doi:10.1016/j.jmb.2011.03.013 J. Mol. Biol. (2011) 408, 616–627

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

COMMUNICATION

The Immunoglobulin-like Domains 1 and 2 of the Protein Tyrosine Phosphatase LAR Adopt an Unusual Horseshoe-like Conformation

Bridget H. Biersmith[1], Michal Hammel[2], Erika R. Geisbrecht[1] and Samuel Bouyain[3]*

[1]*Division of Cell Biology and Biophysics, School of Biological Sciences, University of Missouri-Kansas City, 5100 Rockhill Road, Kansas City, MO 64110, USA*
[2]*Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[3]*Division of Molecular Biology and Biochemistry, School of Biological Sciences, University of Missouri-Kansas City, 5100 Rockhill Road, Kansas City, MO 64110, USA*

Received 17 December 2010; received in revised form 3 March 2011; accepted 4 March 2011
Available online 12 March 2011

Edited by I. Wilson

Keywords:
cell adhesion;
receptor protein tyrosine phosphatase;
heparan sulfate proteoglycans;
immunoglobulin-like domains;
crystal structure

Neurogenesis depends on exquisitely regulated interactions between macromolecules on the cell surface and in the extracellular matrix. In particular, interactions between proteoglycans and members of the type IIa subgroup of receptor protein tyrosine phosphatases underlie crucial developmental processes such as the formation of synapses at the neuromuscular junction and the migration of axons to their appropriate targets. We report the crystal structures of the first and second immunoglobulin-like domains of the Drosophila type IIa receptor Dlar and its mouse homolog LAR. These two domains adopt an unusual antiparallel arrangement that has not been reported in tandem repeats of immunoglobulin-like domains and that is presumably conserved in all type IIa receptor protein tyrosine phosphatases.

Receptor protein tyrosine phosphatases (RPTPs) are a family of cell surface receptors involved in the growth and development of the nervous system from worms to humans.[1,2] These molecules are typically composed of intracellular tyrosine phosphatase domains that antagonize tyrosine kinase signaling and large, modular extracellular domains that resemble those found in the ectodomains of cell adhesion molecules. The particular architecture of RPTPs is akin to that of receptor tyrosine kinases, which suggests that the binding of extracellular ligands can control the intracellular phosphatase activity. However, most RPTPs remain orphan receptors and a unifying mechanism of how ligand binding is transduced into intracellular tyrosine dephosphorylation, if it exists, is still lacking.

Studies carried out in Drosophila were the first to uncover the involvement of RPTPs in neurogenesis.[3,4] In particular, it was demonstrated that the motor axons of flies lacking the leukocyte-antigen related receptor

*Corresponding author. E-mail address: bouyains@umkc.edu.

Abbreviations used: Ig, immunoglobulin-like domain; FNIII, fibronectin type III domain; HS, heparan sulfate; HSPG, heparan sulfate proteoglycan; hGH, human growth hormone; SAXS, small angle X-ray scattering; Dscam, Down syndrome cell adhesion molecule.

(Dlar) were unable to reach their appropriate targets.[3] These pathfinding defects were later linked to aberrant signaling through the tyrosine kinase Abl, which is normally antagonized by the intracellular tyrosine phosphatase activity of Dlar.[5] Furthermore, it was shown that the formation of synapses at the neuromuscular junction depended on interactions between ligands expressed at the surface of muscle cells and the N-terminus of Dlar expressed on nerve cells.[6] In contrast, the appropriate targeting of photoreceptor neuron R7 axons in the Drosophila visual system does not depend on the N-terminus of Dlar or its phosphatase activity, but instead requires the presence of domains within the membrane-proximal region of the ectodomain of Dlar.[7] Consequently, these results indicate that some of the characterized physiological functions of Dlar are mediated by distinct modules in its extracellular region.

Dlar and its three vertebrate homologs LAR/PTPRF, RPTPδ/PTPRD and RPTPσ/PTPRS share a similar modular architecture and belong to the type IIa subgroup of RPTPs. These type I transmembrane proteins are composed of three immunoglobulin domains (Ig) and eight or nine fibronectin type III (FNIII) repeats followed by a transmembrane region and tandem cytoplasmic tyrosine phosphatase domains (Fig. 1a).[2,8] Mice lacking either of the vertebrate homologs of Dlar display neurological abnormalities such as learning disabilities (LAR and PTPRD[9,10]), posture/motor defects (PTPRD[10]) or spastic movement, tremors and decreased brain size (PTPRS[11,12]). In addition, mice deficient for both PTPRD and PTPRS suffer from severe motoneuron defects and die soon after birth.[13] Taken together, these findings suggest that Dlar and its vertebrate homologs LAR, PTPRD and PTPRF play crucial roles in neurogenesis across species.

In addition to their similar architecture and their roles in neural development, type IIa RPTPs interact with related extracellular ligands. In Drosophila, Dlar

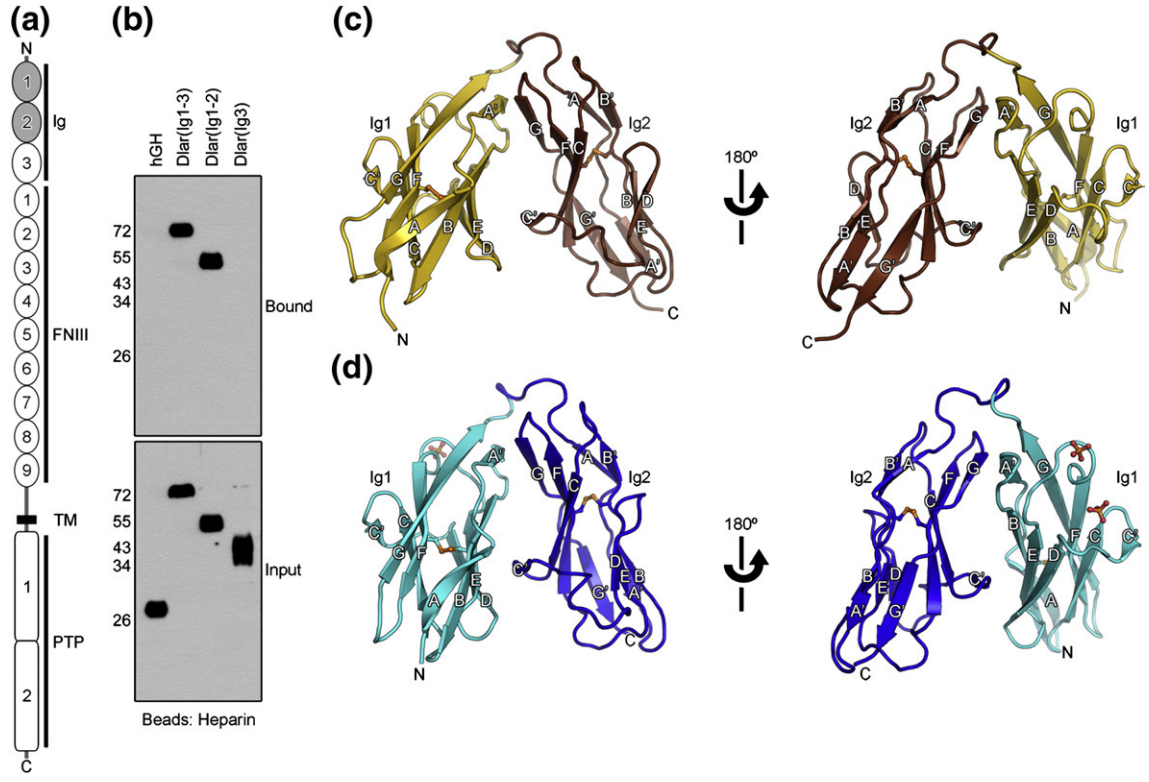

Fig. 1. Crystal structures of the Ig1-Ig2 domains of Dlar and mouse LAR. (a) A representation of type IIa RPTPs. Ig, immunoglobulin-like domains; FNIII, fibronectin type III domains; TM, transmembrane region; PTP, protein tyrosine phosphatase domains. The Ig domains that were crystallized are shaded. (b) Identification of a heparin-binding region within the Ig domains of Dlar. Fragments of Dlar were fused to human growth hormone (hGH) and expressed transiently in HEK293 cells. Conditioned media were incubated with Heparin Sepharose™. Bound fusion proteins were visualized by immunoblotting against hGH. (c) A ribbon diagram of Dlar(Ig1-2). The letters N and C indicate the N- and C-termini, respectively. Each β-strand is labeled. Disulfide bonds are shown as orange ball-and-stick models. The first and second Ig domains are colored gold and brown, respectively. (d) A ribbon diagram of mouse LAR(Ig1-2). The first and second Ig domains are colored cyan and blue, respectively. Two bound sulfate ions are shown in ball-and-stick representation. Oxygen and sulfur atoms are colored red and orange, respectively. Structural images were prepared with PyMOL (www.pymol.org).

interacts *in vivo* with the heparan sulfate proteoglycans (HSPGs) syndecan and dally-like protein during the formation of synapses at the Drosophila neuromuscular junction.[6,14] HSPGs are composed of a protein core and are decorated with heparan sulfate (HS) chains, which are polymers of negatively charged disaccharide units. The interactions between Dlar and the HSPGs syndecan and dally-like protein have dissociation constants within the nanomolar range and are sensitive to treatment with the HS-degrading enzyme heparinase, indicating that the HS chains present on these molecules play a crucial role in mediating these interactions.[6] In addition, a form of Dlar lacking the first three Ig domains was unable to bind to HSPGs and could not rescue the synaptic defects observed in *dlar*$^{-/-}$ flies, suggesting that the HSPG-binding site localizes to this region and that these protein–carbohydrate binding events are crucial for fly motoneuron development. In vertebrates, the avian PTPRS also binds with high affinity to the HSPGs agrin and collagen XVIII.[15] Dissociation constants in the nanomolar range were measured and were nearly indistinguishable from those determined for the binding of the artificial ligand heparin–albumin to PTPRS. These combined studies have highlighted the importance of the interactions between proteoglycans and the extracellular regions of type IIa RPTPs during the growth and development of the nervous system. As a first step towards providing a molecular basis for these protein–carbohydrate interactions, we undertook structural studies of the first two Ig domains of Dlar and its mouse homolog LAR.

Crystal structures of Ig1-Ig2 repeats of Dlar and mouse LAR

We initially pursued the crystallization of the entire Ig region of Dlar for subsequent structural analysis. The protein was purified from the conditioned medium of transiently transfected N-acetylglucosaminyltransferase I-negative HEK293S cells, followed by removal of the N-linked oligosaccharides using endoglycosidase H.[16] However, attempts to grow crystals of Dlar(Ig1-3) were unsuccessful, mostly because of the instability of this protein fragment. Incubation of the protein at 4 °C for several days resulted in a reduction of molecular mass of about ~10 kDa as judged by SDS-PAGE. Subsequent analysis by mass spectrometry indicated that the stable form of Dlar included residues 32–237, which span Ig domains 1 and 2. These serendipitous findings resulted in the identification of a discrete form of Dlar that was more amenable to crystallographic analysis because of its improved stability. Earlier, mutational analysis of PTPRS demonstrated that a cluster of basic residues in domain Ig1 are crucial for binding to HSPGs;[15] therefore, we predicted that the absence of Ig3 would not impair the ability of Dlar to bind to HS chains. To test this hypothesis, we incubated secreted fragments of Dlar fused to human growth hormone (hGH) expressed in HEK293 cells with Heparin Sepharose resin (Fig. 1b). In this assay, the third Ig domain of Dlar does not appear to have any intrinsic heparin-binding property and its absence does not impair the ability of Dlar to associate with heparin; therefore, we focused our efforts on determining the crystal structure of the first two Ig repeats of Dlar.

Instead of producing Dlar(Ig1-2) in HEK293S cells, we opted for a more expeditious approach and expressed this fragment in *Escherichia coli* strain Origami2(DE3) to promote the formation of disulfide bonds within each Ig domain. A similar strategy was used to produce the Ig1-Ig2 fragment of mouse LAR. Both bacterially expressed Dlar (Ig1-2) and mouse LAR(Ig1-2) bind to heparin agarose in an affinity isolation assay (Supplementary Data Fig. S1), indicating that these fragments retain the heparin-binding activity observed for Dlar expressed in mammalian cells. Small plate-like crystals of Dlar(Ig1-2) appeared after two rounds of microseeding, whereas crystals of LAR (Ig1-2) grew readily. First, the crystal structure of LAR(Ig1-2) was determined by molecular replacement and refined to 2.0 Å resolution (R_{work} 22.0 % and R_{free} 24.0 %) (Table 1). The structures of Ig1 and Ig2 of LAR were then used as independent search models to calculate initial phases for Dlar (Ig1-2) and the final model was refined to 2.3 Å resolution (R_{work} 22.3 % and R_{free} 26.1 %).

Overall, domains Ig1 and Ig2 of Dlar adopt an I set topology[17] and superimpose with rmsd 1.6 Å over 92 residues. Ig1 is most similar to the A168 Ig domain of titin (rmsd 1.2 Å over 95 residues, 29% identity) whereas the closest structural neighbor of Ig2 is the second Ig domain of muscle-specific kinase (rmsd 2.4 Å over 97 residues, 38% identity). Strikingly, Ig1 and Ig2 fold back onto one another to adopt a constrained conformation in which the individual Ig domains are arranged in an antiparallel manner (Fig. 1c and d). This horseshoe-like conformation is in contrast to the extended arrangement observed for titin and muscle-specific kinase or in other tandem Ig domains that harbor glycosaminoglycan-binding activity, such as Robo and fibroblast growth factor receptors.[18–20]

In essence, Dlar(Ig1-2) resembles a triangle of dimensions of ~45 Å × 41 Å × 49 Å as measured from the N-terminus to the middle of the linker region, from the middle of the linker region to the C-terminus and from the C-terminus to the N-terminus. The four molecules in the asymmetric unit of Dlar(Ig1-2) all adopt this conformation and superimpose with rmsd 0.4–0.5 Å for 187–199 equivalent C^{α} positions, indicating that this arrangement of Ig1 and Ig2 in Dlar is rigid. Mouse

Table 1. Data collection and refinement statistics

Beamline	APS 22-ID	APS 22-ID
Crystal	Dlar(Ig1-2)	LAR(Ig1-2)
Wavelength (Å)	0.97242	0.97242
Number of unique reflections	39,287	15,225
Resolution (Å)	50 – 2.3	50 – 2.0
Space group	$P2_1$	$P3_221$
Unit cell		
a, b, c (Å)	72.96, 77.51, 81.73	77.45, 77.45, 68.50
α, β, γ (°)	90.00, 101.08, 90.00	90.00, 90.00, 120.00
R^a_{sym}	0.117 (0.415)	0.069 (0.364)
Completeness (%)	98.7 (98.2)	92.6 (65.4)
Redundancy	2.7 (2.0)	15.4 (8.7)
I/σI	5.7 (1.9)	28.2 (3.6)
Refinement		
RCSB accession number	3PXJ	3PXH
Molecules in asymmetric unit	4	1
Resolution (Å)	41.4 – 2.3	33.5 – 2.0
Reflections	37,084	14,492
R^c_{work} / R_{free}	0.223 / 0.261	0.220 / 0.240
Number of protein atoms	5,965	1,494
Number of water atoms	138	85
r.m.s. deviation from ideal bonds (Å)	0.032	0.030
r.m.s. deviation from ideal angles (°)	1.66	1.67
Average B factors (Å^2)	58.2	53.4
Protein	58.4	53.6
Water	51.2	49.0
Sulfate	-	69.6
Ramachandran statistics		
Favored (%)	97.9	99.5
Allowed (%)	2.0	0.5
Outlier (%)	0.1	-

Methods: cDNA fragments encoding Dlar(Ig1-2) (residues 32–237) and mouse LAR(Ig1-2) (residues 29–235) were amplified from a Drosophila cDNA library and a mouse embryonic day 13.5 cDNA library, respectively. These fragments were subcloned into a modified pET32 plasmid (Novagen, La Jolla, CA) that directs the expression of thioredoxin, a His_6 tag, a human rhinovirus 3C protease cleavage site and the protein of interest. Proteins were expressed in *Escherichia coli* strain Origami 2(DE3) and purified by immobilized nickel-affinity chromatography, ion-exchange chromatography, heparin-affinity chromatography and size-exclusion chromatography as described.[16] Crystals were grown at 20 °C by the hanging-drop, vapor-diffusion method. Plate-like crystals of Dlar(Ig1-2) were grown in 50 mM imidazole–HCl pH 7.0, 10 mM $(NH_4)_2SO_4$ and 20% (w/v) PEG2000 MME after two consecutive rounds of microseeding. Crystals of LAR(Ig1-2) were grown in 50 mM sodium citrate pH 5.5, 200 mM Li_2SO_4 and 25% (w/v) PEG1500. For cryoprotection, crystals were transferred to mother liquor containing 15% – 20% (v/v) glycerol. Diffraction data were collected at beamline 22-ID of the Advanced Photon Source of Argonne National Laboratory and processed with HKL2000.[37] The structure of LAR(Ig1-2) was solved by molecular replacement with PHASER as implemented in PHENIX using the structures of telokin (PDB ID code IFHG, residues 40–123) and MuSK (PDB ID code 2IEP, residues 129–210) as two independent search models for the first and second Ig domains of LAR(Ig1-2), respectively. The final model for LAR(Ig1-2) was obtained after manual model building using COOT and refinement in PHENIX and consists of residues 30–226, 85 water molecules and two sulfate ions.[38,39] The Ig domains from LAR(Ig1-2) were then used as two independent search models to obtain a molecular replacement solution for Dlar(Ig1-2). The final model for Dlar(Ig1-2) consists of four chains with residues 32–230 in molecule A, residues 32–230 in molecule B, residues 33–76, 81–114, 120–131 and 134–230 in molecule C, residues 33–131 and 134–230 in molecule D and 138 water molecules. Ramachandran statistics were calculated using RAMPAGE as implemented in CCP4.[40] Values in parentheses apply to the high-resolution shell.

[a] $R_{sym} = \Sigma_h \Sigma_i | I_i(h) - \langle I(h) \rangle | / \Sigma_h \Sigma_i I(h)$, where $I(h)$ is the ith measurement of reflection h and $\langle I(h) \rangle$ is a weighted mean of all measurements of h.

[b] $R = \Sigma_h | F_{obs(h)} - F_{calc(h)} | / \Sigma_h | F_{obs} |$. R_{work} and R_{free} were calculated from the working and test reflection sets, respectively. The test set was 5% of the total reflections and these were not used in the refinement.

LAR adopts a very similar, albeit more compact, conformation with dimensions of ~44 Å× 41 Å×43 Å. The Ig1-2 regions of Dlar and LAR superimpose with rmsd 2.3 Å for 193 equivalent Cα positions (49% identity) whereas Ig1 and Ig2 superimpose with rmsd 1.5 Å (95 C^α positions) and 1.3 Å (89 C^α positions), respectively, indicating that much of the difference between the two molecules resides in the relative orientations of the individual Ig domains.

The antiparallel arrangement of Ig1 and Ig2

The most striking feature of the structures of Dlar(Ig1-2) and LAR(Ig1-2) is the presence of the horseshoe-like conformation of the two Ig domains. This antiparallel arrangement is made possible because of the flexible seven amino acid linker region between the domains (EGDKTPA in Dlar and EEDQLPSG in LAR) and the extensive

contacts between Ig1 and Ig2 (Fig. 2a and b). Overall, the interface buries 1199 Å^2 with a shape complementarity coefficient of 0.56 in Dlar(Ig1-2) and 1295 Å^2 with a shape complementarity coefficient of 0.69 in LAR(Ig1-2). These values are consistent, albeit slightly inferior in the case of Dlar(Ig1-2), with those found in known biological interfaces.[21,22] The contacts between Ig1 and Ig2 include mostly van der Waals interactions and are localized to a minor and a major site (Fig. 2a and b). The minor contact site, located at the linker region between the two Ig domains, includes R48, Y130, P136, F139 and V141 in Dlar(Ig1-2) and L45, L125, L130 and P131 in LAR(Ig1-2). The major contact site, which involves residues in the ABED face of the β-sheet in Ig1 and the GFCC' face of the β-sheet in Ig2 (Fig. 1c and d), encompasses amino acid residues Q45, V52, Y56, L85, R96, Q177, K179, E208, T217, E218 and H219 in Dlar and S51, I81, V90, R92, F169, F172, L173, P174, E206 and Y217 in LAR. There is little conservation in the residues involved in Ig1-Ig2 contacts in Dlar and mouse LAR (Fig. 2c): a single salt bridge (R96/E208 in Dlar(Ig1-2) and R92/E206 in LAR(Ig1-2)) is the sole interaction that is strictly conserved between the two proteins.

The fact that the Ig1-Ig2 fragments from two orthologous receptors adopt a similar antiparallel arrangement while crystallizing in distinct lattice environments indicates strongly that this conformation is not a consequence of crystallization, but is likely to represent the biologically active form of these proteins. To test this hypothesis directly, we analyzed the conformation of mouse LAR(Ig1-2) produced in *E. coli* in solution by small angle X-ray scattering (SAXS) (Fig. 2d and e). This technique is sensitive to the low-resolution shape of a molecule in solution and it can be used to discern to what extent LAR(Ig1-2) adopts an extended or compact conformation independently of the crystal lattice constraints.[23] Data sets were acquired at different concentrations and exposures and merged to obtain an experimental SAXS profile for LAR(Ig1-2). This profile was compared to the theoretical profile of LAR(Ig1-2), which was calculated by using the crystal structure in which it adopts a horseshoe-like shape (Fig. 2d and e). The experimental and theoretical scattering profiles match closely, indicating that the antiparallel conformation found in LAR crystals reflects its conformation in solution.

Because of the homology between LAR and Dlar, the results from our SAXS analysis suggest strongly that Dlar also adopts an antiparallel arrangement in solution. We engineered a mutant form of Dlar(Ig1-2) to test this hypothesis directly. In the crystal structure of Dlar(Ig1-2), residues V52 in Ig1 and H219 in Ig2 form van der Waals interactions across the interdomain interface (Fig. 2a). On the basis of our structural analysis, substitution of these residues for cysteine would result in the formation of an interdomain disulfide bond that would constrain the protein in its horseshoe-like conformation. This mutant protein was expressed in HEK293 cells as a fusion protein with hGH. It migrates slightly faster than the wild type Dlar(Ig1-2) in SDS-PAGE under non-reducing conditions, indicating that the disulfide bridge is indeed formed. Furthermore, introduction of the interdomain disulfide bond had no effect on the binding activity of Dlar(Ig1-2), as Fc fusions of wild type and the cysteine mutant of Dlar(Ig1-2) were able to stain cells in the ventral nerve cord of Drosophila larvae to a similar extent (Fig. 2g and h).[14,24] Taken together, these findings demonstrate that the biologically active forms of Dlar(Ig1-2) and LAR(Ig1-2) are the compact conformations identified in their crystal structures. Furthermore, all the residues involved in the interdomain contact in LAR are strictly conserved in its homologs PTPRD and PTPRS (Fig. 2c), indicating that these receptors are likely to contain a similar antiparallel arrangement. Overall, this analysis suggests strongly that the horseshoe-like conformation identified in Dlar(Ig1-2) and LAR(Ig1-2) is a hallmark of all type IIa RPTPs.

Comparison with other horseshoe-like conformations

Horseshoe-like arrangements of Ig domains have been reported for several Ig-containing proteins. These cell surface receptors include the natural killer cell receptor p58[25] and, more recently, Ig superfamily proteins such as the insect hemolymph protein hemolin, the neuronal receptors axonin/TAG1/contactin-2 (CNTN2), CNTN4 and isoforms of the Down syndrome cell adhesion molecule (Dscam).[16,26–30] In these latter structures, the four N-terminal Ig domains adopt closely related horseshoe-like conformations mediated by extensive contacts between Ig1 and Ig4 on one hand and Ig2 and Ig3 on the other hand. In addition, Ig5 and Ig6 in Dscam adopt an antiparallel arrangement distinct from that adopted by Ig2 and Ig3.[29] Hence, it was of interest to compare the structure of Dlar(Ig1-2) with these known structures containing horseshoe-like motifs. To this end, we superimposed Ig1 of Dlar with Ig1 of p58 (rmsd 2.7 Å over 81 equivalent C^α positions) with Ig2 of CNTN4 (rmsd 4.5 Å over 93 equivalent C^α positions) or with Ig5 of Dscam (rmsd 3.5 Å over 96 equivalent C^α positions) to assess to what extent the C-terminal domains of the tandem overlapped (Fig. 3a). Proteins were superimposed using the DaliLite server[31] and the program SUPERPOSE as implemented in CCP4.[32] CNTN4 was chosen in this case because its second Ig domain shares the greatest degree of sequence identity (26%) with Ig1 of Dlar among hemolin, CNTN2, CNTN4 and Dscam. This analysis shows that the Ig1-Ig2 conformation observed in Dlar is distinct from those

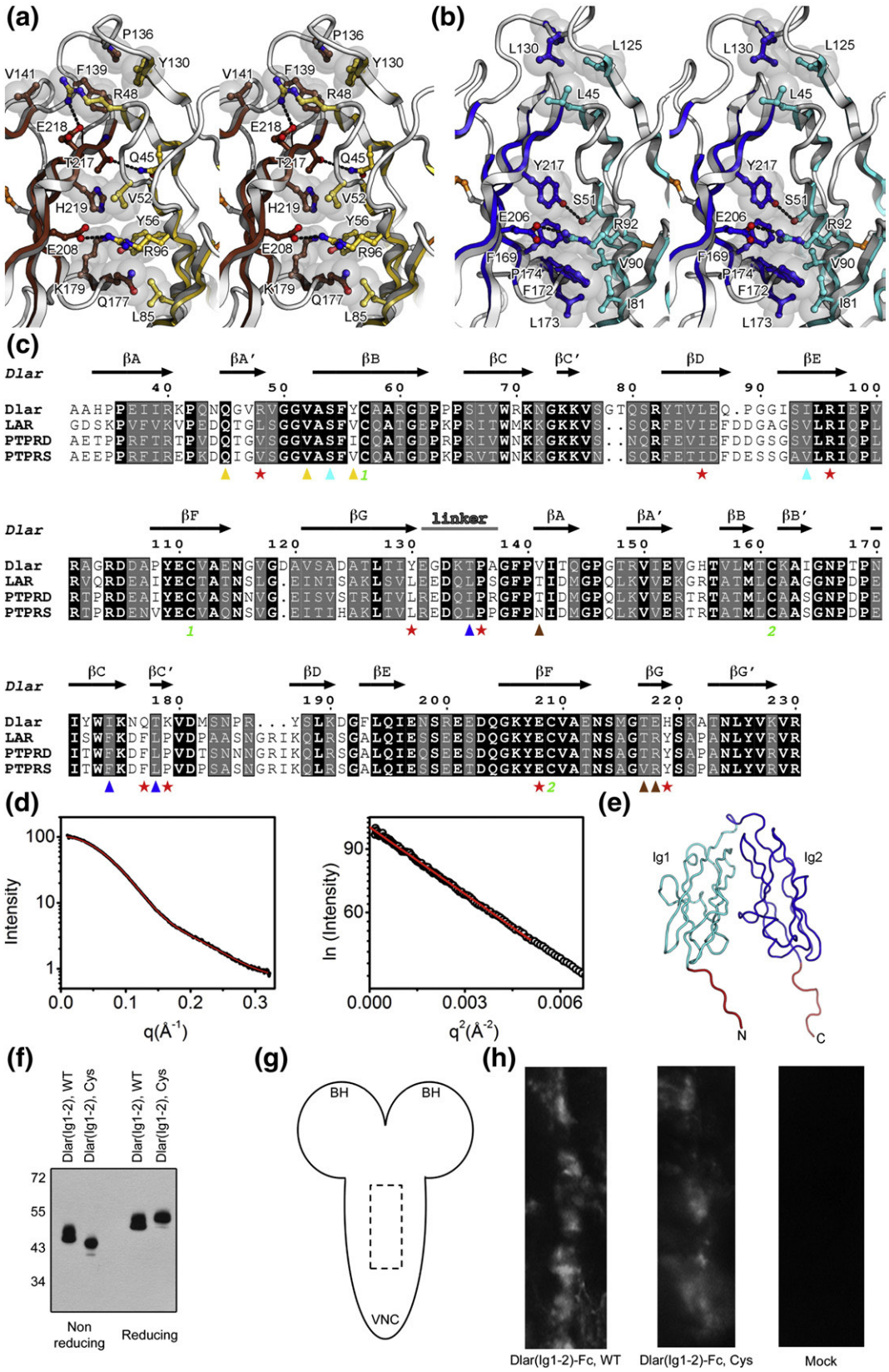

Fig. 2 (*legend on next page*)

reported for p58, CNTN4 and Dscam. The differences in these structures can be rationalized when comparing the interfaces between the Ig domains. In Dlar, residues at the ABED face of the β-sheet in Ig1 are in contact with residues in the GFCC' face of the β-sheet in Ig2. In contrast, in p58, the contacting faces involve strands G, F and A' in Ig1 and strands F, C and C' in Ig2, whereas in CNTN4, the contacting faces involve strands B, E and D in Ig2 and strands A, B and D in Ig3. The comparison with Ig5-Ig6 of Dscam shows a striking change of direction with, essentially, the backbone 'turning left' in Dscam and 'turning right' in Dlar. Likewise, the contacting faces in Ig5 and Ig6 are distinct from those in Ig1 and Ig2 of Dlar as they involve residues in strands F, C and C' in Ig5 and strands A and B' in Ig6. These comparisons with other horseshoe-like conformations of Ig domains indicate that the arrangement of Ig1 and Ig2 in type IIa RPTPs is so far unique among tandem Ig repeats.

In contrast, the structures of Dlar(Ig1-2) and LAR (Ig1-2) are reminiscent of the arrangement adopted by the FNIII repeats 1 and 2 of iHog, a Drosophila Ig superfamily protein that functions as a coreceptor for the morphogen hedgehog (Fig. 3c).[33] Using secondary structure alignment,[34] Ig1 of Dlar can be superimposed onto FN1 of iHog with rmsd 3.2 Å over 60 equivalent C^{α} positions and, overall, the Ig1-Ig2 pair of Dlar can be superimposed onto the FN1-FN2 pair of iHog with rmsd 4.0 Å over 143 equivalent C^{α} positions, even though the extent of sequence identity is only 6% (Fig. 3d). Similarly, the Ig1-Ig2 fragment of mouse LAR can be superimposed onto the FN1-FN2 pair of iHog with rmsd 3.6 Å over 147 residues. Furthermore, the interdomain interface in iHog encompasses residues located in the ABE face of the β-sheet in FN1 and strands FCC' in FN2,[33] which is similar to the contacting faces in Ig1 and Ig2 of Dlar (Fig. 3a and c). Overall, these structural findings suggest that horseshoe-like conformations are recurring motifs in Ig superfamily proteins involved in cell adhesion and signaling.

Binding to heparin

A noteworthy feature of the Dlar(Ig1-2) and LAR (Ig1-2) structures is the presence of a shallow groove bordered by positively charged residues that

Fig. 2. The Ig1–Ig2 interface in Dlar and mouse LAR. (a) A stereo view of the interface between Ig domains 1 and 2 in Dlar(Ig1-2). Residues at the interface between the two domains are shown as ball-and-sticks and colored gold (Ig1) or brown (Ig2). This view is in the same orientation as the right-hand view in Fig. 1c. Transparent gray spheres and broken lines denote residues involved in van der Waals contacts and potential hydrogen bonds and salt bridges, respectively. (b) A stereo view of the interface between Ig domains 1 and 2 in LAR(Ig1-2). Residues at the interface between the two domains are shown as ball-and-sticks and colored cyan (Ig1) or blue (Ig2). This view is in the same orientation as the right-hand view in Fig. 1d. (c) Alignment of amino acid sequences of Ig domains 1-2 of Dlar and mouse type IIa RPTPs. Strictly conserved residues are shaded in black and similar residues are colored gray. The numbering refers to Dlar. Cysteine residues involved in disulfide bridges are numbered in green below the sequences. Gold and brown triangles indicate the residues involved in interactions between Ig domains 1 and 2 of Dlar, respectively, that are unique to Dlar. Cyan and blue triangles indicate the residues involved in interactions between Ig domains 1 and 2 of mouse LAR, respectively, that are unique to LAR. Red stars indicate the positions of residues involved in interactions between Ig1 and Ig2 of both Dlar and mouse LAR. (d) Analysis of the interference-free SAXS curve for LAR(Ig1-2). The experimental scattering profile (black) for LAR(Ig1-2) and the theoretical scattering (red line, χ^2 1.2) calculated from the LAR(Ig1-2) crystal structure are shown in the left-hand panel. The right-hand panel shows the Guinier plots with linear fit (red line). (e) The crystal structure of LAR(Ig1-2) used to calculate the theoretical scattering profile in which the red regions indicate disordered residues at the N- and C-termini that have been added in extended form and optimized by BILBOMD.[41] These residues are GPGSSRG at the N-terminus and VRRVAPRFS at the C-terminus. SAXS data were collected at the ALS beamline 12.3.1 LBNL Berkeley, CA.[42] Tunable wavelength λ 1.0–1.5 Å and the sample-to-detector distances were set to 1.5 m, resulting in scattering vectors, q, ranging from 0.01 $Å^{-1}$ to 0.32 $Å^{-1}$. The scattering vector q is defined as $4\pi \sin\theta/\lambda$, where 2θ is the scattering angle. All experiments were performed at 20°C and the data were processed as described.[42] Data acquired for short and long exposure (0.5 and 5 s) were merged for calculations using the entire scattering profile. The experimental SAXS data for different protein concentrations were investigated for aggregation using Guinier plots.[43] The radius of gyration R_G is 20.5 ± 0.1 Å and was derived by the Guinier approximation $I(q) = I(0) \exp(-q^2R_G^2/3)$ with the limits $qR_G < 1.6$. The theoretical SAXS profile and the corresponding fit to the experimental data were calculated using the program FoXS.[44] (f) Fusion proteins of hGH with wild type Dlar(Ig1-2) (labeled Dlar(Ig1-2), WT) and a mutant form of Dlar(Ig1-2) with cysteine residues at positions 52 and 219 (labeled Dlar(Ig1-2), Cys)) were analyzed by SDS-PAGE (9% (w/v) polyacrylamide gel) under reducing and non-reducing conditions followed by immunoblotting against hGH. (g) A drawing of the Drosophila larval central nervous system, illustrating the positions of the two brain hemispheres (BH) and the ventral nerve cord (VNC). (h) Confocal micrographs of the VNC region shown in (g) (dotted box). Fc fusion proteins of histidine-tagged wild type and mutant Dlar(Ig1-2) were expressed in HEK293 cells and purified from conditioned media by cobalt-affinity chromatography. The brain/VNC complex was dissected from second instar larvae in phosphate-buffered saline (PBS) and incubated with the indicated fusion proteins for 30 min. After fixation for 30 min in PBS/formaldehyde, the larval VNC was washed with PBS, 0.05% (v/v) Triton X-100 and incubated with protein G-Alexa Fluor 488. Both wild type Dlar (Ig1-2) (left-hand panel) and its cysteine mutant (middle panel) showed a staining pattern consistent with ventral nerve cord staining that was not observed in mock (cobalt-affinity chromatography eluate of conditioned media from untransfected HEK293 cells, right-hand panel).

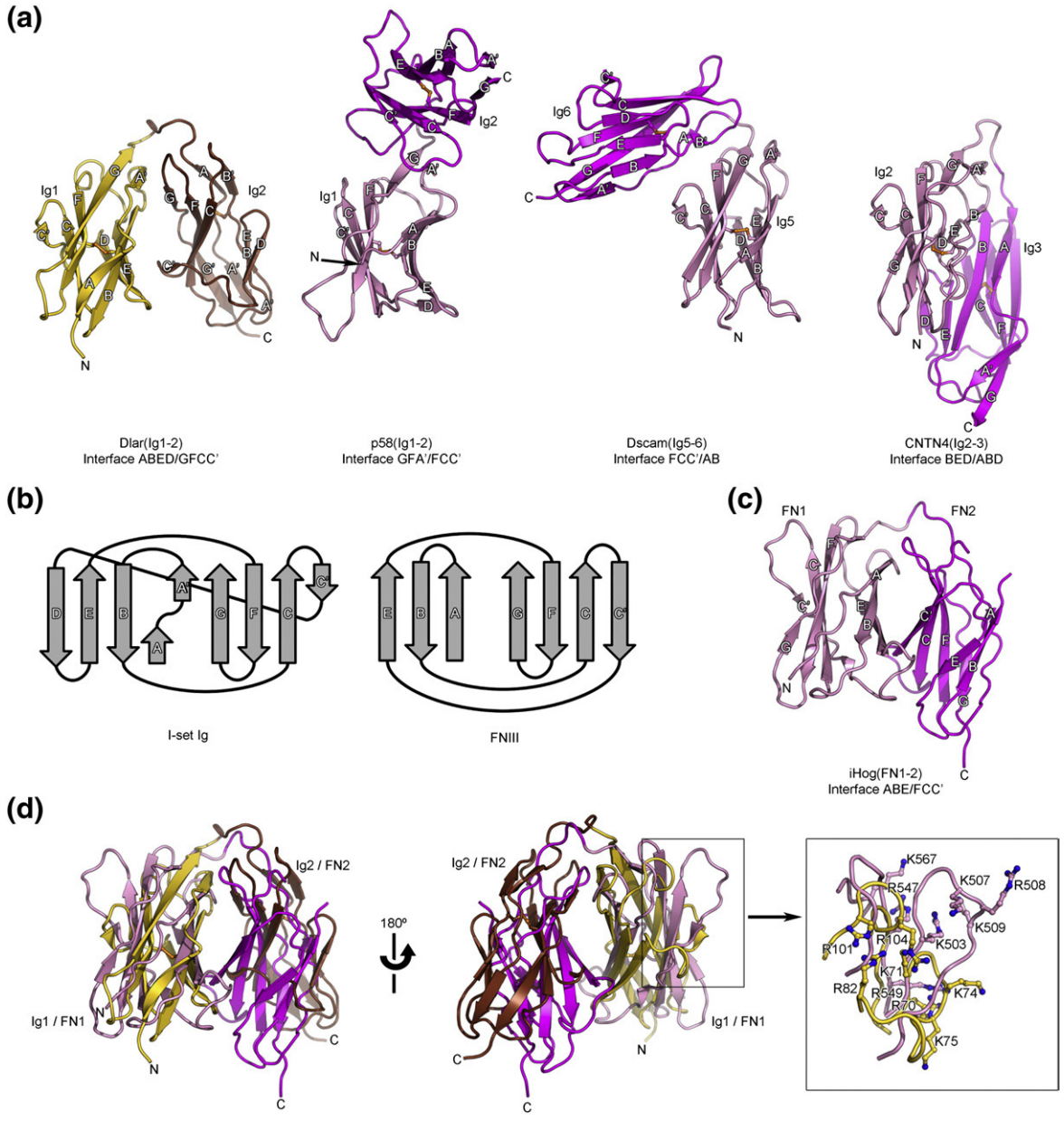

Fig. 3. Structural comparisons of the antiparallel arrangement of Ig domains 1 and 2 of Dlar with other horseshoe-like structures. (a) Comparison of the Dlar(Ig1-2) structure (left, colored gold and brown) with three structures of tandem Ig repeats that adopt a horseshoe-like conformation. The distinct conformations of horseshoe-like arrangements of these tandem β-sandwiches repeats arise from interactions between residues found in distinct strands, which are listed below each structure. In each of the compared structures, the N-terminal and C-terminal Ig repeats are colored pink and magenta, respectively. Disulfide bonds are shown as orange ball-and-stick models. The letters N and C indicate the N- and C-termini, respectively. These views of Dlar(Ig1-2), p58(Ig1-2), CNTN4(Ig2-3) and Dscam(Ig5-6) were obtained by superimposing their N-terminal Ig repeats in an effort to highlight the distinct positions of the C-terminal Ig repeats that result from changes in the contacting faces in each of the horseshoe-like structures. (b) Topology diagrams for I set Ig domains and FNIII domains highlighting the similarities between the two folds.[45] (c) A ribbon diagram depicting the horseshoe-like arrangement observed for the tandem FNIII repeats of Drosophila iHog. The first and second FNIII repeats of iHog are colored pink and magenta, respectively. This view was obtained by superimposing the first FNIII iHog and the first Ig repeat of Dlar using secondary structure matching. (d) Overlay of the Ig1-Ig2 fragment of Dlar with the first (pink) and second (magenta) FNIII domains of Drosophila iHog. The HSPG-binding site for iHog and the presumed glycosaminoglycan-binding site for Dlar are partially overlapped and are shown in the boxed view.

localizes to the first Ig domain in each structure and could function as a glycosaminoglycan-binding site. In Dlar, this site encompasses residues R70, K71, K74, K75 and R82 located at the C-terminal end of the C strand and on strand C' as well as residues R101 and R104 found in the loop between strands E and F (Figs. 1c and 4a). Similarly, in mouse LAR, the surface is composed of the equivalent residues K68, K69, K71, K72, R77, R97 and R100 (Figs. 1d and 4b). In addition, two sulfate ions are bound to this positively charged region in the vicinity of R97 and R100. Interestingly, the location of this putative glycosaminoglycan-binding site in the region between strands C and D corresponds to the heparin-binding sites found in the 14th FNIII repeat of fibronectin[35] and in the 1st FNIII repeat of iHog (Fig. 3d),[33] further emphasizing the resemblance between Dlar and iHog. Furthermore, the basic residues in LAR are conserved in PTPRD and PTPRS and introduction of pairs of alanine residues into avian PTPRS in place of the residues equivalent to positions K68/K69, K71/K72 and R97/R100 in LAR impaired binding to HSPGs (Fig. 4b).[15] Taken together, these observations indicate that this positively charged region is the likely binding site for glycosaminoglycans and that it is a conserved surface feature of type IIa RPTPs.

The positively charged region identified in the crystal structures of Dlar and LAR localizes to the first Ig repeat and no such region was observed in the second Ig repeat, which would suggest that only the first Ig domain is required to bind heparin. Therefore, we used a heparin affinity isolation assay to test this hypothesis and to determine to what extent the horseshoe conformation adopted by Dlar is important for heparin binding. Secreted fragments of Dlar fused to hGH were expressed in HEK293 cells and incubated with Heparin Sepharose. As

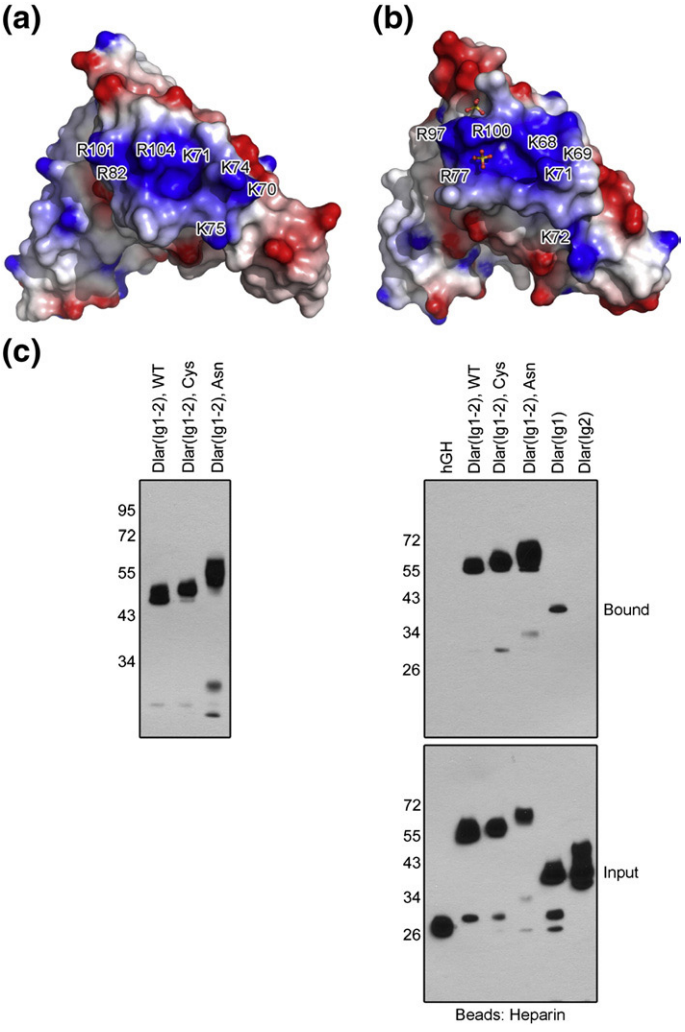

Fig. 4. (a) Electrostatic surface representation of Dlar(Ig1-2). This view is related to the left-hand view in panel c by a counterclockwise rotation of 120° along a vertical axis. Regions with negative electrostatic potential are colored red and regions with positive electrostatic potential are colored blue (scale ± 5 e/kT). (b) Electrostatic surface representation of mouse LAR(Ig1-2). Electrostatic potentials were calculated with DELPHI.[46,47] (c) Interactions between mutants of Dlar(Ig1-2) and heparin. Fragments of Dlar fused to hGH were expressed transiently in HEK293 cells. These fragments include wild type Dlar(Ig1-2) (labeled Dlar(Ig1-2), WT), Dlar(Ig1-2) with cysteine residues at positions 52 and 219 (labeled Dlar(Ig1-2), Cys), with asparagine residues at positions 52 and 218 to introduce N-linked carbohydrates (Dlar(Ig1-2), Asn), Dlar(Ig1) and Dlar(Ig2). The left-hand panel shows wild type Dlar(Ig1-2) along with its cysteine and asparagine mutants to illustrate the difference in size between the proteins upon introduction of consensus N-linked glycosylation sites in Dlar(Ig1-2). Samples were resolved by SDS-PAGE (9% (w/v) polyacrylamide gel) and fusion proteins were visualized by immunoblotting against hGH. The right-hand panel shows the results of heparin affinity isolation assays. Conditioned media were incubated with Heparin Sepharose™ in PBS, 1 % (v/v) Tween-20 at room temperature for 1 h. Resins were washed in the same buffer and samples were resolved by SDS-PAGE (12% (w/v) polyacrylamide gel). Bound fusion proteins were visualized by immunoblotting against hGH.

expected from our previous analysis (Fig. 2g and h), a mutant form of Dlar(Ig1-2) locked into a horseshoe conformation by formation of a disulfide bridge across the Ig1-Ig2 interface supports heparin binding. Similar results were obtained for Ig1, whereas repeat Ig2 of Dlar appears unable to bind heparin (Fig. 4c). To determine to what extent disruption of the Ig1-Ig2 interface would impair heparin binding, we introduced asparagine residues at positions V52 and E218 to create N-linked carbohydrate consensus sites N52AS in Ig1 and N218HS in Ig2, respectively. We would expect that these changes would prevent Dlar from adopting a horseshoe conformation because N-linked glycans at these positions would present a steric barrier to the formation of the Ig1-Ig2 interface. This mutant migrates slightly slower than wild-type Dlar(Ig1-2) and the band obtained is broader, indicating that additional N-linked carbohydrates were indeed appended (Fig. 4c, left-hand panel). As expected, our affinity isolation assay shows that this mutant form of Dlar binds to heparin, which is in line with the fact that Ig1 alone supports heparin binding. Overall, these results indicate that the horseshoe conformation is not necessary to bind to heparin and, presumably, the glycosaminoglycan ligands of Dlar and LAR, thus raising the question of what is the exact role of this unusual antiparallel arrangement.

Conclusion

Dlar and its vertebrate counterparts LAR, PTPRD and PTPRS mediate crucial interactions with proteoglycans during neurogenesis. The crystal structures of the Ig1-Ig2 fragments of Dlar and mouse LAR make it possible to identify a group of conserved lysine and arginine residues that form a positively charged patch in the first Ig domain, which is the likely glycosaminoglycan-binding site in type IIa RPTPs. Interestingly, PTPRS has been shown to bind to two different classes of proteoglycans with dissociation constants in the nanomolar range. Indeed, avian PTPRS associates with the HSPGs agrin and collagen XVIII,[15] whereas murine PTPRS is a physiological receptor for chondroitin sulfate proteoglycans (CSPGs).[36] Much like HSPGs, CSPGs are composed of a core protein to which chondroitin sulfate (CS) chains are attached. However, CS chains differ in their carbohydrate composition from HS chains in that they include repeats of a disaccharide formed by glucuronic acid and *N*-acetyl galactosamine instead of repeats of either glucuronic acid or its epimer iduronic acid linked to *N*-acetyl glucosamine found in HS chains. In addition, HS disaccharide units can be sulfated on up to four of the available hydroxyl groups, whereas CS chains are usually less sulfated than HS chains. The interactions between proteoglycans and PTPRS were sensitive to treatment with either the HS-degrading enzyme heparinase or the CS-degrading enzyme chondroitinase ABC, suggesting that the glycosaminoglycan chains, not the protein cores, are primarily responsible for these binding events.[15,36] In addition, mutations of a cluster of basic residues in the first Ig domain of avian PTPRS that impaired binding to HSPGs had an identical effect on the binding of CSPGs to mouse PTPRS. Avian and murine PTPRS share more than 80% amino acid sequence identity and are virtually identical in the Ig1-Ig2 region so that differences in ligand-binding specificities are unlikely to reflect species differences. Taken together, these findings indicate that PTPRS in particular and type IIa RPTPs in general might not discriminate between HSPGs and CSPGs and appear to harbor a more generic glycosaminoglycan-binding site, which would be the positively charged surface identified in the work presented here.

Conversely, the unexpected resemblance between type IIa RPTPs and the hedgehog coreceptor iHog could suggest a more complex role for the positively charged surface on Ig1. In particular, iHog also binds to HSPGs and its HS-binding site is located in its FN1 repeat in a loop region that partially overlaps the putative HSPG-binding site in Dlar and LAR (Fig. 3d). Even more interesting is the fact that iHog forms a ternary complex with hedgehog and HSPGs.[33] Although this would appear to be in contrast to the idea of a generic glycosaminoglycan-binding site on Ig1, it is tempting to speculate that type IIa RPTPs might also associate with protein ligands in an HSPG-dependent manner. In this case, there might be two distinct classes of ligands for type IIa RPTPs, CSPGs on one hand and HSPGs/HSPG-dependent protein ligands on the other hand, which could in turn elicit distinct biological functions. However, no HSPG-dependent protein ligand has been identified for type IIa RPTPs. In the future, structural investigation of glycosaminogycan-bound forms of type IIa receptors as well as renewed efforts to identify physiological ligands for these proteins could shed light on the protein–carbohydrate interactions that underlie crucial cell adhesion and signaling events during the development and growth of the nervous system.

Protein Data Bank accession numbers

The atomic coordinates and structure factors have been deposited in the Protein Data Bank† under accession codes 3PXH for mouse LAR(Ig1-2) and 3PXJ for Dlar(Ig1-2).

Supplementary materials related to this article can be found online at doi:10.1016/j.jmb.2011.03.013

† www.pdb.org

Acknowledgements

The authors thank Brian Geisbrecht for helpful comments on the manuscript. This work was supported by the National Institute Of General Medical Sciences (award number R01GM088806). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract no. W-31-109-Eng-38. Data were collected at Southeast Regional Collaborative Access Team beamlines at the Advanced Photon Source, Argonne National Laboratory‡. X-ray scattering technologies at the Lawrence Berkeley National Laboratory SIBYLS beamline of the Advanced Light Source are supported, in part, by the U.S. Department of Energy program Integrated Diffraction Analysis Technologies (IDAT).

References

1. Chang, C., Yu, T. W., Bargmann, C. I. & Tessier-Lavigne, M. (2004). Inhibition of netrin-mediated axon attraction by a receptor protein tyrosine phosphatase. *Science*, **305**, 103–106.
2. Johnson, K. G. & Van Vactor, D. (2003). Receptor protein tyrosine phosphatases in nervous system development. *Physiol. Rev.* **83**, 1–24.
3. Desai, C. J., Gindhart, J. G., Goldstein, L. S. & Zinn, K. (1996). Receptor tyrosine phosphatases are required for motor axon guidance in the Drosophila embryo. *Cell*, **84**, 599–609.
4. Krueger, N. X., Van Vactor, D., Wan, H. I., Gelbart, W. M., Goodman, C. S. & Saito, H. (1996). The transmembrane tyrosine phosphatase DLAR controls motor axon guidance in Drosophila. *Cell*, **84**, 611–622.
5. Wills, Z., Bateman, J., Korey, C. A., Comer, A. & Vactor, V. (1999). The tyrosine kinase Abl and its substrate enabled collaborate with the receptor phosphatase Dlar to control motor axon guidance. *Neuron*, **22**, 301–312.
6. Johnson, K. G., Tenney, A. P., Ghose, A., Duckworth, A. M., Higashi, M. E., Parfitt, K. *et al.* (2006). The HSPGs Syndecan and Dallylike bind the receptor phosphatase LAR and exert distinct effects on synaptic development. *Neuron*, **49**, 517–531.
7. Hofmeyer, K. & Treisman, J. E. (2009). The receptor protein tyrosine phosphatase LAR promotes R7 photoreceptor axon targeting by a phosphatase-independent signaling mechanism. *Proc. Natl Acad. Sci. USA*, **106**, 19399–19404.
8. Andersen, J. N., Mortensen, O. H., Peters, G. H., Drake, P. G., Iversen, L. F., Olsen, O. H. *et al.* (2001). Structural and evolutionary relationships among protein tyrosine phosphatase domains. *Mol. Cell. Biol.* **21**, 7117–7136.
9. Kolkman, M. J., Streijger, F., Linkels, M., Bloemen, M., Heeren, D. J., Hendriks, W. J. & Van der Zee, C. E. (2004). Mice lacking leukocyte common antigen-related (LAR) protein tyrosine phosphatase domains demonstrate spatial learning impairment in the two-trial water maze and hyperactivity in multiple behavioural tests. *Behav. Brain Res.* **154**, 171–182.
10. Uetani, N., Kato, K., Ogura, H., Mizuno, K., Kawano, K., Mikoshiba, K. *et al.* (2000). Impaired learning with enhanced hippocampal long-term potentiation in PTPdelta-deficient mice. *EMBO J.* **19**, 2775–2785.
11. Elchebly, M., Wagner, J., Kennedy, T. E., Lanctôt, C., Michaliszyn, E., Itié, A., Drouin, J. & Tremblay, M. L. (1999). Neuroendocrine dysplasia in mice lacking protein tyrosine phosphatase sigma. *Nat. Genet.* **21**, 330–333.
12. Wallace, M. J., Batt, J., Fladd, C. A., Henderson, J. T., Skarnes, W. & Rotin, D. (1999). Neuronal defects and posterior pituitary hypoplasia in mice lacking the receptor tyrosine phosphatase PTPsigma. *Nat. Genet.* **21**, 334–338.
13. Uetani, N., Chagnon, M. J., Kennedy, T. E., Iwakura, Y. & Tremblay, M. L (2006). Mammalian motoneuron axon targeting requires receptor protein tyrosine phosphatases sigma and delta. *J. Neurosci.* **26**, 5872–5880.
14. Fox, A. N. & Zinn, K. (2005). The heparan sulfate proteoglycan syndecan is an in vivo ligand for the Drosophila LAR receptor tyrosine phosphatase. *Curr. Biol.* **15**, 1701–1711.
15. Aricescu, A. R., McKinnell, I. W., Halfter, W. & Stoker, A. W. (2002). Heparan sulfate proteoglycans are ligands for receptor protein tyrosine phosphatase sigma. *Mol. Cell. Biol.* **22**, 1881–1892.
16. Bouyain, S. & Watkins, D. J. (2010). The protein tyrosine phosphatases PTPRZ and PTPRG bind to distinct members of the contactin family of neural recognition molecules. *Proc. Natl Acad. Sci. USA*, **107**, 2443–2448.
17. Harpaz, Y. & Chothia, C. (1994). Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. *J. Mol. Biol.* **238**, 528–539.
18. Fukuhara, N., Howitt, J. A., Hussain, S.-A. & Hohenester, E. (2008). Structural and functional analysis of slit and heparin binding to immunoglobulin-like domains 1 and 2 of Drosophila Robo. *J. Biol. Chem.* **283**, 16226–16234.
19. Morlot, C., Thielens, N. M., Ravelli, R. B. G., Hemrika, W., Romijn, R. A., Gros, P. *et al.* (2007). Structural insights into the Slit-Robo complex. *Proc. Natl Acad. Sci. USA*, **104**, 14923–14928.
20. Schlessinger, J., Plotnikov, A. N., Ibrahimi, O. A., Eliseenkova, A. V., Yeh, B. K. & Yayon, A. (2000). Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol. Cell*, **6**, 743–750.
21. Lawrence, M. C. & Colman, P. M. (1993). Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* **234**, 946–950.
22. Lo Conte, L., Chothia, C. & Janin, J. (1999). The atomic structure of protein–protein recognition sites. *J. Mol. Biol.* **285**, 2177–2198.
23. Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate

‡For a list of supporting member institutions, see www.ser-cat.org/members.html

macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40**, 191–285.
24. Tian, S. S., Tsoulfas, P. & Zinn, K. (1991). Three receptor-linked protein-tyrosine phosphatases are selectively expressed on central nervous system axons in the Drosophila embryo. *Cell*, **67**, 675–685.
25. Fan, Q. R., Mosyak, L., Winter, C. C., Wagtmann, N., Long, E. O. & Wiley, D. C. (1997). Structure of the inhibitory receptor for human natural killer cells resembles haematopoietic receptors. *Nature*, **389**, 96–100.
26. Freigang, J., Proba, K., Leder, L., Diederichs, K., Sonderegger, P. & Welte, W. (2000). The crystal structure of the ligand binding module of axonin-1/TAG-1 suggests a zipper mechanism for neural cell adhesion. *Cell*, **101**, 425–433.
27. Meijers, R., Puettmann-Holgado, R., Skiniotis, G., Liu, J. H., Walz, T., Wang, J. H. & Schmucker, D. (2007). Structural basis of Dscam isoform specificity. *Nature*, **449**, 487–491.
28. Mortl, M., Sonderegger, P., Diederichs, K. & Welte, W. (2007). The crystal structure of the ligand-binding module of human TAG-1 suggests a new mode of homophilic interaction. *Protein Sci.* **16**, 2174–2183.
29. Sawaya, M. R., Wojtowicz, W. M., Andre, I., Qian, B., Wu, W., Baker, D. *et al.* (2008). A double S shape provides the structural basis for the extraordinary binding specificity of Dscam isoforms. *Cell*, **134**, 1007–1018.
30. Su, X. D., Gastinel, L. N., Vaughn, D. E., Faye, I., Poon, P. & Bjorkman, P. J. (1998). Crystal structure of hemolin: a horseshoe shape with implications for homophilic adhesion. *Science*, **281**, 991–995.
31. Holm, L., Kääriäinen, S., Rosenström, P. & Schenkel, A. (2008). Searching protein structure databases with DaliLite v.3. *Bioinformatics*, **24**, 2780–2781.
32. Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.
33. McLellan, J. S., Yao, S., Zheng, X., Geisbrecht, B. V., Ghirlando, R., Beachy, P. A. & Leahy, D. J. (2006). Structure of a heparin-dependent complex of Hedgehog and Ihog. *Proc. Natl Acad. Sci. USA*, **103**, 17208–17213.
34. Krissinel, E. & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D*, **60**, 2256–2268.
35. Sharma, A., Askari, J. A., Humphries, M. J., Jones, E. Y. & Stuart, D. I. (1999). Crystal structure of a heparin- and integrin-binding segment of human fibronectin. *EMBO J.* **18**, 1468–1479.
36. Shen, Y., Tenney, A. P., Busch, S. A., Horn, K. P., Cuascut, F. X., Liu, K. *et al.* (2009). PTPσ Is a receptor for chondroitin sulfate proteoglycan, an inhibitor of neural regeneration. *Science*, **326**, 592–596.
37. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
38. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W. *et al.* (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D*, **58**, 1948–1954.
39. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D*, **60**, 2126–2132.
40. Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I., Word, J. M., Prisant, M. G. *et al.* (2003). Structure validation by Cα geometry: ϕ, φ and Cβ deviation. *Proteins: Struct. Funct. Genet.* **50**, 437–450.
41. Pelikan, M., Hura, G. L. & Hammel, M. (2009). Structure and flexibility within proteins as identified through small angle X-ray scattering. *Gen. Physiol. Biophys.* **28**, 174–189.
42. Hura, G. L., Menon, A. L., Hammel, M., Rambo, R. P., Poole, F. L. n., Tsutakawa, S. E. *et al.* (2009). Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat. Methods*, **6**, 606–612.
43. Guinier, A. & Fournet, F. (1955). *Small Angle Scattering of X-rays.* Wiley Interscience, New York, NY.
44. Schneidman-Duhovny, D., Hammel, M. & Sali, A. (2010). FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res.* **38**, W540–W544.
45. Leahy, D. J., Hendrickson, W. A., Aukhil, I. & Erickson, H. P. (1992). Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionyl protein. *Science*, **258**, 987–991.
46. Rocchia, W., Alexov, E. & Honig, B. (2001). Extending the applicability of the nonlinear Poisson-Boltzman equation: multiple dielectric constants and multivalent Ions. *J. Phys. Chem. B*, **105**, 6507–6514.
47. Rocchia, W., Sridharan, S., Nicholls, A., Alexov, E., Chiabrera, A. & Honig, B. (2002). Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. *J. Comput. Chem.* **23**, 128–137.