Supporting Online Material

for

Proteoglycan-specific molecular switch for RPTPσ clustering and neuronal extension

Charlotte H. Coles,[1,*] Yingjie Shen,[2,*] Alan P. Tenney,[2,3,‡] Christian Siebold,[1,‡] Geoffrey C. Sutton,[1] Weixian Lu,[1] John T. Gallagher,[4] E. Yvonne Jones[1,†], John G. Flanagan,[2,†] A. Radu Aricescu[1,†]

Materials and methods

Construct design and cloning

A series of type IIa RPTP constructs were cloned into the pHLsec vector (*S1*), introducing an N-terminal secretion signal sequence and a C-terminal hexahistadine tag. Human RPTPσ short isoform (NCBI Ref. Seq. NM_130853.2): Ig1-2 (amino acids 30-231), Ig1-3 (30-321), Ig1-FN3 (30-602), sEcto (30-839). Human RPTP LAR (NCBI Ref. Seq. NM_002840.3): Ig1-2 (30-231), Ig1-FN3 (30-602). Human RPTPδ (NCBI Ref. Seq. BC106713.1): Ig1-2 (21-220), Ig1-FN3 (21-594). *Drosophila* RPTP LAR (NCBI Ref. Seq. NM_078880.3): Ig1-2 (33-232), Ig1-FN3 (33-606). Chick RPTPσ (NCBI Ref. Seq. NM_205407.1) Ig1-2 29-226. A human RPTPσ Ig1-3 R227Q+R228N mutant construct was designed to prevent proteolytic cleavage during Ig1-3 crystallisation trials. This construct was generated by two-step overlapping PCRs. The Glypican2-AP construct was generated by subcloning the coding sequence for amino acids 1-555 of mouse Glypican-2 (NCBI Ref. Seq. BC083180) into the APTag5 vector. Mouse RPTPσ sEcto-Fc and Neurocan-AP fusion proteins were generated as described previously (*S2*).

Protein purification and crystallisation

All Ig1-2 constructs were expressed in HEK-293T cells following transient transfection using polyethylenimine (*S1*). Human RPTPσ Ig1-3 R227Q+R228N was expressed in HEK-293T cells, in the presence of kifunensine (Toronto Research Chemicals), a specific inhibitor of α-mannosidase (*S3, S4*). Ig1-FN3 and sEcto constructs were expressed in either HEK-293T cells in the presence of kifunensine or in GnTI$^-$ HEK293S cells (*S5*). All proteins were purified from filtered cell culture media by immobilised nickel affinity chromatography (Chelating Sepharose Fast Flow, GE Healthcare) followed by heparin affinity chromatography (HiTrap Heparin HP column, GE Healthcare) and finally gel filtration

(Superdex resin, GE Healthcare) in 10 mM HEPES, 150 mM NaCl, pH 7.5. In addition, the human RPTPσ Ig1-3 R227Q+R228N protein, which contains two N-linked glycosylation sites, was incubated for 3 hours at 37 °C with endoglycosidase F1 (*S4, S6*) to cleave the N-linked $Man_9GlcNAc_2$ oligosaccharides to a single GlcNAc residue prior to the heparin affinity chromatography step of the purification. Expression of selenomethionine labelled protein was carried out as described previously (*S1*) and the protein was subsequently purified as described for native proteins.

Crystallisation trials, using 100 nl protein solution plus 100nl reservoir solution in sitting drop vapour diffusion format were set up in 96-well Greiner plates using a Cartesian Technologies robot (*S7*). Crystallisation plates were maintained at 20.5 °C in a TAP Homebase storage vault and imaged *via* a Veeco visualisation system (*S8*). Crystallisation conditions for all proteins are provided in Table S1.

Data collection and processing

All crystals obtained were cryo-protected as stated (Table S1) and then flash frozen at 100 K. X-ray diffraction data was collected at the Diamond Light Source, Oxfordshire, UK or the European Synchotron Radiation Facility, France as indicated. The diffraction images were indexed, integrated, scaled and merged using either the HKL2000 (*S9*) or the xia2 (*S10*) data processing suites. The chick RPTPσ Ig1-2 structure was determined by SAD analysis, using images collected from a selenomethionine labelled protein crystal at the selenium peak wavelength (λ = 0.9783 Å). Using SHELXD (*S11*) *via* the autoSHARP interface (*S12*), the positions of two selenium atoms were determined, before subsequent phase calculation, improvement and preliminary model building was performed. These preliminary models were used together with a high resolution (1.65 Å) native dataset in iterative rounds of automatic model building using RESOLVE (*S13*), ARP/wARP (*S14*) and manual building

using Coot (*S15*). The data obtained for all subsequent type IIa RPTP proteins were phased by molecular replacement using the chicken RPTPσ Ig1-2 solution as a search model in Phaser (*S16*). The resulting electron density maps were of high quality and enabled manual model adjustment in Coot (*S15*) and refinement in Refmac (*S17*), Phenix (*S18*) and Buster (*S19*). For structural validation PROCHECK (*S20*) and MolProbity (*S21*) were used to assess the stereochemical properties of the models and refinement statistics are given in Table S2. The superimposition of atomic models to compare the domain architecture between different structures, was performed using SHP (*S22*), based on Cα positions. Crystallographic figures were created using PyMOL (*S23*) and APBS was used to calculate the electrostatic potential of solvent accessible surfaces (*S24*).

Multi-angle light scattering (MALS)

MALS experiments were carried out on a Wyatt MALS/AFFFF System (Wyatt Technologies). All proteins were purified as described above, and samples were incubated alone or with an oligosaccharide ligand for an hour at room temperature prior to use in MALS analysis. For size defined heparin fragments (Iduron dp4-dp30; product codes HO04-HO30) a five-fold molar excess of oligosaccharide to protein was used, except for the series of experiments presented in Fig. S9, where a range of RPTPσ constructs were tested with a separate batch of dp10 at a two-fold molar excess (Table S3, highlighted in blue). Equivalent amounts (equal mg/ml concentrations) of HS (Iduron, GAG-HS01) and CS (Sigma, C4384) were used for MALS experiments as had been for heparin dp10 (in giving a five-fold molar excess). "x5" indicates that a five-fold greater quantity of HS or CS was used. All Ig1-FN3 proteins were injected at approximately 10 μM. Size exclusion chromatography was performed in 10 mM Tris, 50 mM NaCl, pH 7.5 on a Superdex75 or Superdex200 HR10/30 column (GE Healthcare), attached to an Agilent chromatography system. An Optilab rEX Refractive Index detector and a Dawn Helios II Multi-Angle Light Scattering (MALS)

detector recorded the refractive index and light scattering of the samples upon elution from the size exclusion column. The Wyatt software ASTRA was used to analyse all the data collected.

Native Mass Spectrometry

10 μM human RPTPσ Ig1-2 (desalted previously into 100 mM ammonium acetate pH 7.5) and 50 μM dp10 (ammonium salt) were incubated in a final volume of 500 μl for an hour. Gel filtration in 100 mM ammonium acetate pH 7.5 was carried out with this complex sample and the Ig1-2 protein alone, and fractions taken from the peaks on the UV traces and diluted to a concentration of 8 μM were used for analysis. Electrospray ionisation-mass spectrometry (ESI-MS) was carried out on a Waters Quadrupole Time-of-Flight (Q-Tof) Micro instrument with direct infusion of the protein samples. For native analysis, the protein-heparin complex spectrum was obtained using a cone voltage of 200 V.

Analytical Ultracentrifugation

Sedimentation equilibrium and velocity experiments were carried out at 20 ˚C in an Optima XL-I analytical ultracentrifuge (Beckman Instruments) utilising a scanning absorbance of 280 nm and interference optics. Samples of human RPTPσ Ig1-2 alone and in complex to heparin dp10, were taken directly from gel filtration in 20 mM Tris, 50 mM NaCl pH 7.5. Sedimentation velocity experiments were performed and analysed using the time derivative g(s*) method (*25*). Samples were equilibrated at 20 ˚C in the rotor, before the velocity was sequentially increased through speeds of 12, 18, 25, 35 and 50 krpm, 50 scans being made at each speed. ProFit (QuantumSoft) was used to fit Gaussian curves to the g(s*) plots obtained, and to describe the distribution of individual species.

Solid phase binding assays

Analysis of binding between RPTPσ sEcto-Fc and Glypican-2 or Neurocan alkaline phosphate (AP) fusion proteins was performed as described previously (*S2, S26*). All fusion proteins were produced in transiently transfected 293T cells. Briefly, the RPTPσ Fc-fusion protein was immobilised on 96-well Reactibind Protein A-coated plates (Pierce) and incubated with soluble proteoglycan-AP proteins which had been normalised for equal AP activity. AP activity was determined by measuring substrate turnover on a microplate reader as described (*S27, S28*). For experiments involving treatment with chondroitinase ABC (Sigma) or heparitinase III (HPNIII, Sigma), chondoitinase-treated, heparitinase-treated or mock-treated proteoglycan-APs were incubated at 37 °C for 2 hours prior to addition to the RPTPσ-Fc coated plates.

DRG neurite outgrowth assay and immunofluorescence

DRGs of P8-12 *RPTPσ*$^{+/+}$ or *RPTPσ* $^{-/-}$ mice were cut off from all roots. Neurons were dissociated and subjected to neurite outgrowth assays and analysis as described previously (*S2*). For experiments in Fig. 1, recombinant mouse Neurocan (5800-NC, R&D Systems) was used at 5 μg/ml and recombinant mouse Glypican-2 (2355-GP, R&D Systems) at 30 μg/ml. In Fig. S1, 30 μg/ml of Glypican-2 was pre-incubated with 1 unit/ml of Heparitinase III (HPNIII, H8891, Sigma) at 37 °C for 4 hrs where indicated, prior to use in outgrowth assays. For experiments in Fig. S15, recombinant mouse glypican-2 was used at 5 μg/ml, and Chondroitin Sulfate Proteoglycan mixture (CSPG, CC117, Millipore) was at 125 μg/ml. Proteoglycans were added to neuronal cultures 30 min after seeding and were replenished 24 hrs later. At 48 hrs after seeding, cells were fixed and immunostained with anti-GAP43 antibody (NB300-143, Novus). The whole of each well was scanned with a fluorescence microscope and the images were then analyzed for total neurite outgrowth using the Metamorph program.

For immunolocalization, dissociated DRG neurons were cultured without treatment then fixed and permeabilized, to allow co-staining with the cytoplasmic marker GAP43, at about 48 hrs after seeding. Where indicated, the fixed samples were then treated at 37 °C for 2 hrs with either 1 unit/ml of Chondroitinase ABC (C3667, Sigma) or a mixture of Heparinase I, II and III (H2519, H6512, H8891, Sigma), each at 1 unit/ml. Subsequently, samples were subjected to immunostaining using anti-chondroitin sulfate antibody (C8035, Sigma), anti-heparin/heparan sulfate antibody (OBT1698, AbD Serotec), anti-RPTPσ antibody (AF3430, R&D Systems) and anti-GAP43 antibody (NB300-143, Novus).

Statistics

Error bars show SEMs; *P* values were calculated by Student's t-test.

Supporting Figures

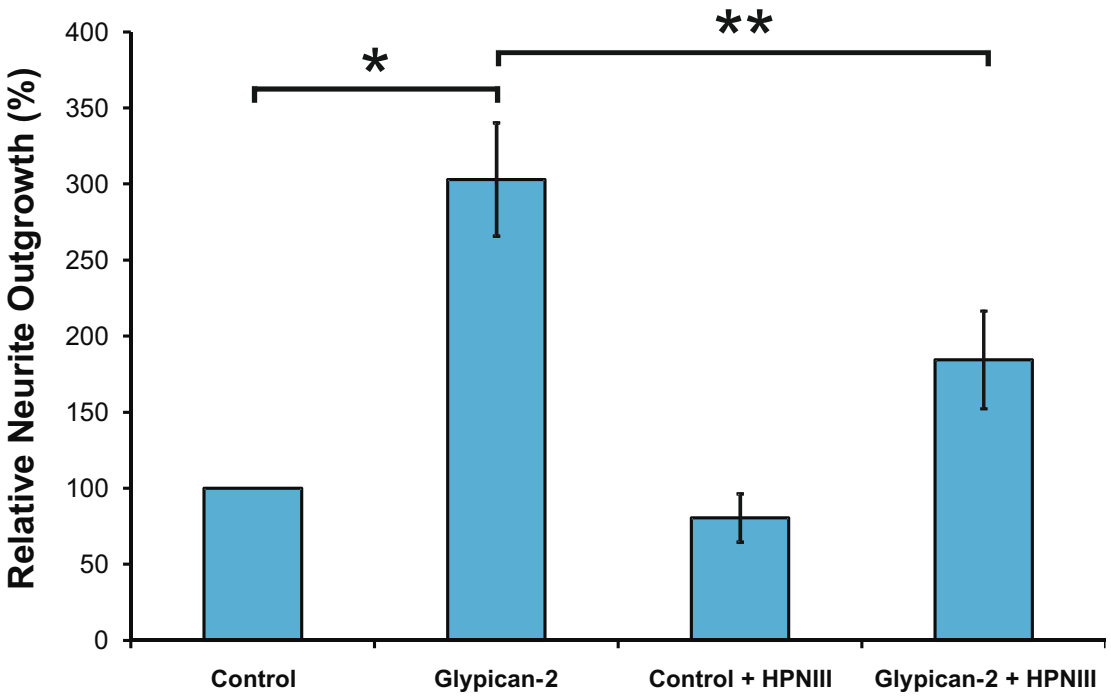


Figure S1. Glypican-2 induced neurite outgrowth is largely dependent upon the presence of its heparan sulphate moieties. Wild-type P8 mouse dorsal root ganglion (DRG) neurons were grown on substrates containing a poly-D-lysine/laminin mixture (control) or supplemented with glypican-2, either with or without heparitinase III (HPNIII) treatment. The outgrowth of DRG neurons, relative to the control (assigned 100 % outgrowth) was quantified. Treatment with HPNIII did not completely eliminate the growth promoting effect of the glypican-2 substrate, which may reflect either novel interactions involving the deglycosylated glypican core or incomplete enzymatic digestion of the heparan sulphate chains. Error bars show SEMs. **$p<0.005$ and *$p<0.05$, Student's t-test. The necessity of the chondroitin sulphate groups of neurocan for CSPG-mediated inhibition of DRG outgrowth, has been reported previously (*S2*).

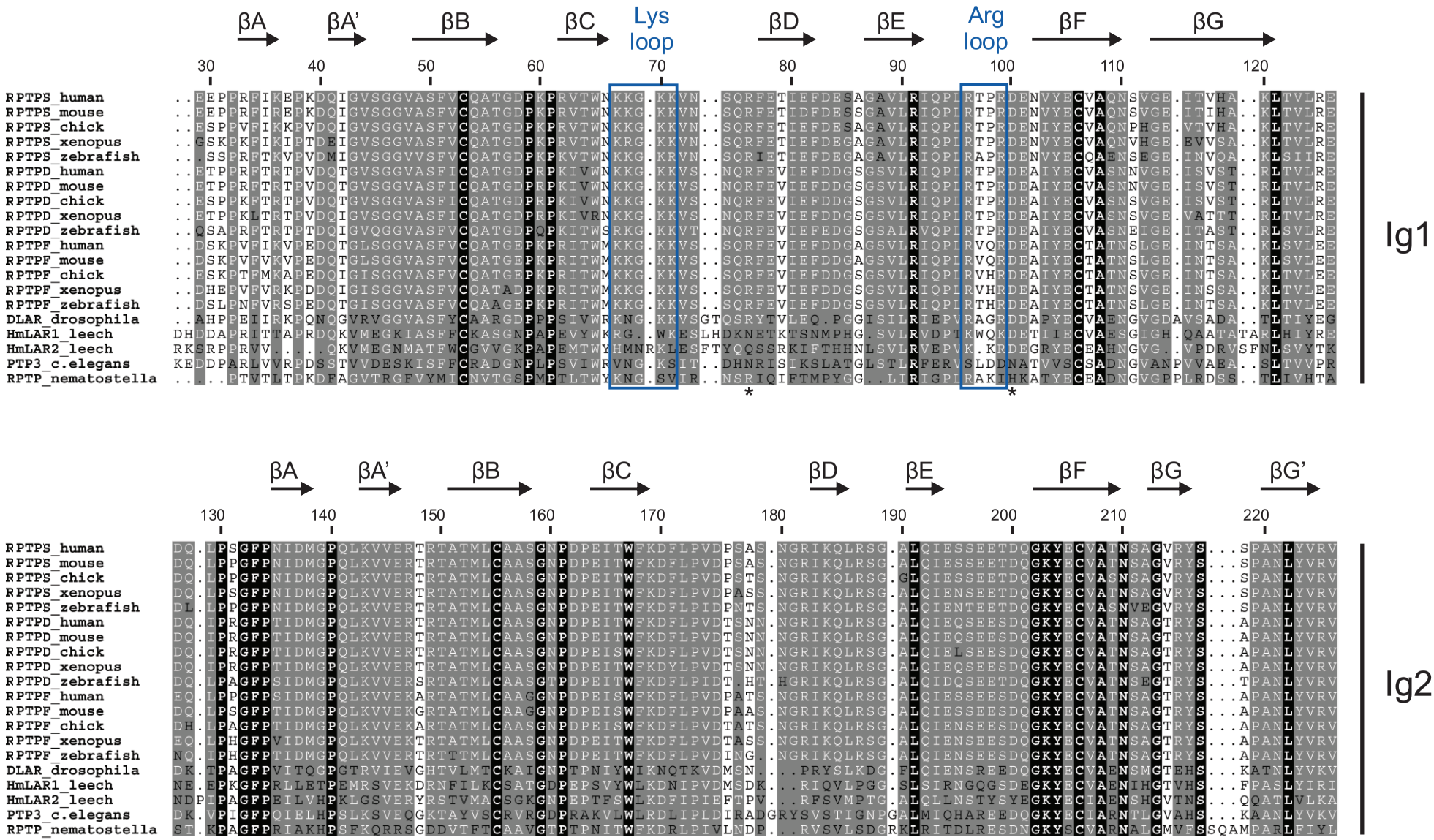
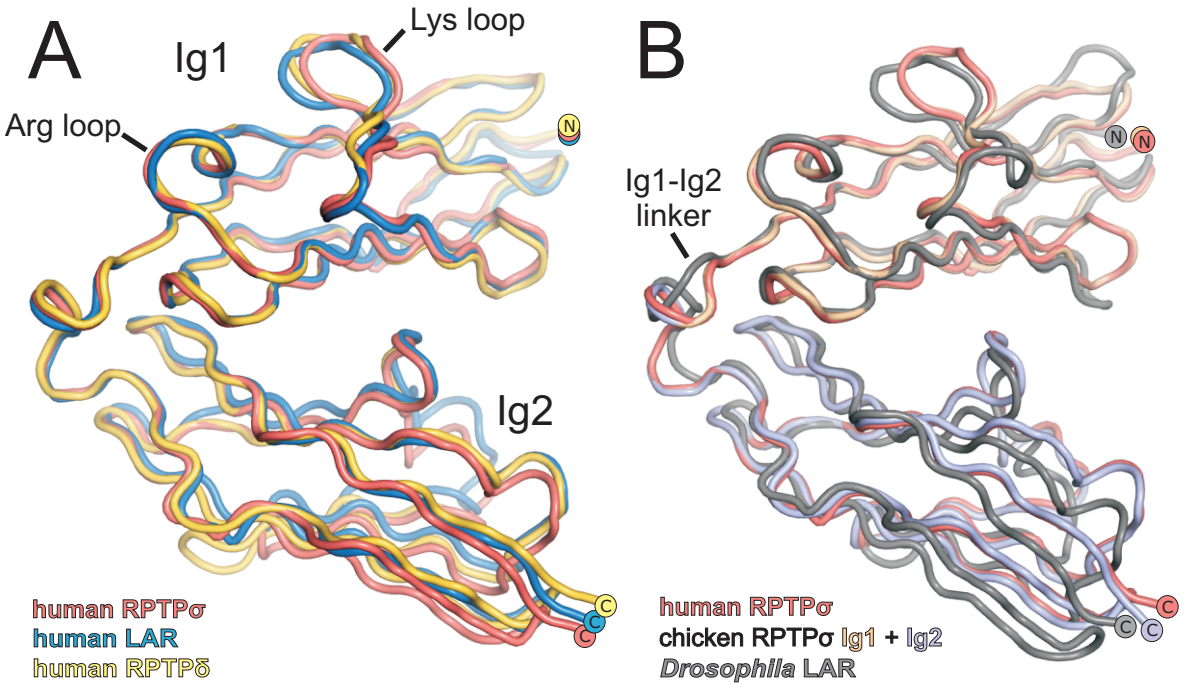

βA
βA'
βB
βC
Lys loop
βD
βE
Arg loop
βF
βG
βG'
Ig1
Ig2
RPTPS_human
RPTPS_mouse
RPTPS_chick
RPTPS_xenopus
RPTPS_zebrafish
RPTPD_human
RPTPD_mouse
RPTPD_chick
RPTPD_xenopus
RPTPD_zebrafish
RPTPF_human
RPTPF_mouse
RPTPF_chick
RPTPF_xenopus
RPTPF_zebrafish
DLAR_drosophila
HmLAR1_leech
HmLAR2_leech
PTP3_c.elegans
RPTP_nematostella

Figure S2. Sequence alignment of the two N-terminal Ig domains of the type IIa RPTP family members across species. Sequences correspond to the RPTP isoforms lacking the MeA and MeB exons, based on amino acid sequences taken from the following sources: RPTPσ human (NM_130854.2), mouse (BC052462.1), chicken (NM_205407.1), xenopus (NM_001141992.1) and zebrafish (XP_002666198.1); RPTPδ human (BC106713.1), mouse (EDL31049.1), chick (NP_990738.1), xenopus (NM_001090381.1) and zebrafish (NP_001159520.1); RPTP LAR human (NM_002840.3), mouse (NM_011213.2), chicken (XP_001233494.1), xenopus (NP_001081987) and zebrafish (NP_001077045.1); *Drosophila* LAR (NM_078880.3), leech HmLAR1 (AF017084.1) and HmLAR2 (AF017083.1), *C. elegans* PTP-3A (AF316539.1) and *Nematostella* RPTP (XP_001639024.1). Numbers above the sequence alignment correspond to amino acid residue numbers relative to the chicken RPTPσ sequence, where residue 1 is the initial methionine. Black arrows above the sequence alignment indicate the location of the β-strands within the two immunoglobulin domains, based on the structure of chick Ig1-2, assigned using ksdssp (*S29*). Blue boxes highlight the lysine loop between β strands C-D (containing K67, K68, K70 and K71) and the arginine loop between β strands E-F (containing R96 and R99). Black asterisks above the alignments highlight R76 and D100, the two residues forming the salt bridge which is disrupted upon binding of human LAR Ig1-2 to sucrose octasulphate.



C

Reference Structure	Moving Structure	Residue range used for alignment	Equivalent residues	Sequence identity of equivalences (%)	rmsd of equivalences (Å)
Human RPTPσ $I4_122$ Ig1	Chicken RPTPσ $I4_122$ Ig1	29 to 125	97	92	0.27
Human RPTPσ $I4_122$ Ig1	Human RPTPσ C2 Ig1	30 to 126	96	100	0.45
Human RPTPσ $I4_122$ Ig1	Human LAR $P3_221$ Ig1	30 to 126	97	71	0.76
Human RPTPσ $I4_122$ Ig1	Human RPTPδ $P3_221$ Ig1	21 to 117	97	73	0.70
Human RPTPσ $I4_122$ Ig1	Human RPTPδ C2 A Ig1	21 to 117	95	75	1.00
Human RPTPσ $I4_122$ Ig1	Human RPTPδ C2 B Ig1	21 to 117	97	73	0.84
Human RPTPσ $I4_122$ Ig1	*Drosophila* LAR C2 Ig1	35 to 228	91	44	0.99
Human RPTPσ $I4_122$ Ig2	Chicken RPTPσ $I4_122$ Ig2	132 to 225	94	96	0.19
Human RPTPσ $I4_122$ Ig2	Human RPTPσ C2 Ig2	133 to 226	94	100	0.76
Human RPTPσ $I4_122$ Ig2	Human LAR $P3_221$ Ig2	133 to 226	94	89	0.50
Human RPTPσ $I4_122$ Ig2	Human RPTPδ $P3_221$ Ig2	124 to 117	88	82	1.41
Human RPTPσ $I4_122$ Ig2	Human RPTPδ C2 A Ig2	124 to 117	89	83	1.25
Human RPTPσ $I4_122$ Ig2	Human RPTPδ C2 B Ig2	124 to 117	87	82	1.45
Human RPTPσ $I4_122$ Ig2	*Drosophila* LAR C2 Ig2	138 to 228	92	46	1.26

Structure	Interdomain interface (Å^2)	ΔG (kcal/mol)	Ig1-Ig2 angle deviation (°)	Surface complementarity statistic
Human RPTPσ $I4_122$	640	-6.9	0.0	0.681
Chicken RPTPσ $I4_122$	659	-7.8	2.2	0.723
Human RPTPσ C2	662	-7.4	-5.0	0.739
Human LAR $P3_221$	662	-7.7	-7.8	0.718
Human RPTPδ $P3_221$	654	-6.9	-5.9	0.693
Human RPTPδ C2 A	694	-6.0	-10.2	0.700
Human RPTPδ C2 B	644	-6.9	-3.7	0.683
Drosophila LAR C2	588	0.0	11.3	0.590

Figure S3. Structural alignment of the type IIa RPTPs illustrates the conserved architecture of the two N-terminal domains across the type IIa RPTP family. (**A**) Ig1 domains from human RPTP LAR ($P3_221$) and RPTPδ ($P3_221$) crystal structures were aligned with the Ig1 domain from human RPTPσ ($I4_122$) using SHP (*S22*), to obtain an overlay of the human RPTP structures. (**B**) Ig1 from the chicken RPTPσ ($I4_122$) and DLAR (C2) structures were similarly aligned with the Ig1 from human RPTPσ ($I4_122$). (**C**) Structural comparison of the Ig1-Ig2 domain interface for the type IIa RPTPs. Ig1-Ig2 angle deviation represents the angle required to realign the Ig2 domains (containing the residues listed) from the Ig1-2 structures on human RPTPσ $I4_122$ Ig2 after having first superposed each Ig1-2 structure to align with human RPTPσ $I4_122$ Ig1 and was calculated from the rotation matrix obtained for the Ig2 domain transformation in SHP . Interface surface areas per domain and the change in free energy for the Ig1-Ig2 interaction were estimated for each structure (using residues equivalent to 29-130 and 131-226 as molecule A and molecule B respectively), in the PISA prediction program (EBI, EMBL) and the surface complementarity was calculated using SC (*S30*). Although the buried surface area at the Ig1-Ig2 domain interface is comparable and the interdomain angle variation is small across the structures, the ΔG value and surface complementarity score for this domain arrangement in the DLAR structure are both notably less favourable.

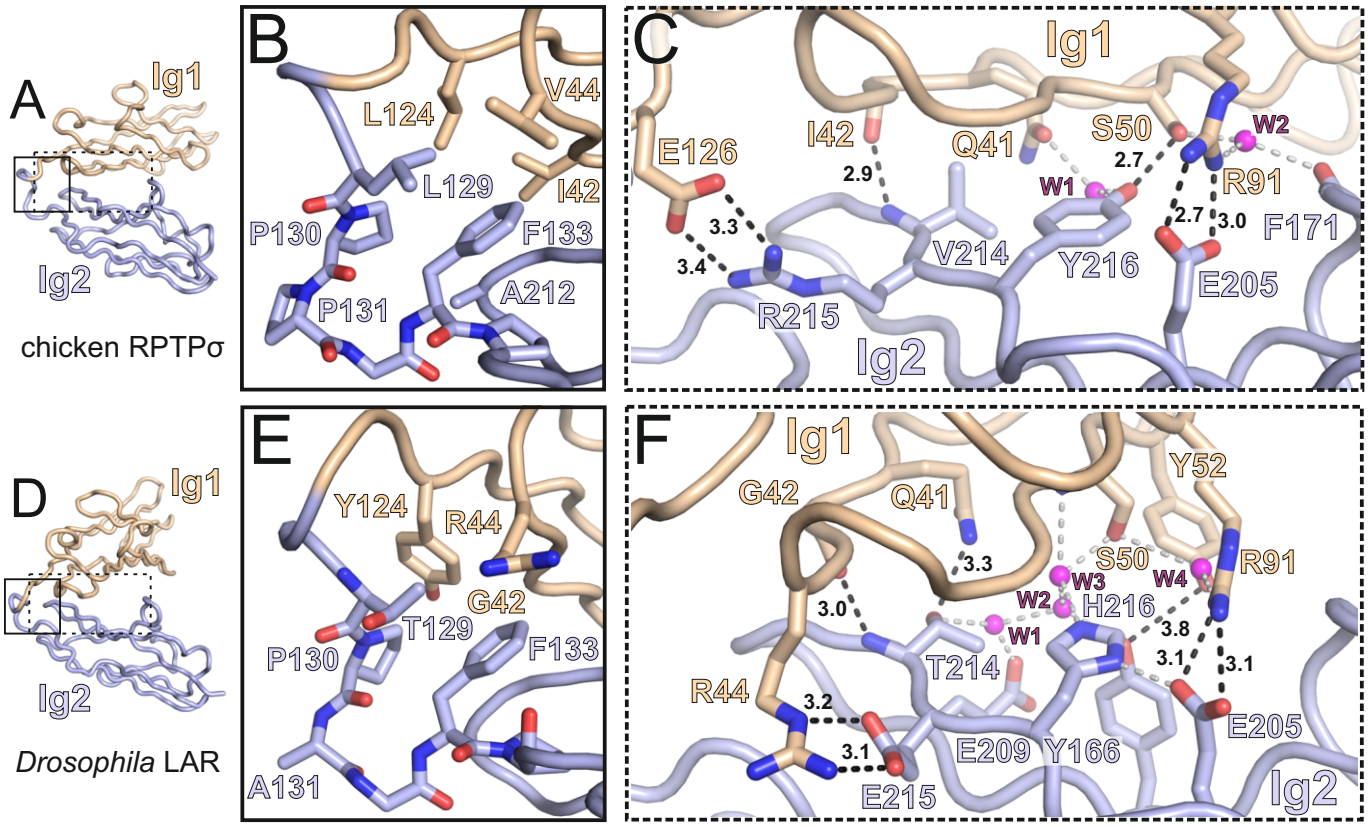


Figure S4. Comparison of the Ig1-Ig2 interdomain interactions observed in chicken RPTPσ and *Drosophila* LAR (DLAR) Ig1-2 crystal structures. Chicken RPTPσ: (**A**) Ribbon representation of the Ig1-2 crystal structure, highlighting the linker region (solid box) and the domain interface (dashed box). (**B**) Sidechains of hydrophobic residues L124, L129, P130 and F133, which lie on the Ig1-Ig2 Pro-rich loop, also pack closely with I42, V44 and A212 in a hydrophobic interdomain region, (**C**) Two salt bridges R91-E205 and E126-R215 and a network of hydrogen bonds involving the hydroxyl groups of S50 and Y216, the sidechain amide of Q41, the backbone carbonyls of I42 and F171, the backbone amide of V214 and two water molecules (purple spheres), hold Ig1 and Ig2 in a rigid arrangement. The interdomain interactions, including the two highlighted water molecules, observed in the chick RPTPσ Ig1-2 structure are present in the three human RPTP structures and the residues involved are also highly conserved across species (Fig. S2). DLAR: (**D**) Ribbon representation of the Ig1-2 crystal structure, highlighting the linker region (solid box) and the domain interface (dashed box), (**E**) Sidechains of P130 and F133 from the Ig1-Ig2 loop, pack with the carbon backbones of the Y124 and R44 sidechains, but this hydrophobic region is less extensive than in the chicken RPTPσ structure. (**F**) Two salt bridges R91-E205 and R44-E215 and a network of hydrogen bonds involving the backbone carbonyl of G42, the backbone amide of T214, the sidechain amide of Q41, the carboxyl group of E209, the imidazole group of H216, the hydroxyl groups of S50, Y52, Y166 and T214 and four water molecules, hold Ig1 and Ig2 in a similarly rigid arrangement. Interdomain interactions directly between protein residues and through water molecules are illustrated with black and grey dashed lines respectively. Blue and red atoms represent nitrogen and oxygen. Notably, there is a charge swap in residue 215 between the vertebrate (R215) and *Drosophila* (D215) crystal structures, which forms an interdomain salt bridge in all proteins.

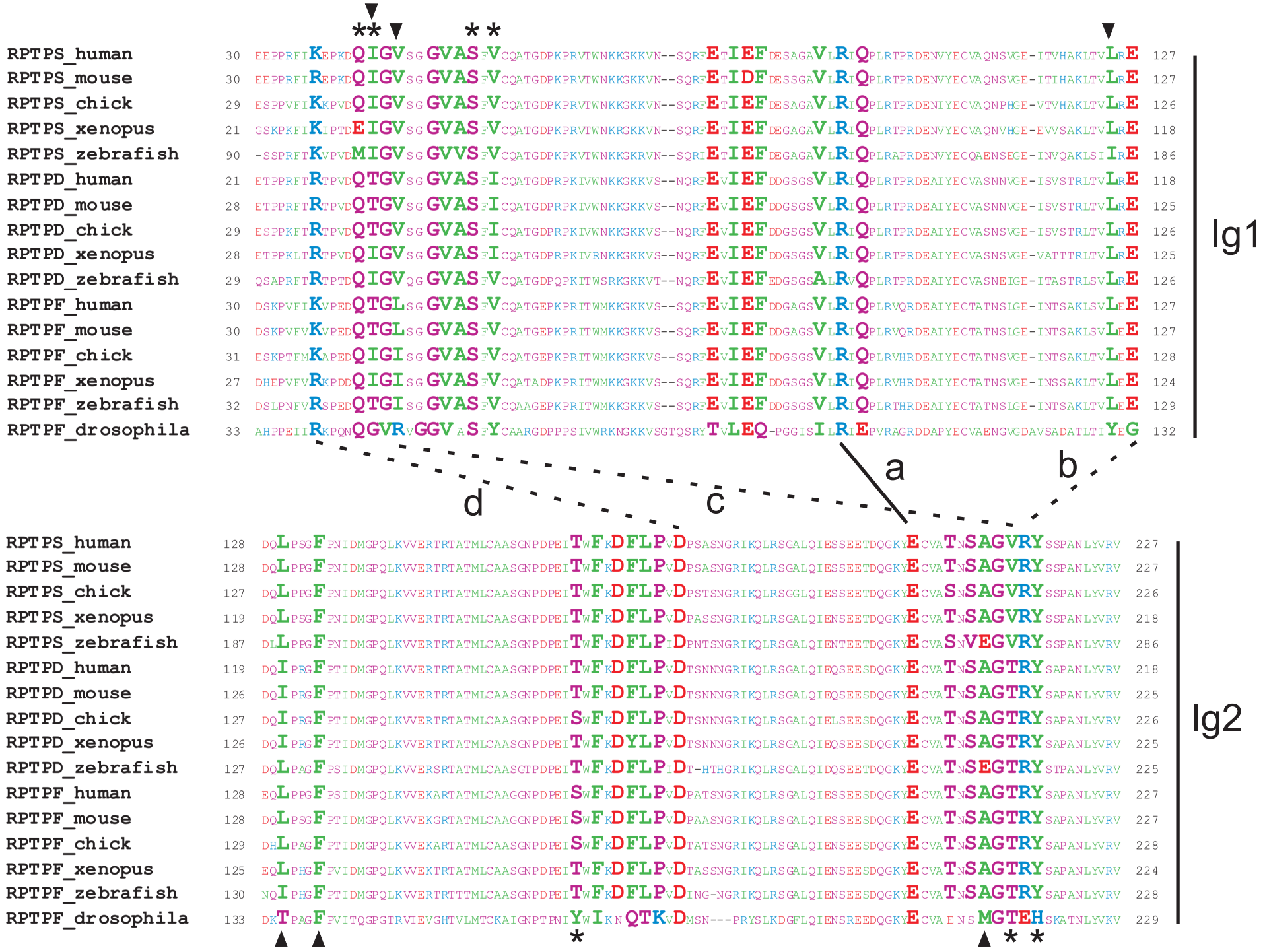

RPTPS_human 30 EEPPRFIKEPKDQIGVSGGVASFVCQATGDPKPRVTWNKKGKKVN--SQRFETIEFDESAGAVLRIQPLRTPRDENVYECVAQNSVGE-ITVHAKLTVLRE 127
RPTPS_mouse 30 EEPPRFIREPKDQIGVSGGVASFVCQATGDPKPRVTWNKKGKKVN--SQRFETIDFDESSGAVLRIQPLRTPRDENVYECVAQNSVGE-ITIHAKLTVLRE 127
RPTPS_chick 29 ESPPVFIKKPVDQIGVSGGVASFVCQATGDPKPRVTWNKKGKKVN--SQRFETIEFDESAGAVLRIQPLRTPRDENIYECVAQNPHGE-VTVHAKLTVLRE 126
RPTPS_xenopus 21 GSKPKFIKIPTDEIGVSGGVASFVCQATGDPKPRVTWNKRGKKVN--SQRFETIEFDEGAGAVLRIQPLRTPRDENVYECVAQNVHGE-EVVSAKLTVLRE 118
RPTPS_zebrafish 90 -SSPRFTKVPVDMIGVSGGVVSFVCQATGDPKPKVTWNKKGKRVN--SQRIETIEFDEGAGAVLRIQPLRAPRDENVYECQAENSEGE-INVQAKLSIIRE 186
RPTPD_human 21 ETPPRFTRTPVDQTGVSGGVASFICQATGDPRPKIVWNKKGKKVS--NQRFEVIEFDDGSGSVLRIQPLRTPRDEAIYECVASNNVGE-ISVSTRLTVLRE 118
RPTPD_mouse 28 ETPPRFTRTPVDQTGVSGGVASFICQATGDPRPKIVWNKKGKKVS--NQRFEVIEFDDGSGSVLRIQPLRTPRDEAIYECVASNNVGE-ISVSTRLTVLRE 125
RPTPD_chick 29 ESPPKFTRTPVDQTGVSGGVASFICQATGDPRPKIVWNKKGKKVS--NQRFEVIEFDDGSGSVLRIQPLRTPRDEAIYECVASNSVGE-ISVSTRLTVLRE 126
RPTPD_xenopus 28 ETPPKLTRTPVDQIGVSGGVASFICQATGDPRPKIVRNKKGKKVS--NQRFEVIEFDDGSGSVLRIQPLRTPRDEAIYECVASNSVGE-VATTTRLTVLRE 125
RPTPD_zebrafish 29 QSAPRFTRTPTDQIGVQGGVASFVCQATGDPQPKITWSRKGKKVT--NQRFEVIEFEDGSGSALRVQPLRTPRDEAIYECVASNEIGE-ITASTRLSVLRE 126
RPTPF_human 30 DSKPVFIKVPEDQTGLSGGVASFVCQATGEPKPRITWMKKGKKVS--SQRFEVIEFDDGAGSVLRIQPLRVQRDEAIYECTATNSLGE-INTSAKLSVLEE 127
RPTPF_mouse 30 DSKPVFVKVPEDQTGLSGGVASFVCQATGEPKPRITWMKKGKKVS--SQRFEVIEFDDGAGSVLRIQPLRVQRDEAIYECTATNSLGE-INTSAKLSVLEE 127
RPTPF_chick 31 ESKPTFMKAPEDQIGISGGVASFVCQATGEPKPRITWMKKGKKVS--SQRFEVIEFDDGSGSVLRIQPLRVHRDEAIYECTATNSVGE-INTSAKLTVLEE 128
RPTPF_xenopus 27 DHEPVFVRKPDDQIGISGGVASFVCQATADPKPRITWMKKGKKVS--SQRFEVIEFDDGSGSVLRIQPLRVHRDEAIYECTATNSVGE-INSSAKLTVLEE 124
RPTPF_zebrafish 32 DSLPNFVRSPEDQTGISGGVASFVCQAAGEPKPRITWMKKGKKVS--SQRFEVIEFDDGSGSVLRIQPLRTHRDEAIYECTATNSLGE-INTSAKLTVLEE 129
RPTPF_drosophila 33 AHPPEIIRKPQNQGVRVGGVASFYCAARGDPPPSIVWRKNGKKVSGTQSRYTVLEQ-PGGISILRIEPVRAGRDDAPYECVAENGVGDAVSADATLTIYEG 132
Ig1
a
b
c
d
RPTPS_human 128 DQLPSGFPNIDMGPQLKVVERTRTATMLCAASGNPDPEITWFKDFLPVDPSASNGRIKQLRSGALQIESSEETDQGKYECVATNSAGVRY SSPANLYVRV 227
RPTPS_mouse 128 DQLPPGFPNIDMGPQLKVVERTRTATMLCAASGNPDPEITWFKDFLPVDPSASNGRIKQLRSGALQIESSEETDQGKYECVATNSAGVRY SSPANLYVRV 227
RPTPS_chick 127 DQLPPGFPNIDMGPQLKVVERTRTATMLCAASGNPDPEITWFKDFLPVDPSTSNGRIKQLRSGGLQIESSEETDQGKYECVASNSAGVRY SSPANLYVRV 226
RPTPS_xenopus 119 DQLPSGFPNIDMGPQLKVVERTRTATMLCAASGNPDPEITWFKDFLPVDPASSNGRIKQLRSGALQIENSEETDQGKYECVATNSAGVRY SSPANLYVRV 218
RPTPS_zebrafish 187 DLLPPGFPNIDMGPQLKVVERTRTATMLCAASGNPDPEITWFKDFLPIDPNTSNGRIKQLRSGALQIENTEETDQGKYECVASNVEGVRY SSPANLYVRV 286
RPTPD_human 119 DQIPRGFPTIDMGPQLKVVERTRTATMLCAASGNPDPEITWFKDFLPVDTSNNNGRIKQLRSGALQIEQSEESDQGKYECVATNSAGTRY SAPANLYVRV 218
RPTPD_mouse 126 DQIPRGFPTIDMGPQLKVVERTRTATMLCAASGNPDPEITWFKDFLPVDTSNNNGRIKQLRSGALQIEQSEESDQGKYECVATNSAGTRY SAPANLYVRV 225
RPTPD_chick 127 DQIPRGFPTIDMGPQLKVVERTRTATMLCAASGNPDPEISWFKDFLPVDTSNNNGRIKQLRSGALQIELSEESDQGKYECVATNSAGTRY SAPANLYVRV 226
RPTPD_xenopus 126 DQIPRGFPTIDMGPQLKVVERTRTATMLCAASGNPDPEITWFKDYLPVDTSNNNGRIKQLRSGALQIEQSEESDQGKYECVATNSAGTRY SAPANLYVRV 225
RPTPD_zebrafish 127 DQLPAGFPSIDMGPQLKVVERSRTATMLCAASGTPDPEITWFKDFLPIDT-HTHGRIKQLRSGALQIDQSEETDQGKYECVATNSEGTRY STPANLYVRV 225
RPTPF_human 128 EQLPPGFPSIDMGPQLKVVEKARTATMLCAAGGNPDPEISWFKDFLPVDPATSNGRIKQLRSGALQIESSEESDQGKYECVATNSAGTRY SAPANLYVRV 227
RPTPF_mouse 128 DQLPSGFPTIDMGPQLKVVEKGRTATMLCAAGGNPDPEISWFKDFLPVDPAASNGRIKQLRSGALQIESSEESDQGKYECVATNSAGTRY SAPANLYVRV 227
RPTPF_chick 129 DHLPAGFPTIDMGPQLKVVEKARTATMLCAASGNPDPEISWFKDFLPVDTATSNGRIKQLRSGALQIENSEESDQGKYECVATNSAGTRY SAPANLYVRV 228
RPTPF_xenopus 125 EQLPHGFPVIDMGPQLKVVEKTRTATMLCAASGNPDPEITWFKDFLPVDTASSNGRIKQLRSGALQIENSEESDQGKYECVATNSAGTRY SAPANLYVRV 224
RPTPF_zebrafish 130 NQIPHGFPTIDMGPQLKVVERTRTTTMLCAASGNPDPEITWFKDFLPVDING-NGRIKQLRSGALQIENSEESDQGKYECVATNSAGTRY SAPANLYVRV 228
RPTPF_drosophila 133 DKTPAGFPVITQGPGTRVIEVGHTVLMTCKAIGNPTPNIYWIKNQTKVDMSN---PRYSLKDGFLQIENSREEDQGKYECVAENSMGTEH SKATNLYVKV 229
Ig2

Figure S5. Sequence alignment of the two N-terminal Ig domains of the type IIa RPTP family members across species (sequences from Fig. S2). Hydrophobic amino acids (AVFMILWY) are coloured green, acidic (DE) are red, basic (RKH) are blue and residues (STCNGQP) are magenta. The bold and dashed lines represent interdomain salt bridges that are either conserved or non-conserved between RPTP Ig1-2 crystal structures. Residues highlighted with black asterisks appear to be involved in interdomain hydrogen bonding interactions and those indicated with black arrows pack in a hydrophobic region near to the rigid interdomain linker.

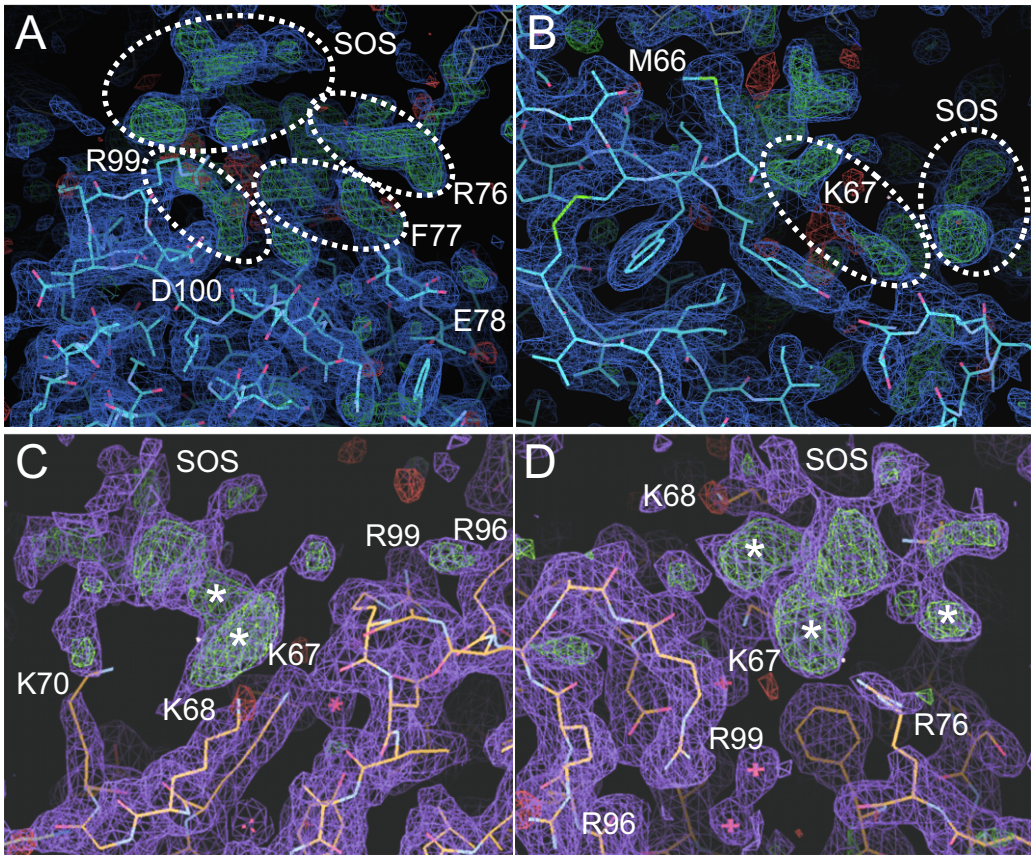


Figure S6. Model building for the 2.05 Å human LAR Ig1-2-sucrose octasulphate (SOS) crystal structure. (**A**) and (**B**): SigmaA-weighted electron density maps from refinement of the initial human LAR Ig1-2 model after molecular replacement in Phaser (*S16*). $2F_o$-F_c maps (blue) and F_o-F_c maps (green and red) are contoured at 1 σ and ± 3 σ respectively. Unmodelled features in the electron density maps are highlighted by white dashed lines. (**C**) and (**D**): SigmaA-weighted electron density maps after refinement in Phenix (*S18*) prior to addition of the SOS ligand to the model. $2F_o$-F_c maps (purple) and F_o-F_c maps (green and red) are contoured at 1 σ and ± 3 σ respectively. Asterisks mark the density into which three sulphate groups of the SOS ligand were initially placed, and onto which a full SOS molecule was superposed. The five-membered ring appeared to fit well into the electron density, but this left the six-membered ring of the ligand pointing into the solvent and consequently without electron density. Therefore the SOS ring was split at the disaccharide linker and just the five-membered ring was included in the model for refinement in Phenix (*S18*). The R_{work} and R_{free} decreased from 20.0 % and 23.4 % to 19.5 % and 22.5 % respectively upon addition of the five-membered ring from SOS to the model, supporting the inclusion of the ligand in the final crystal structure. Extensive co-crystallization attempts with fragments of chondroitin sulfate or heparin (widely-used to mimic highly sulfated regions of heparan sulfate) and various type IIa RPTP constructs did not produce interpretable electron density maps for the glycans, presumably due to their structural heterogeneity.

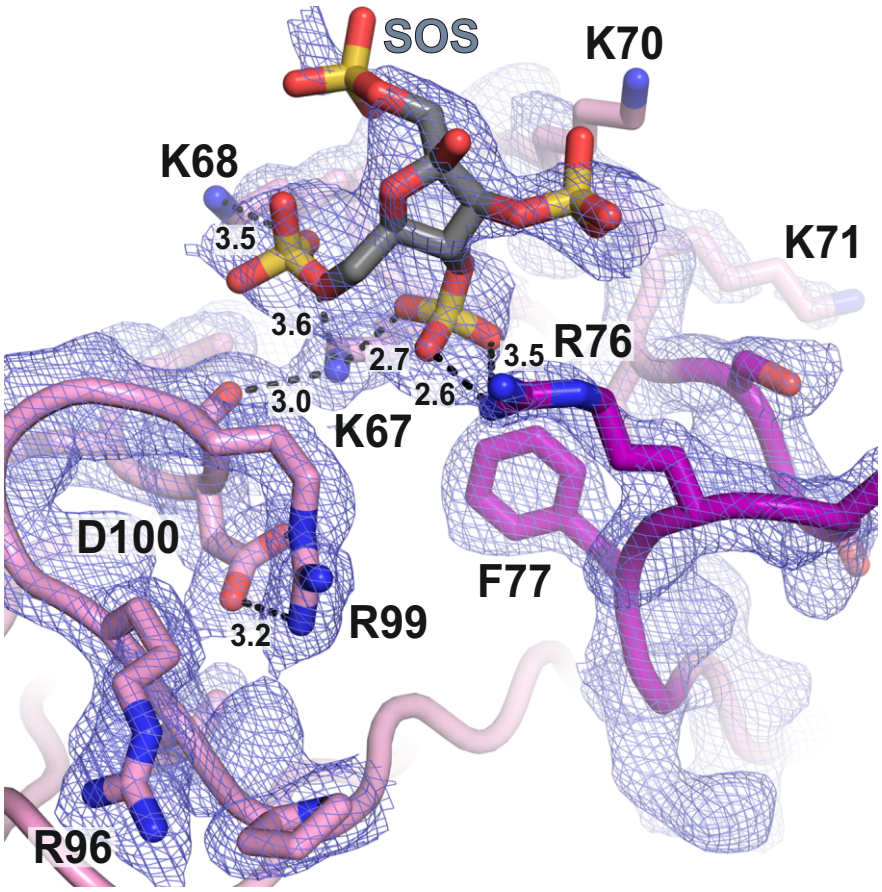


Figure S7. Electron density map for the 2.05 Å crystal structure of human LAR Ig1-2 in complex with sucrose octasulphate (SOS). The flexible proteoglycan binding loop is coloured in purple while the remainder of the LAR protein is coloured in pink. The carbon backbone of the SOS ligand is shown in grey. Non-carbon atoms are highlighted as follows: nitrogen, blue; oxygen, red; sulphur, yellow. The sidechains of K67, K68 and R76 are well ordered in the SOS-bound LAR crystal structure and are suitably positioned to form electrostatic interactions with the negatively charged sulphate groups of the SOS ligand (distances less than 3.6 Å between oppositely charged groups are indicated by black dashed lines). R96 and R99 which lie on the Arg-loop are also well ordered, but play no role in SOS-binding in this crystal structure. R99 instead assumes the role of R76 in the apo-LAR crystal structure, by forming a salt bridge with D100 after the R76-D100 salt bridge is broken and R76 becomes involved in ligand binding. K70 is disordered and clear density is not visible for the sidechain. Blue mesh represents the SigmaA-weighted $2F_o$-F_c electron density map contoured at 1 σ, after the final round of structure refinement in Phenix (*S18*).

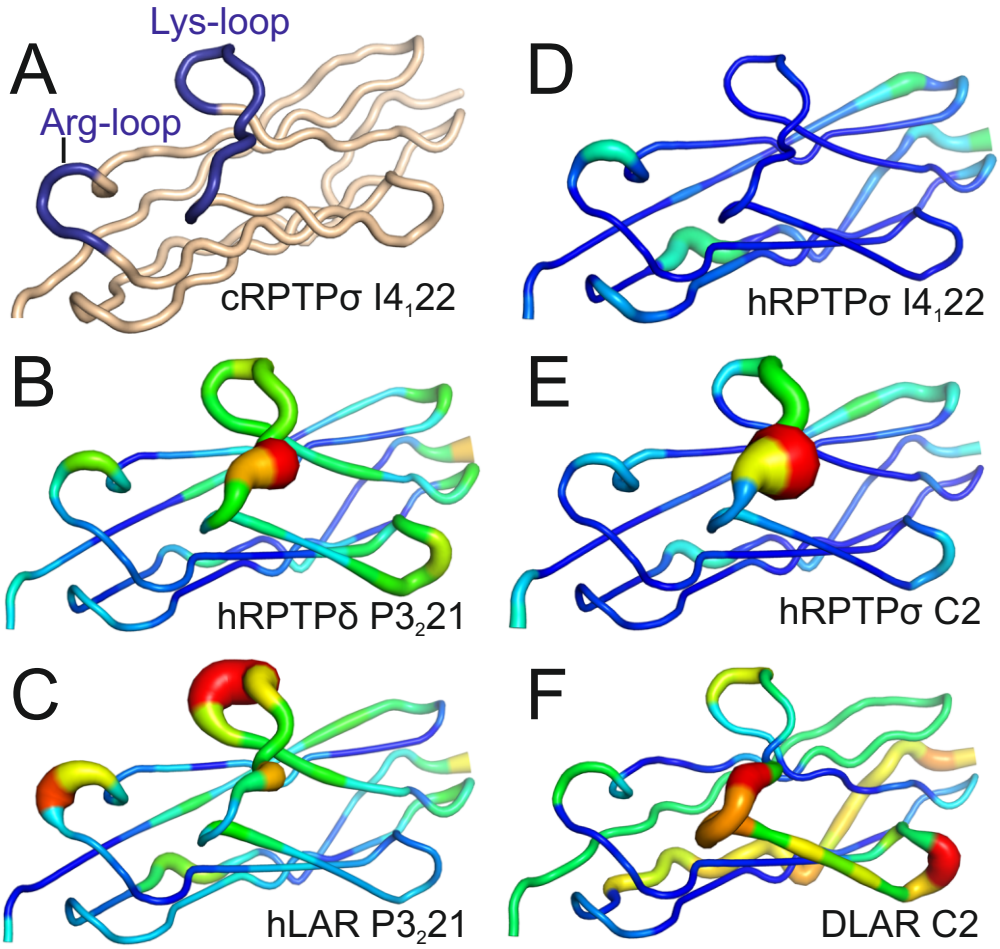


Figure S8. Conformational flexibility of the proteoglycan binding loop across type IIa RPTP structures. (**A**) Ribbon representation of the chicken RPTPσ Ig1 backbone. The "Lys"-loop and "Arg"-loop which harbour the crucial heparin binding residues, are highlighted in blue. The rmsd between the Ig1 C_α positions of chicken RPTPσ ($I4_122$ space group) and either (**B**) human RPTPδ ($P3_221$ space group), (**C**) human LAR ($P3_221$ space group), (**D**) human RPTPσ ($I4_122$ space group), (**E**) human RPTPσ (C2 space group) or (**F**) DLAR (C2 space group) was measured using SHP (*S22*) and is plotted upon the chicken RPTPσ Ig1 structure; increasing rmsd values are shown using a rainbow scale (blue, low rmsd; red, high rmsd) and the increasing thickness of the protein backbone. The Lys-loop displays the greatest movement across crystal structures, indicating that this region has an inherent flexibility, which may facilitate binding of the type IIa RPTP receptors to different GAG ligands.

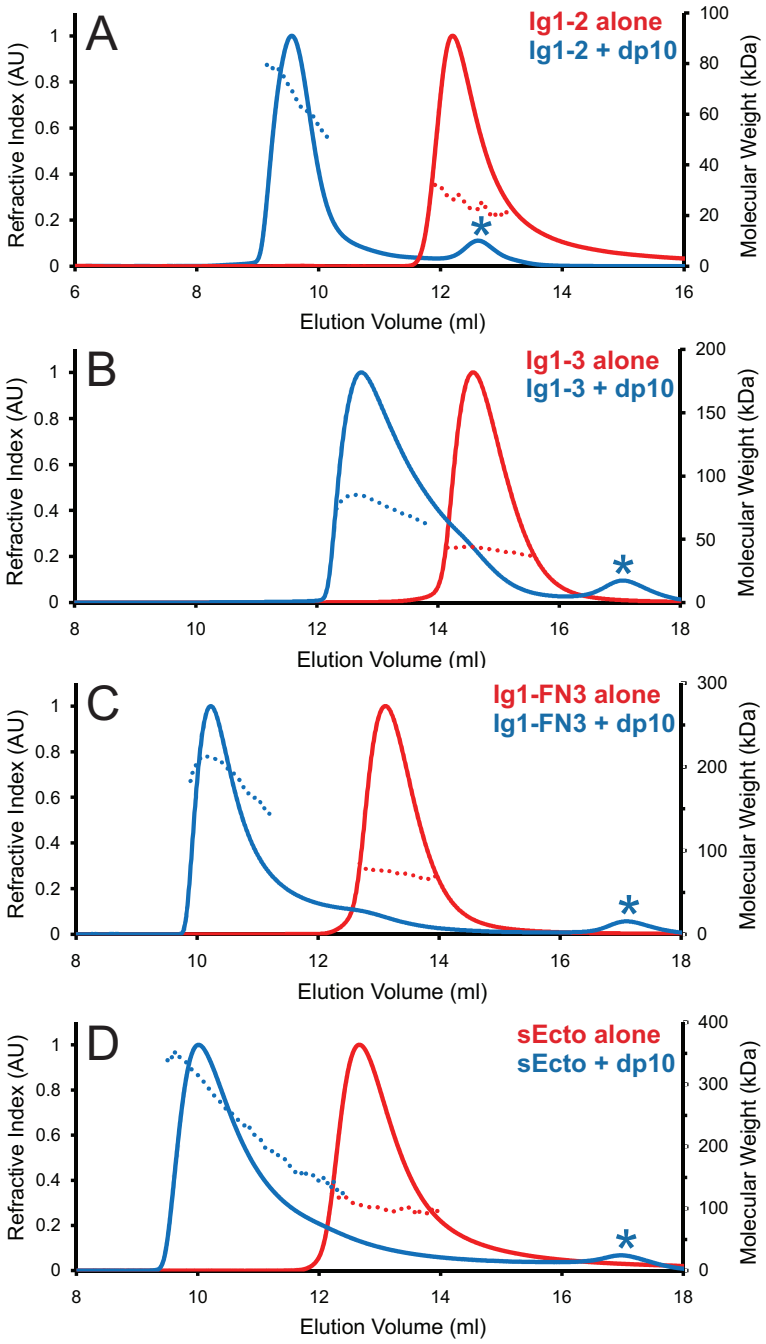


Figure S9. Heparin-induced dimerisation of human RPTPσ ectodomain constructs. A series of human RPTPσ constructs were incubated with (blue) and without (red) a two-fold molar excess of heparin dp10 before SEC-MALS (size-exclusion chromatography-multi-angle light scattering) analysis; (**A**) Ig1-2, (**B**) Ig1-3, (**C**) Ig1-FN3, (**D**) sEcto. Refractive index traces (scaled within each panel) are shown by bold lines and the measured molecular weights are shown by dotted lines. Peaks in the refractive index that are indicated by an asterisk correspond to excess dp10 ligand. A Superdex 75 column (1 cm x 30 cm) was used for the Ig1-2 construct (A) while all other constructs (B-D) were analysed using a Superdex 200 (1 cm x 30 cm) column.

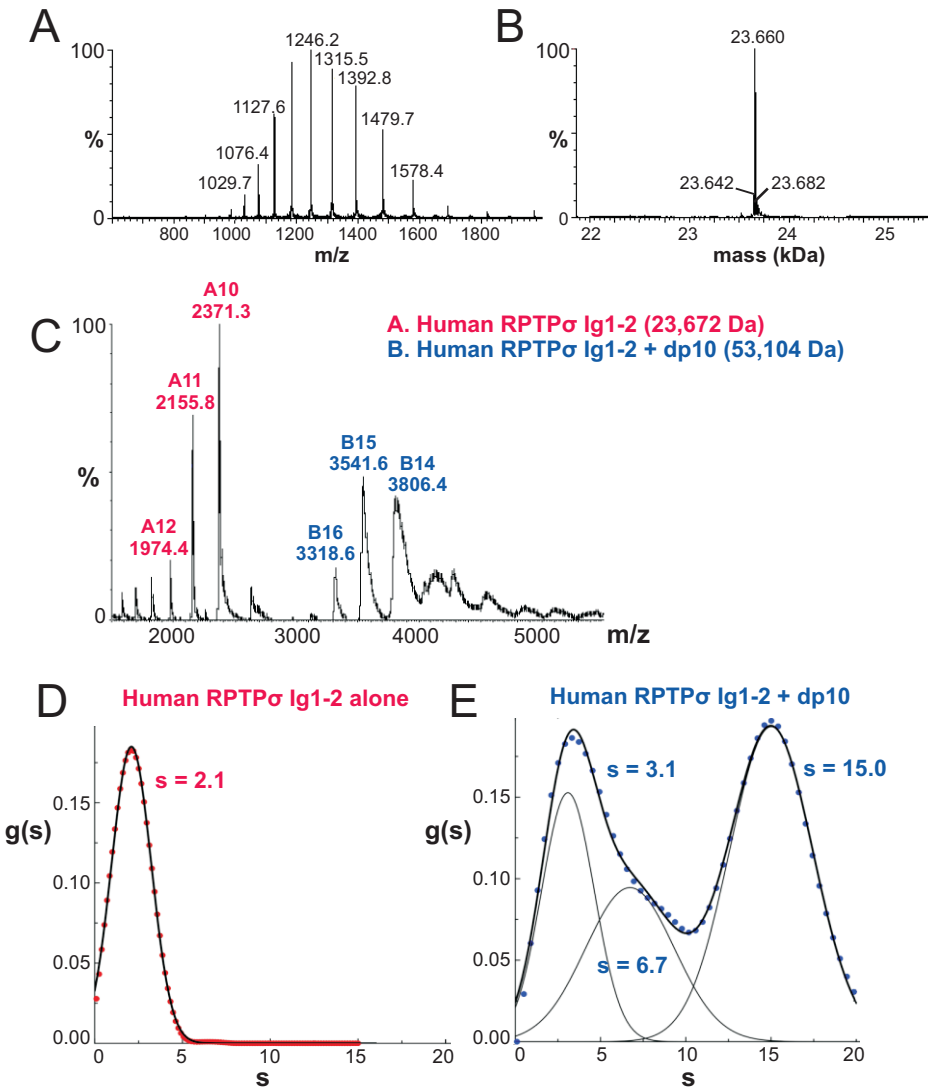


Figure S10. Investigation of type IIa RPTP clustering in solution. After incubation of human RPTPσ alone or with heparin dp10 (Iduron H010), electrospray ionisation-mass spectrometry (ESI-MS) and analytical ultracentrifugation (AUC) were used to analyse the oligomerisation state of the protein. The mass of human RPTPσ Ig1-2 alone was confirmed using ESI-MS under denaturing conditions; (**A**) an ion series corresponding to different charge states of the proteins was obtained and (**B**) this ion series was deconvoluted to give the molecular weight of the protein 23,660 Da (estimated molecular weight based on sequence is 23,676 Da). (**C**) Human RPTPσ Ig1-2 in complex with heparin dp10 and analysed using ESI-MS under native conditions. Two main ion series are observed; a series at lower m/z values corresponding to the protein alone (red) with a deconvoluted mass of 23, 672 Da and a series at higher m/z values (blue) with a deconvoluted mass of 53,104 Da which would correspond to a protein-dp10 complex. Further ion series at higher m/z values still are also present, which most likely represent higher order oligomers, but these series couldn't be deconvoluted to obtain mass values due to the overlapping arrangement and low abundance of the peaks. A series of sedimentation velocity AUC experiments were performed with the human RPTPσ Ig1-2 protein either alone (**D**) or in complex with heparin dp10 (**E**). The coloured circles represent the measured data points, while the black lines represent the fit of the data (the normal black lines correspond to the individual species included in the model and the bold lines to the fit of the overall model to the data). A single homogeneous species, with a Svedberg value indicating a monomeric protein was observed in the sample containing the protein alone. However the protein-heparin complex sample appears to contain three different species, which may match the monomeric, dimeric and higher order oligomeric states also observed in the native ESI-MS experiments.

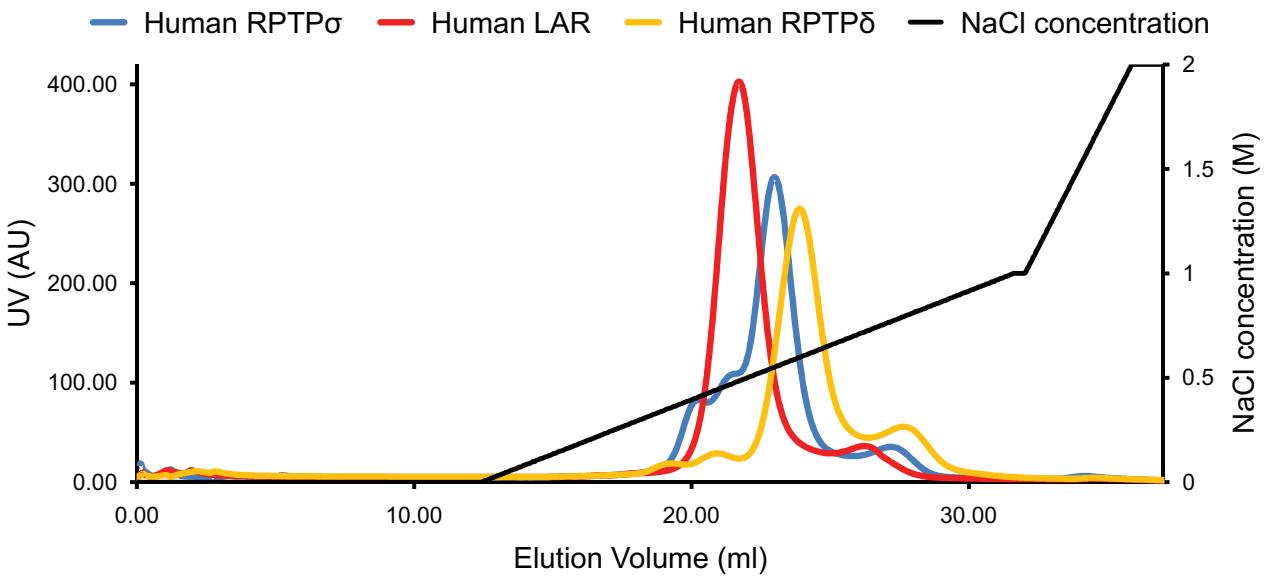


Figure S11. The ability of the type IIa RPTPs to bind to a heparin affinity column was measured. Purified Ig1-2 constructs of human RPTPσ (blue), RPTP LAR (red) and RPTPδ (yellow) were sequentially injected over the column and eluted upon addition of 550 mM, 490 mM and 600 mM sodium chloride respectively. All proteins were freshly purified by SEC, then desalted in 50 mM HEPES, 50 mM sodium chloride, pH 7.5 prior to injection onto a 1 ml heparin column, before elution with a gradient of 50 mM HEPES, 2M sodium chloride, pH 7.5 (black).

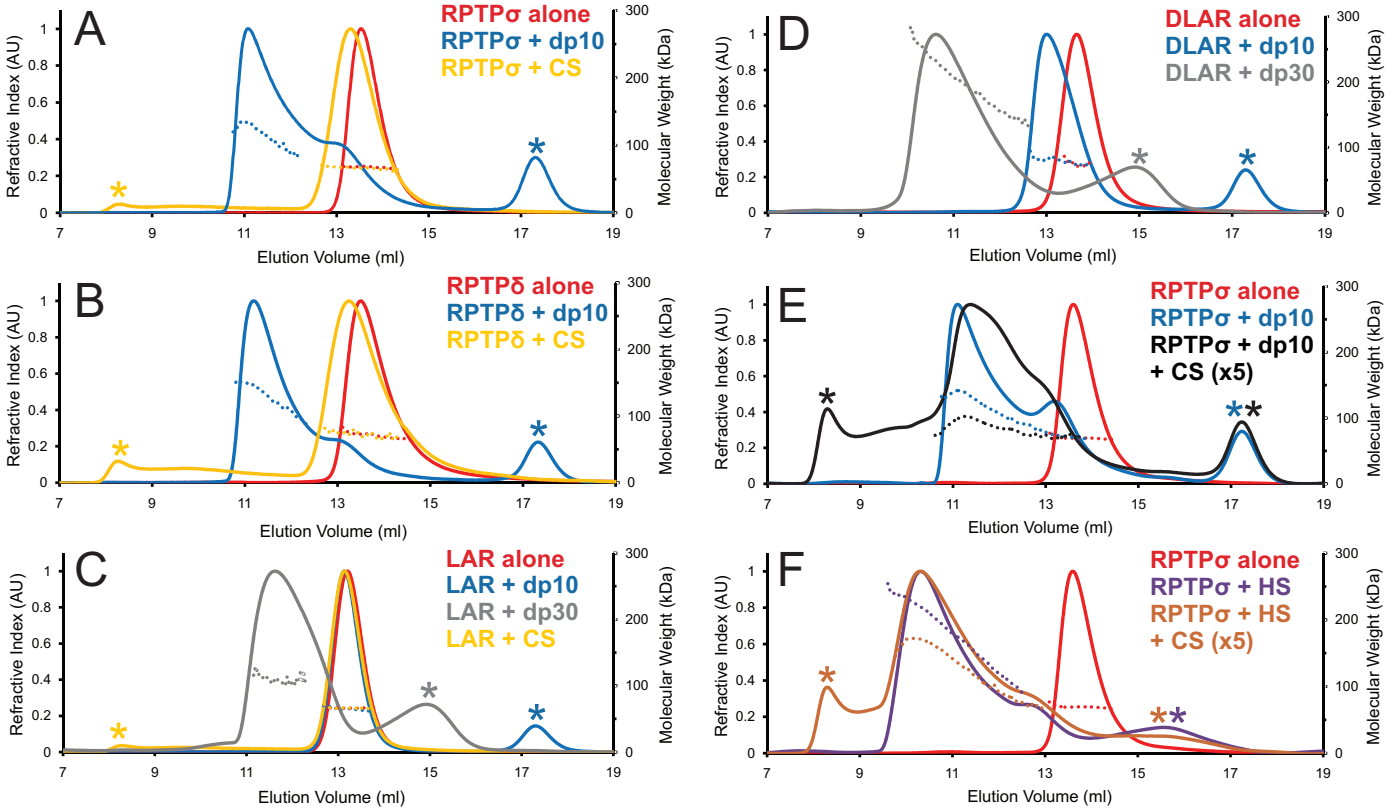


Figure S12. GAG-induced oligomerisation of the type IIa RPTP family. Alone (red), or after incubation with a five-fold molar excess of heparin dp10 (blue) or an equivalent amount of chondroitin sulphate (yellow), human RPTPσ Ig1-FN3 (**A**), RPTPδ Ig1-dFN3 (**B**) or LAR Ig1-dFN3 (**C**) were analysed by SEC-MALS. Similarly to RPTPσ, RPTPδ was observed to oligomerise upon addition of heparin dp10, however a longer oligosaccharide (dp30 in a five-fold molar excess, grey) was required to induce oligomerisation of LAR. In contrast to heparin, the introduction of chondroitin sulphate did not induce the oligomerisation of any of the human type IIa family members. Heparin promotes the oligomerisation of DLAR Ig1-FN3 (**D**), but this protein also requires a longer minimal heparin unit than RPTPσ. Incubating human RPTPσ Ig1-FN3 with a mixture of either heparin dp10 and a five-fold greater amount of chondroitin sulphate (**E**) or heparan sulphate and a five-fold greater amount of chondroitin sulphate (**F**) resulted in oligomers of reduced mass relative to the addition of heparin dp10 or heparan sulphate alone. Refractive index traces (scaled within each panel) are shown by bold lines and the measured molecular weights are shown by dotted lines. Peaks in the refractive index that are indicated by an asterisk correspond to excess oligosaccharide ligand. A Superdex 200 (1 cm x 30 cm) column was used for all constructs.

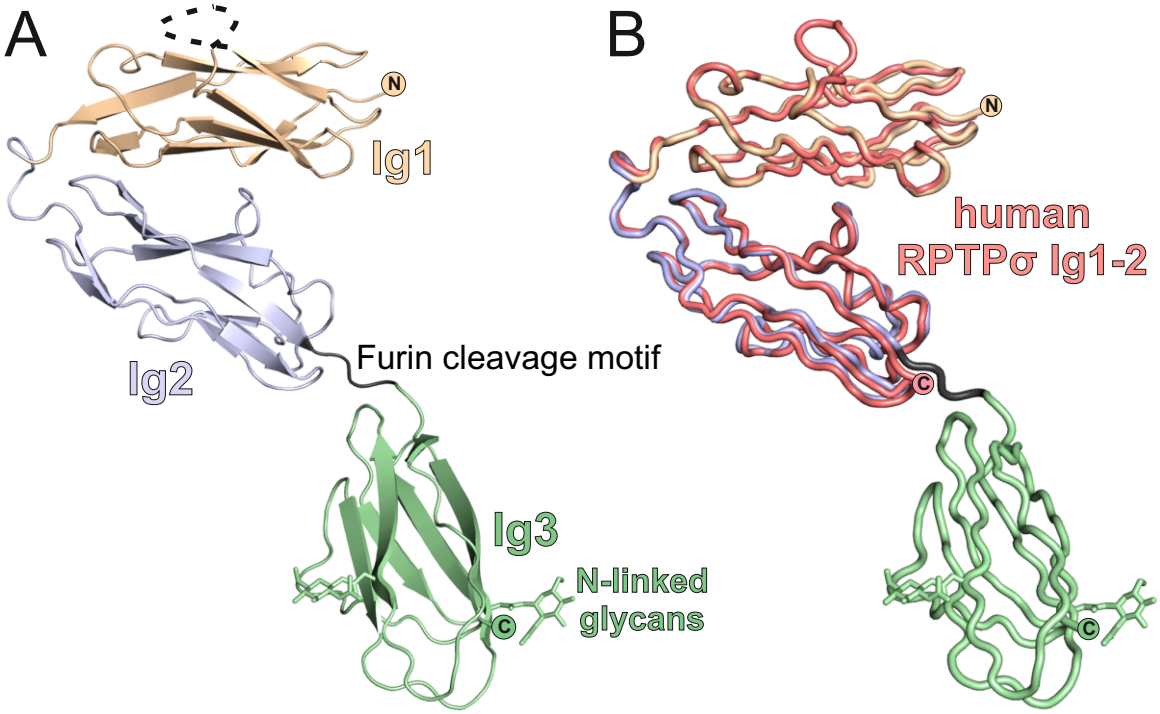


Figure S13. Crystal structure of human RPTPσ Ig1-3. (**A**) Ribbon representation of the three N-terminal Ig domains of a human RPTPσ R227Q+R228N double mutant. Only after introduction of these two point mutations to disrupt a potential furin cleavage motif RVRR, could crystals of intact Ig1-3 protein be obtained. The structure reveals that this motif is located in an exposed, flexible linker (black) between Ig domains 2 and 3. Residues 67-73 of the "Lys"-loop are not included in the structure as they were not well resolved in the electron density; they are indicated here by a dashed black line. (**B**) An overlay of human RPTPσ Ig1-2 (pink) and Ig1-3 illustrates the preserved architecture of the first two Ig domains in the Ig1-3 structure.

Removal of Ig1-2, which contains the glycosaminoglycan binding site, either at or prior to arrival at the cell surface, could represent an additional mode of regulation of the type IIa RPTPs. However, RPTP isoforms which contain the MeB exon (which encodes four amino acid residues, ELRE, and lies within the RVRR motif), would be unable to utilise such a regulatory mechanism. For example, an isoform of RPTPδ lacking MeB is found to be expressed in the kidney, while an isoform with the MeB insert is expressed in the brain (*S31, S32*), therefore isoform expression patterns would determine whether this proteolytic cleavage event occurs in specific tissues. Additional membrane proximal and non-isoform dependent protease cleavage sites have been identified for type IIa RPTPs (*S33*), which result in the shedding of almost the entire ectodomain and would also desensitise the receptors to proteoglycans.

A novel RPTPσ isoform lacking Ig3 has been identified in human ulcerative colitis patients (*S34*). The deletion of this domain may result in disruption of RPTPσ-ligand interactions, though as yet no binding partner for the Ig3 has been identified. Alternatively, it is possible that the resultant direct fusion of the rigid Ig1-2 unit onto the remainder of the ectodomain, limits the range of movement of the Ig1-2 unit, hampering efficient proteoglycan binding and receptor clustering.

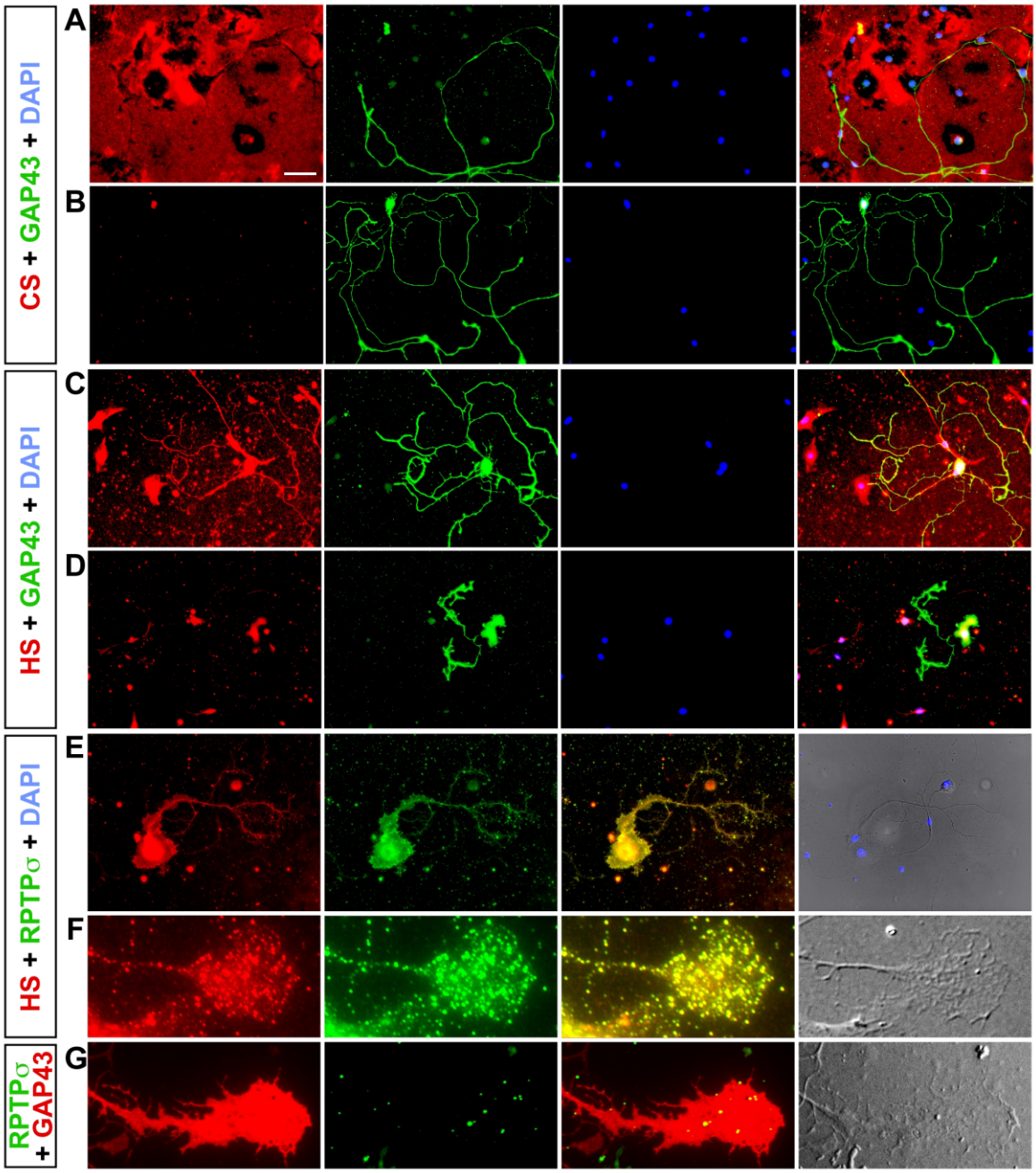


Figure S14. Immunolocalization of endogenous HS, CS and RPTPσ in DRG neuron cultures. Treatments are as described in Fig. 4 and (*27*), with additional controls shown here. (**A** and **B**) Immunolocalization of CS (red), with GAP43 neuronal marker (green) and DAPI nuclear stain (blue). Colors are merged in the right-hand panel. Panel A, shows strong CS labeling over the ECM, whereas areas corresponding to cells showed little or no CS labeling above background. Panel B shows a control treated with chondroitinase ABC to remove CS. (**C** and **D**) Immunolocalization of HS (red), with GAP43 (green) and DAPI (blue). Colors are merged in the right-hand panel. Panel C, shows strong HS labeling over cells including neurites, with additional faint labeling above background over the ECM. Panel D shows a control treated with combined heparinase I, II and III, to remove HS. (**E-G**) Immunolocalization of RPTPσ (green), combined with HS labeling (red) in panels E and F, or GAP43 neuronal marker (red) in panel G. The right-hand panels show DIC images. Panel G shows a negative control for RPTPσ labeling with secondary antibody only. RPTP and HS overlapped in similar although not identical patterns in cells, as well as in puncta within the growth cone. Some additional scattered dots were seen outside the cells and may be cell fragments based on DIC and DAPI images. A similar pattern, with punctate labeling in the growth cone and scattered dots outside the cell, was reported for RPTPσ previously (*35*). Scale bar: 60 μm (panels A-E), 6 μm (panels F and G).

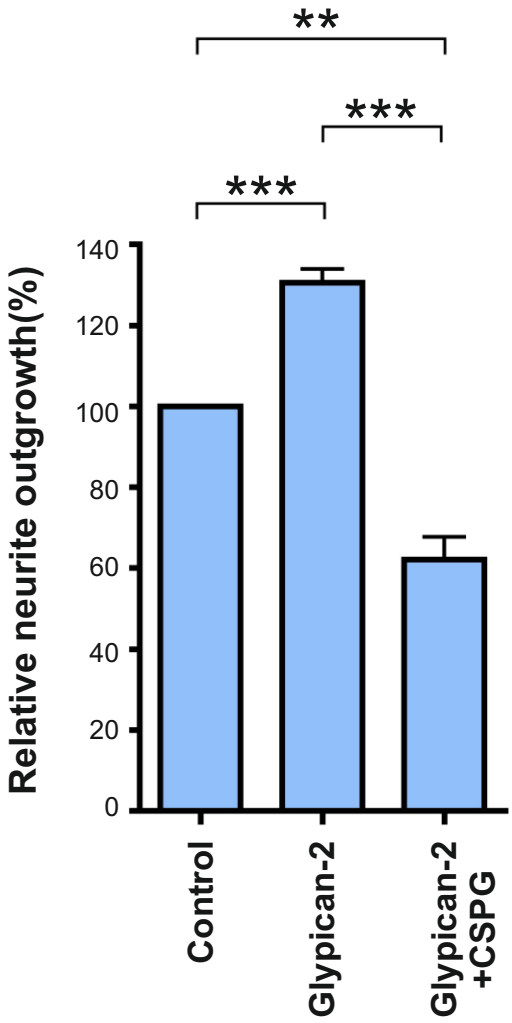


Figure S15. CSPG competition for HSPG-induced neurite outgrowth. Wild-type P10 mouse dorsal root ganglion (DRG) neurons were grown on substrates containing a poly-D-lysine/laminin mixture (control), or supplemented with either 5 μg/ml glypican-2, or with a mixture of 5 μg/ml glypican-2 plus 125 μg/ml CSPG. The outgrowth of DRG neurons, relative to control (assigned 100 % outgrowth) was quantified. Treatment with 5 μg/ml glypican-2 promoted outgrowth to a lesser extent than the higher concentration (30 μg/ml) of glypican-2 used in Figs. 1 and S1. CSPG out-competed the outgrowth promoting effect of glypican-2, leading to outgrowth inhibition relative to the control. Error bars show SEM. ***$p<0.001$, **$p<0.01$, Student's t-test.

Table S1. Crystallisation conditions and freezing methods

Construct	Space Group	Protein Concentration	Crystallisation Condition	Temp (°C)	Freezing Procedure
Chick RPTPσ Ig1-2 SeMet	$I4_122$	7 mg/ml	10 % w/v PEG 6k 0.1 M citrate pH 5	6.5	30 % PEG 400
Chick RPTPσ Ig1-2	$I4_122$	7 mg/ml	20 % w/v PEG 8k 0.1 M CHES pH 9.6	20.5	30 % PEG 400
Human RPTPσ Ig1-2 †	$I4_122$	11 mg/ml	0.2 M sodium sulphate 20 % w/v PEG 3350 pH 6.6	20.5	Perfluoropolyether oil PFO-X125/03
Human RPTPσ Ig1-2	C2	10.5 mg/ml	0.2 M NH_4I 20 % w/v PEG 3350 pH 6.2	20.5	Perfluoropolyether oil PFO-X125/03
Human RPTPδ Ig1-2	$P3_221$	6 mg/ml	0.085 M tri-sodium citrate dihydrate 25% PEG 4k 15% glycerol pH5.6	20.5	Extra 15 % glycerol
Human RPTPδ Ig1-2	C2	6 mg/ml	20 % w/v PEG 3350 0.2 M ammonium di-hydrogen phosphate	20.5	30 % ethylene glycol
Human RPTP LAR Ig1-2	$P3_221$	9 mg/ml	0.1 M sodium acetate trihydrate 20 % isopropanol 20 % w/v PEG 4k 0.25 mM heparin dp10 pH 5.6	20.5	25 % ethylene glycol
Drosophila RPTP LAR Ig1-2	C2	12mg/ml	25 % w/v PEG 1500 0.1 M SPG System 35 mM ammonium sulphate pH 5.0	20.5	25 % glycerol
Human RPTPσ Ig1-3 R227Q+R228N	I222	6.7 mg/ml	20% w/v PEG 3350 0.2 M sodium iodide 0.1 M bis-Tris propane pH 7.5	20.5	30 % glycerol
Human RPTP LAR Ig1-2 + 25 mM sucrose octasulphate	$P3_221$	9 mg/ml	0.1 M MES 30 % w/v PEG 6k pH 6.0	20.5	30 % propylene glycol

† Crystallisation trials were set up with human RPTPσ Ig1-3 wildtype protein. Crystals were obtained within 80 days. Subsequent structure determination showed clear density for Ig1-2, but no space within the crystal lattice for a third Ig domain. This suggested a proteolytic cleavage of the protein within the flexible linker between Ig domains two and three, which contains a putative ^{225}RVRR^{228} furin motif. To obtain crystals of human RPTPσ Ig1-3 a double point mutation R227Q+R228N was necessary to disrupt this motif.

Table S2. Crystallographic data collection and refinement statistics

Data collection										
Protein	Chick RPTPσ Ig1-2 SeMet	Chick RPTPσ Ig1-2	Human RPTPσ Ig1-2	Human RPTPσ Ig1-2	Human RPTPδ Ig1-2	Human RPTPδ Ig1-2	Human RPTP LAR Ig1-2	*Drosophila* RPTP LAR Ig1-2	Human RPTP LAR Ig1-2 + SOS	Human RPTPσ Ig1-3
PDB accession code	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Beamline	ESRF-BM14	ESRF-BM14	Diamond-I03	ESRF-ID14EH3	Diamond-I02	Diamond-I02	Diamond-I03	ESRF-ID23EH2	Diamond-I04	Diamond-I04
Resolution (Å)	50.0-3.1 (3.2-3.1)	74-1.65 (1.7-1.65)	40-2.3 (2.38-2.3)	29-2.55 (2.6-2.55)	50-2.0 (2.06-2.0)	49-1.35 (1.4-1.35)	48-2.2 (2.26-2.2)	50-1.8 (1.85-1.8)	40-2.05 (2.12-2.05)	50-2.6 (2.7-2.6)
Spacegroup	$I4_122$	$I4_122$	$I4_122$	C2	C2	$P3_221$	$P3_221$	C2	$P3_221$	I222
Cell dimensions *a* , *b* , *c* (Å) α, β, γ (°)	 104.9, 104.9, 92.0 90.0, 90.0, 90.0	 104.8, 104.8, 94.6 90.0, 90.0, 90.0	 104.2, 104.2, 94.9 90.0, 90.0, 90.0	 124.2, 29.4, 61.4 90.0, 92.6, 90.0	 161.2, 31.7, 92.3 90.0, 112.0, 90.0	 48.7, 48.7, 145.9 90.0, 90.0 , 120.0	 77.0, 77.0, 68.7 90.0, 90.0, 120.0	 124.2, 29.1, 49.4 90.0, 107.3, 90.0	 78.8, 78.8, 71.9 90.0, 90.0, 120.0	 71.7, 90.0, 143.3 90.0, 90.0, 90.0
Wavelength (Å)	0.9783	0.9763	1.0600	0.9310	0.9795	0.9795	0.9765	0.8726	0.9763	0.9786
Unique reflections	4922 (475)	31882 (2319)	12003 (1176)	7438 (474)	30633 (2991)	45165 (3237)	12272 (872)	16271 (1610)	16614 (1638)	14979 (1469)
Completeness (%)	100 (100)	100 (100)	100 (99.9)	98.1 (84.8)	99.8 (99.8)	99.9 (99.8)	99.7 (98.2)	100 (100)	100 (100)	99.9 (99.9)
R_{merge} (%)[b]	14.9 (86.2)	11.8 (88.5)	13.7 (66.5)	8.2 (66.6)	7.4 (60.3)	11.4 (81.6)	9.2 (83.1)	8.1 (53.3)	10.2 (94.3)	13.0 (87.0)
I/σI	33.6 (5.6)	26.2 (4.2)	13.6 (3.7)	13.7 (2.6)	16.8 (2.1)	17.4 (3.1)	19.1 (3.6)	19.6 (3.1)	24.7 (2.7)	14.6 (2.3)
Redundancy	35.7 (36.9)	26.0 (18.6)	18.3 (18.3)	3.5 (3.1)	3.6 (3.6)	15.0 (10.4)	10.7 (10.6)	5.6 (5.4)	9.9 (10.0)	7.2 (7.4)
Resolution range (Å)		74-1.65	40-2.3	29-2.55	50-2.0	49-1.35	48-2.2	50-1.8	40-2.05	50-2.6
Number of reflections		30925	11897	7196	28407	42829	11593	16006	15755	13846
R_{work} (%)[c]		17.0	19.9	20.2	21.8	13.7	19.3	18.1	19.5	24.8
R_{free} (%)[d]		20.4	23.3	26.0	24.8	17.0	23.1	22.2	22.7	28.9
Number of protein atoms		1677	1546	1524	3079	1657	1518	1505	1506	2119
Number of chloride ions		1	1	3	-	1	-	-	-	4
Number of sodium ions		-	1	-	-	-	-	-	-	-
Number of sulphate ions		1	-	-	-	-	-	-	-	-
Number of phosphate ions		-	-	-	3	-	-	-	-	-
Number of iodide ions		-	-	4	-	-	-	-	-	4
Number of PEG molecules		3	-	-	-	-	-	-	-	-
Number of citrate molecules		-	-	-	-	1	-	-	-	-
Number of B3P molecules		-	-	-	-	-	-	-	-	1
Number of SOS atoms		-	-	-	-	-	-	-	28	-
Number of NAG molecules		-	-	-	-	-	-	-	-	2
Number of glycine molecules		-	-	-	-	-	-	1	-	-
Number of water molecules		273	81	29	139	259	58	115	97	21
r.m.s.d. bonds(Å)		0.009	0.007	0.005	0.009	0.014	0.004	0.007	0.014	0.006
r.m.s.d. angles (deg)		1.296	1.006	0.831	1.43	1.660	0.848	1.138	1.237	0.955
Average B factor (Å^2):										
- protein		20.2	38.3	30.6	43.8	13.7	40.4	31.1	43.9	69.0
- chloride ions		12.8	31.9	19.2	-	16.4	-	-	-	53.3
- sodium ions		-	25.4	-	-	-	-	-	-	-
- sulphate ions		59.4	-	-	-	-	-	-	-	-
- phosphate ions		-	-	-	68.3	-	-	-	-	-
- iodide ions		-	-	61.1	-	-	-	-	-	45.3
- PEG molecules		40.1	-	-	-	-	-	-	-	-
- citrate molecules		-	-	-	-	24.3	-	-	-	-
- B3P molecules		-	-	-	-	-	-	-	-	51.5
- SOS molecules		-	-	-	-	-	-	-	147.3	-
- NAG molecules		-	-	-	-	-	-	-	-	76.0
- Glycine molecules		-	-	-	-	-	-	62.1	-	-
- water molecules		29.7	37.5	22.3	39.5	25.3	35.3	35.6	44.0	42.2
Ramachandran plot (% residues) :										
- favoured		98.5	98.0	94.9	97.1	99.0	97.5	98.5	98.0	96.0
- additionally allowed		1.5	1.5	5.1	2.9	1.0	2.5	1.5	2.0	4.0
- disallowed		0	0.5	0	0	0	0	0	0	0

[a]Numbers in parentheses refer to the appropriate outer shell.

[b]$R_{merge} = \Sigma hkl \Sigma j \, |Ij - \langle Ij \rangle| \, / \, \Sigma hkl \Sigma j \, |Ij|$, where hkl specifies unique indices, *j* indicates equivalent observations of hkl, and $\langle Ij \rangle$ is the mean value.

[c]$R_{work} = \Sigma hkl \, || Fo| - |Fc| | \, / \, \Sigma hkl \, |Fo|$, where $|Fo|$ and $|Fc|$ are the observed and calculated structure factor amplitudes, respectively.

[d]R_{free} is calculated as per R_{work} for a 5% subset of reflections which was not used in the crystallographic refinement.

Table S3. Summary of MALS experiments. Heparin experiments in blue or white were performed with different batches of heparin oligomers and tested at 1:2 and 1:5 protein:heparin ratios respectively. The approximate molecular weight of heparin dp10 is ~3.3 kDa, dp20 is ~6.6 kDa and dp30 ~10 kDa. CS (Sigma, C4384) and HS (Iduron, GAG-HS01) mixtures used in these experiments contain GAG fragments 10-50 kDa in mass. All Ig1-FN3 proteins were injected at ~10 μM. The molecular weight values listed in this table (MALS MW) were measured at the peak in the refractive index trace for a given experiment. The error in the measured molecular weights is the uncertainty statistic provided by the Astra software (Wyatt Technologies).

Protein and construct name	GAG added	Protein concentration at peak	Number of N-glyc sites	Theoretical MW (+ N-glyc)	MALS MW (± error)	Dominant oligomeric state
		μg/ml		kDa		
RPTPσ						
sEcto	none	151	2	91.0 (93.7)	107.2 (1.1)	monomer
sEcto	none	709	2	91.0 (93.7)	128.5 (1.0)	monomer
sEcto	heparin-dp10	147	2	91.0 (93.7)	318 (2.4)	di/trimer
Ig1-2	none	211	0	23.7 (23.7)	27.2 (1.4)	monomer
Ig1-2	heparin-dp10	383	0	23.7 (23.7)	69.8 (0.9)	di/trimer
Ig1-3	none	589	2	34.9 (37.6)	44.2 (0.2)	monomer
Ig1-3	heparin-dp10	434	2	34.9 (37.6)	84.7 (0.6)	dimer
Ig1-FN3	none	173	2	64.3 (67.0)	75.3 (0.6)	monomer
Ig1-FN3	heparin-dp10	156	2	64.3 (67.0)	211.6 (1.9)	di/trimer
Ig1-FN3	none	69	2	64.3 (67.0)	68.7 (0.8)	monomer
Ig1-FN3	heparin-dp4	62	2	64.3 (67.0)	73.9 (0.9)	monomer
Ig1-FN3	heparin-dp6	55	2	64.3 (67.0)	77.4 (0.8)	monomer
Ig1-FN3	heparin-dp8	26	2	64.3 (67.0)	113.5 (0.7)	dimer
Ig1-FN3	heparin-dp10	31	2	64.3 (67.0)	119.8 (1.7)	dimer
Ig1-FN3	heparin-dp20	38	2	64.3 (67.0)	207 (1.3)	trimer
Ig1-FN3	heparin-dp30	41	2	64.3 (67.0)	275.7 (2.8)	tetramer
Ig1-FN3	HS	39	2	64.3 (67.0)	219.0 (1.3)	trimer
Ig1-FN3	HS (x5)	38	2	64.3 (67.0)	242.3 (1.5)	tri/tetramer
Ig1-FN3	CS	52	2	64.3 (67.0)	67.7 (1.4)	monomer
Ig1-FN3	CS (x5)	29	2	64.3 (67.0)	80.8 (1.4)	monomer
Ig1-FN3	dp10 + CS (x5)	36	2	64.3 (67.0)	101.8 (1.8)	mono/dimer
Ig1-FN3	HS + CS (x5)	44	2	64.3 (67.0)	171.3 (1.2)	di/trimer
Ig1-FN3 ΔK	none	83	2	64.3 (67.0)	66 (0.3)	monomer
Ig1-FN3 ΔK	heparin-dp10	81	2	64.3 (67.0)	66.5 (0.7)	monomer
Ig1-FN3 ΔK	heparin-dp30	77	2	64.3 (67.0)	70.6 (2.1)	monomer
RPTPδ						
Ig1-FN3	none	52	2	64.1 (66.8)	72 (2.3)	monomer
Ig1-FN3	heparin-dp10	48	2	64.1 (66.8)	147.4 (1.7)	dimer
Ig1-FN3	CS	27	2	64.1 (66.8)	72.6 (3.6)	monomer
LAR						
Ig1-FN3	none	86	2	63.8 (66.5)	66.2 (0.8)	monomer
Ig1-FN3	heparin-dp10	85	2	63.8 (66.5)	66.1 (0.4)	monomer
Ig1-FN3	heparin-dp30	45	2	63.8 (66.5)	107.2 (4.2)	mono/dimer
Ig1-FN3	CS	85	2	63.8 (66.5)	65.9 (0.9)	monomer
DLAR						
Ig1-FN3	none	73	3	64.5 (68.5)	71.3 (2.9)	monomer
Ig1-FN3	heparin-dp10	61	3	64.5 (68.5)	82.3 (1.9)	monomer
Ig1-FN3	heparin-dp20	42	3	64.5 (68.5)	171.6 (10.1)	di/trimer
Ig1-FN3	heparin-dp30	48	3	64.5 (68.5)	234 (2.7)	trimer

Supporting references

S1. A. R. Aricescu, W. Lu, E. Y. Jones, *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1243 (2006).

S2. Y. Shen *et al.*, *Science* **326**, 592 (2009).

S3. A. D. Elbein, J. E. Tropea, M. Mitchell, G. P. Kaushal, *J. Biol. Chem.* **265**, 15599 (1990).

S4. V. T. Chang *et al.*, *Structure* **15**, 267 (2007).

S5. P. J. Reeves, N. Callewaert, R. Contreras, H. G. Khorana, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13419 (2002).

S6. F. Grueninger-Leitch, A. D'Arcy, B. D'Arcy, C. Chène, *Protein Sci.* **5**, 2617 (1996).

S7. T. S. Walter *et al.*, *Acta Crystallogr. D Biol. Crystallogr.* **61**, 651 (2005).

S8. C. J. Mayo *et al.*, *Structure* **13**, 175 (2005).

S9. Z. Otwinowski, Minor, Wladek, *Methods Enzymol.*, 307 (1997).

S10. G. Winter, *J. Appl. Crystallogr.* **43**, 186 (2010).

S11. G. Sheldrick, *Acta Crystallogr. A* **64**, 112 (2008).

S12. E. B. C. Vonrhein, P. Roversi, G. Bricogne, *Methods Mol. Biol.*, 215 (2007).

S13. T. Terwilliger, *Acta Crystallogr. D Biol. Crystallogr.* **59**, 45 (2003).

S14. A. Perrakis, R. Morris, V. S. Lamzin, *Nat. Struct. Mol. Biol.* **6**, 458 (1999).

S15. P. Emsley, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126 (2004).

S16. A. J. McCoy, R. W. Grosse-Kunstleve, L. C. Storoni, R. J. Read, *Acta Crystallogr. D Biol. Crystallogr.* **61**, 458 (2005).

S17. G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240 (1997).

S18. P. D. Adams *et al.*, *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213 (2010).

S19. G. Bricogne, *Acta Crystallogr. D Biol. Crystallogr.* **49**, 37 (1993).

S20. R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **26**, 283 (1993).

S21. I. W. Davis *et al.*, *Nucl. Acids Res.* **35**, W375 (2007).

S22. D. I. Stuart, M. Levine, H. Muirhead, D. K. Stammers, *J. Mol. Biol.* **134**, 109 (1979).

S23. W. L. DeLano, *DeLano Scientific LLC, Palo Alto, CA, USA.* (2008).

S24. N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10037 (2001).

S25. W. F. Stafford, E. H. Braswell, *Biophys Chem.* **108**, 273 (2004).

S26. K. G. Johnson *et al.*, *Neuron* **49**, 517 (2006).

S27. J. G. Flanagan, H.-J. Cheng, *Methods Enzymol.* **327**, 198 (2000).

S28. J. G. Flanagan *et al.*, *Methods Enzymol.* **327**, 19 (2000).

S29. W. Kabsch, C. Sander, *Biopolymers* **22**, 2577 (1983).

S30. M. C. Lawrence, P. M. Colman, *J. Mol. Biol.* **234**, 946 (1993).

S31. R. Pulido, N. X. Krueger, C. Serra-Pagès, H. Saito, M. Streuli, *J. Biol. Chem.* **270**, 6722 (1995).

S32. R. Pulido, C. Serra-Pagès, M. Tang, M. Streuli, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11686 (1995).

S33. B. Aicher, M. M. Lerch, T. Muller, J. Schilling, A. Ullrich, *J. Cell Biol.* **138**, 681 (1997).

S34. A. M. Muise *et al.*, *Curr. Biol.* **17**, 1212 (2007).

S35. K. M. Thompson *et al.*, *Mol. Cell. Neuro.* **23**, 681 (2003).