$]Calculated using Procheck

REFERENCES

1. Kamtekar, S. & Hecht, M. H. Protein Motifs. 7. The four-helix bundle: what determines a fold? *FASEB J.* **9,** 1013–1022 (1995).

2. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38,** W545–W549 (2010).

3. Hamaoka, B. Y. & Ghosh, P. Structure of the Essential Plasmodium Host Cell Traversal Protein SPECT1. *PLoS ONE* **9,** e114685 (2014).

4. Ishino, T., Yano, K., Chinzei, Y. & Yuda, M. Cell-Passage Activity Is Required for the Malarial Parasite to Cross the Liver Sinusoidal Cell Layer. *PLoS Biol* **2,** e4 (2004).

5. Ishino, T., Chinzei, Y. & Yuda, M. A Plasmodium sporozoite protein with a membrane attack complex domain is required for breaching the liver sinusoidal cell layer prior to hepatocyte infection†. *Cell. Microbiol.* **7,** 199–208 (2005).

6. Han, L. *et al.* Divergent evolution of vitamin B9 binding underlies Juno-mediated adhesion of mammalian gametes. *Curr. Biol.* **26,** R100–R101 (2016).

7. Fisher, R. G. & Sweet, R. M. Treatment of diffraction data from crystals twinned by merohedry. *Acta Crystallogr. Sect. A* **36,** 755–760 (1980).

8. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the *CCP* 4 program suite. *Acta Crystallogr. D Biol. Crystallogr.* **59,** 1131–1137 (2003).

9. Yeates, T. O. Simple statistics for intensity data from twinned specimens. *Acta Crystallogr. A* **44,** 142–144 (1988).

10. Kabsch, W. *XDS*. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132 (2010).

11. Adams, P. D. *et al. PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221 (2010).