Neural immunoglobulin superfamily interaction networks
Kai Zinn\textsuperscript{1} and Engin Özkan\textsuperscript{2}

The immunoglobulin superfamily (IgSF) encompasses hundreds of cell surface proteins containing multiple immunoglobulin-like (Ig) domains. Among these are neural IgCAMs, which are cell adhesion molecules that mediate interactions between cells in the nervous system. IgCAMs in some vertebrate IgSF subfamilies bind to each other homophilically and heterophilically, forming small interaction networks. In *Drosophila*, a global ‘interactome’ screen identified two larger networks in which proteins in one IgSF subfamily selectively interact with proteins in a different subfamily. One of these networks, the ‘Dpr-ome’, includes 30 IgSF proteins, each of which is expressed in a unique subset of neurons. Recent evidence shows that one interacting protein pair within the Dpr-ome network is required for development of the brain and neuromuscular system.

**Addresses**
\textsuperscript{1} Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, United States
\textsuperscript{2} Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637, United States

Corresponding author: Zinn, Kai (zinnk@caltech.edu)

**Current Opinion in Neurobiology** 2017, 45:99–105
This review comes from a themed issue on *Molecular neuroscience*
Edited by Susumu Tomita and Brenda Bloodgood

http://dx.doi.org/10.1016/j.conb.2017.05.010
0959-4388/© 2017 Elsevier Ltd. All rights reserved.

**Introduction**
The immunoglobulin superfamily (IgSF) is a very large family of proteins containing one or more immunoglobulin-like (Ig) domains. Most IgSF proteins are secreted or localized to the cell surface. There are about 500 non-antibody, non-T cell receptor (TCR) IgSF proteins encoded in the human genome, and 130 in *Drosophila melanogaster*. The Ig domain has a sandwich-like fold composed of two sheets of antiparallel β strands. Ig domains were probably selected by evolution as versatile mediators of protein–protein interactions caused by their ability to bind to partners via several different interfaces (Figure 1a). Another domain, the fibronectin type III (FnIII) repeat, also has a β-sandwich structure and mediates protein–protein interactions, and many proteins have both Ig domains and FnIII repeats.

Many cell surface IgSF proteins are homophilic or heterophilic adhesion molecules (IgCAMs) containing multiple Ig domains. IgCAMs mediate interactions among neurons and between neurons and glia. Homophilic IgCAMs include NCAM and L1 in mammals and Fasciclin II and Dscam in *Drosophila*. IgCAM-like proteins are receptors for axon guidance cues such as Netrins and Slits. Finally, neuronal transmembrane signal transduction molecules, such as receptor tyrosine kinases and phosphatases, can have IgCAM-like extracellular (XC) domains.

This mini review does not cover homophilic IgCAMs, axon guidance receptors, or signaling receptors, but focuses on subfamilies of the IgSF that participate in complex networks of interactions. In these networks, individual IgCAMs usually bind to multiple partners. Networks defined thus far in mammals and other vertebrates are composed primarily of proteins within the same IgSF subfamily, since it is straightforward to test members of an identified subfamily for interactions with each other. In *Drosophila*, however, a comprehensive analysis of interactions among all IgSF proteins has uncovered complex networks in which most interactions occur between proteins in different IgSF subfamilies [1\textsuperscript{**}]. It remains to be determined if large interaction networks that include multiple subfamilies also exist in vertebrates.

**IgSF subfamilies participating in interaction networks in mammals**
The IgSF has expanded in the vertebrate lineage, generating many new subfamilies [2]. Here we review three subfamilies that form interaction networks.

**Nectins and Nectin-like proteins**
Nectins and Nectin-like (Necl) proteins form a nine-member subfamily in mammals (Figure 1b) [3\textsuperscript{**}]. *In situ* hybridization analysis (Allen Brain Atlas) shows that all of their genes are expressed at high levels in the adult mouse brain [4]. Nectins and Necls are cell surface receptors with three Ig domains and single transmembrane helices. They are localized to cell–cell junctions, and are commonly, but not always, found in conjunction with cadherins. The Nectin proteins interact with the actin-binding adaptor protein afadin via their cytoplasmic tails, while Necls lack this activity [5].
Nectins and Necls can mediate cell adhesion through homophilic or heterophilic interactions. Nectin and Necl proteins interact homophilically in vitro, and several Nectin and Necl pairs form heterophilic complexes (Figure 1b) [3,6]. Overall, the heterophilic interactions of Nectins/Necls are of higher affinity than the homophilic ones. Recently, a group of five more IgSF proteins were identified as members of an extended Nectin/Necl subfamily that can bind to Nectins and Necls. These proteins (CRTAM, CD96, CD200, CD226, and TIGIT) are primarily known for roles in the immune system [7] (Figure 1b). However, at least one of them, CD200, is also expressed in the brain. In all cases of biophysically characterized interactions, complexes are created solely by binding between the N-terminal variable-type Ig (IgV) domains [6,8,9] (Figure 1a, c). However, there is also evidence for lateral (cis) interactions on the same plasma membrane, mediated by the second and third Ig domains of Nectin-2 [10].

Nectins and Necls can mediate cell adhesion through homophilic or heterophilic interactions. Nectin and Necl proteins interact homophilically in vitro, and several Nectin and Necl pairs form heterophilic complexes (Figure 1b) [3,6]. Overall, the heterophilic interactions of Nectins/Necls are of higher affinity than the homophilic ones. Recently, a group of five more IgSF proteins were identified as members of an extended Nectin/Necl subfamily that can bind to Nectins and Necls. These proteins (CRTAM, CD96, CD200, CD226, and TIGIT) are primarily known for roles in the immune system [7] (Figure 1b). However, at least one of them, CD200, is also expressed in the brain. In all cases of biophysically characterized interactions, complexes are created solely by binding between the N-terminal variable-type Ig (IgV) domains [6,8,9] (Figure 1a, c). However, there is also evidence for lateral (cis) interactions on the same plasma membrane, mediated by the second and third Ig domains of Nectin-2 [10].

Nectin-1 Necl proteins are also known as SynCAMs and Cadms. All of these proteins are localized to synaptic plasma membranes, and are present on both sides of synapses, consistent with their ability to mediate homophilic adhesion [11,12]. Necl-2/SynCAM1 was first identified as a homophilic cell adhesion molecule that can induce synapse formation [11]. Necl-2/SynCAM1 mouse knockout mutants have phenotypes affecting synapse number, excitation/inhibition balance, and learning (e.g., [13,14]). Heterophilic complexes involving all Necls contribute to synaptic organization and function [12]. Necl-4 and Necl-1 mediate Schwann cell–axon interactions during myelination, and Necl-4 mutants have myelination abnormalities [15,16]. Necl-3/SynCAM3 is expressed on floor plate cells during development and interacts with Necls on commissural axons, facilitating their turning responses after they cross the floor plate [17]. The neural functions of Nectins have been studied less extensively than those of the Necls (reviewed by [3**,18]). Nectin-1...
and Nectin-3 are involved in axon–dendrite interactions. Nectin-2 is expressed on both astrocytes and neurons, and glial and neuronal degeneration was observed in Nectin-2 mutants [19].

Nectins are not restricted to the nervous system. They are localized at adherens junctions in epithelia, and function as immune modulators. Nectins and Necls are receptors for poliovirus, herpes simplex virus, and measles virus. Human Nectins are implicated in inherited diseases and cancer (reviewed by Refs [37,18]).

IgLONs
IgLONs are cell surface proteins with three Ig domains that are attached to the membrane by glycosyl-phosphatidylinositol (GPI) linkages. There are five IgLONs in mammals: Neurotrimin, Kilon/Negr1, OBCAM/OPCML, Lsamp/LAMP, and IgLON5. Each IgLON exhibits promiscuous homophilic and heterophilic binding, interacting with every member of the subfamily. IgLONs mediate trans interactions between cells in cell clustering experiments, and cis heterodimers might also exist [20].

The functions of IgLONs are not well understood. IgLONs are broadly expressed in the brain, although there are regional differences in expression levels among them. Neurotrimin and Lsamp are on growing axons and at immature synapses. IgLONs are postsynaptically localized in the mature brain [21–23]. They are released from neurons by the actions of matrix metalloproteases (MMPs). MMP inhibition reduces axonal outgrowth from cortical neurons, and cortical axons can grow on IgLON substrates. It has been suggested that the released IgLONs create a permissive substrate for axonal outgrowth in vivo [24]. IgLON overexpression can affect synapse numbers in neuronal cultures [25]. Analysis of the brain proteome showed that Lsamp is expressed on both neurons and astrocytes, but not on oligodendrocytes and microglia. Thinner axons are prematurely myelinated in Lsamp knockout mutant mice, indicating that Lsamp is a negative regulator of myelination. This regulation could occur through the actions of MMP-released Lsamp, because a surface coated with Lsamp fusion protein is repulsive to oligodendrocytes [26].

CEACAMs
The carcinoembryonic antigen (CEA) IgSF subfamily has 22 members in humans, of which 12 belong to the CEA-related cell adhesion molecule (CEACAM) group and 10 are pregnancy-specific glycoproteins (PSGs). Brain expression has been reported for two CEACAM subfamily members thus far. CEACAM2 is expressed in the hypothalamus, which controls feeding and metabolism. CEACAM2 mutants exhibit hyperphagia, and females are obese [27,28]. PSG16 is also expressed in the brain [29], but its functions are unknown. Subcellular localization of CEACAMs within neurons has not been analyzed.

CEACAMs have been primarily studied outside of the nervous system. The first subfamily member to be discovered was carcinoembryonic antigen (CEA), which corresponds to CEACAM5. This is a tumor marker that is released into serum. CEACAMs are expressed in epithelial, endothelial and in many immune cells, and they have functions in immunity and development. PSGs are secreted by placental cells into the maternal circulation (reviewed by Refs [30,31]).

CEACAM extracellular domains are composed of one to seven immunoglobulin domains, where the N-terminal domain is always a variable-type Ig (IgV), and the rest are usually constant-type Ig domains (similar to the Nectin D2 in Figure 1a). CEACAMs serve as cell adhesion molecules through homophilic and heterophilic interactions. cis Homophilic interactions have also been reported [32,33]. A comprehensive determination of all homophilic and heterophilic interactions is yet to be performed, but several lines of evidence, including cell aggregation assays, analytical ultracentrifugation and crystal structures, show that CEACAM1, CEACAM5 and CEACAM6 can create homodimers, and that CEACAM heterodimers also exist. Both homophilic and heterophilic interactions are mediated by the N-terminal IgV domains (Figure 1c). Of all tested homophilic interactions, CEACAM1 and CEACAM5 form the strongest homodimers, while CEACAM6 dimers are very weak, and CEACAM8 dimerization is too weak to occur under physiologically realistic concentrations. Instead, CEACAM6 and CEACAM8 create heterodimers with moderate affinity [34*].

Recently, an IgSF protein not belonging to the CEA family, TIM-3, was shown to interact with CEACAM1 [35*]. This heterophilic interaction is also mediated by the N-terminal IgV domains of both proteins, and regulates the T-cell inhibition activity of TIM-3, which is central to controlling autoimmunity and anti-tumor immunity.

Global IgSF interaction maps
Since IgSF proteins do not bind only to partners within the same subfamily, defining IgSF binding networks requires determination of the global interaction patterns for all IgSF proteins within a species. Interaction maps for cytoplasmic and nuclear proteins have been defined in many organisms using yeast two-hybrid screening and affinity purification. However, these techniques are usually incapable of detecting interactions among XG domains, which are typically of low affinity (micromolar or weaker) and often occur naturally between clusters of proteins on cell surfaces. Avidity effects (stronger binding caused by clustering) ensure that interactions between cells mediated by IgCAMs can be strong even when the
affinities of IgCAM monomers for their partners are very weak.

To detect interactions among XC domains in vitro, it is necessary to multimerize or cluster one or both of the partners. Multimerized proteins can form more stable complexes caused by avidity effects. Multimerization was required for detection of in vitro interactions between Dscam splice variants [36]. The AVEXIS method detects interactions between a ‘bait’ protein bound to a plate and a pentameric ‘prey’ protein in solution. AVEXIS was used to detect interactions among a group of 110 zebrafish IgSF proteins, and several new binding partners were identified in this screen [37]. It was also used to analyze interactions among 150 zebrafish proteins containing another XC domain, the leucine-rich repeat (LRR) [38].

The Extracellular Interactome Assay (ECIA) is an ELISA-like method that examines interactions between unpurified dimeric Fc fusion protein baits and pentameric alkaline phosphatase fusion protein (AP) preys in cell supernatants. To create a global interaction map for the Drosophila IgSF superfamily, all 130 IgSF proteins were expressed as both dimeric Fc and pentameric AP fusion proteins in transiently transfected Drosophila tissue culture cells, and binding between each pair of proteins was examined using the ECIA. FnIII repeat proteins and LRR proteins were also included. The Drosophila XC interactome assayed in vitro interactions among a total of 202 cell-surface and secreted (CSS) proteins (20,503 protein pairs were tested). 106 binding interactions were detected, 83 of which were new, and cross-subfamily networks involving four IgSF subfamilies were defined [1**].

Despite the success of the interactome project, most Drosophila IgSF proteins, and almost all LRR proteins, remain orphan receptors. There are several possible explanations for this, which are not mutually exclusive. First, the interactome sampled only three domain types, but more than 80 types of XC domains exist in Drosophila. The genome encodes about 1000 cell surface and secreted proteins likely to be involved in cell recognition [39]. Orphan proteins may have partners with XC domains other than IgSF, FnIII, and LRR. Second, binding of some orphans to their partners may require coreceptors. Third, interactions between some partners may be too weak to be detected by the ECIA, and detecting these may necessitate the use of higher-order multimers. For example, analysis of binding of low-affinity T cell receptors to peptide-bound MHC molecules can require the use of large clusters of MHC–peptide complexes assembled on dextran polymers (dextramers) [40]. Fourth, some proteins were expressed at very low levels in cell supernatants, and detecting interactions involving these proteins may require their purification.

The Dpr-ome

The Dpr-ome [41**] represents a subfamily of 21 2-Ig domain cell surface proteins, the Dprs, selectively interacts with a subfamily of 9 3-Ig domain cell surface proteins, the DIPs, forming a network called the Dpr-ome [1**, 41**] (Figure 2a). The only one of these 30 genes that had been previously studied was dpr1, which was identified in a behavioral screen for mutants with reduced aversion to salt [42].

In the Dpr-ome, most Dprs interact with multiple DIPs, and vice versa. DIPs are similar to the vertebrate IgLON subfamily discussed above, while Dprs have no obvious vertebrate counterparts. The Dpr-ome differs from the mammalian networks described above in that DIPs and Dprs seldom interact within their own subfamilies.

The crystal structures of Dpr-DIP complexes show that Dprs and DIPs bind via their D1 (membrane-distal) Ig domains (Figure 1c). The D1s interact with each other using the CC'C' FG faces of the immunoglobulin fold. This topology bears a strong resemblance to several other Ig-type D1-D1 complexes, including Nectin complexes (Figure 1c). The core of the Dpr–DIP interaction surface is strongly hydrophobic and contains few hydrogen bonds and no salt bridges. It is likely that Dpr-DIP binding specificity is determined by shape complementarity. Further work will be required to understand the structural ‘interaction code’ that determines why each Dpr binds to a unique subset of DIPs and vice versa. There are specific hydrophobic interface residues that correlate with binding of Dprs to particular DIPs, but their roles have not yet been tested [41**].

Each Dpr and DIP that has been examined is expressed by a unique small subset of neurons at each stage of neural development (Figure 2b–g). Remarkably, brain neurons expressing a particular Dpr are often presynaptic to neurons expressing a DIP to which that Dpr binds in vitro [41**,**43**]. These data suggest that Dpr–DIP interactions may be important for determination of synaptic connectivity patterns during development. One binding pair, Dpr11 and DIP-γ, has been genetically analyzed in published experiments. dpr11 and DIP-γ loss-of-function mutants have very similar phenotypes in both larvae and pupae, showing that binding in vitro correlates with linked functions in vivo. Analysis of these mutant phenotypes shows that this Dpr-DIP pair regulates signal transduction in larval motor neurons, has a neurotrophic function in pupal optic lobe neurons, and influences formation of synapses [41**]. DIP-γ localizes to the synaptic neuropil of the brain (unpublished results), and Dpr11 is synaptically localized to larval neuromuscular junctions [41**]. Genetic analysis of other members of the Dpr and DIP subfamilies will be required in order to...
The Dpr-ome network and expression of Dprs and DIPs in neuronal subsets.

(a) The ‘extended’ Dpr-ome. Binding interactions between Dprs and DIPs are indicated by lines. cDIP, a LRR protein, interacts with most Dprs and DIPs and with another IgSF protein, Klingon. Two additional LRR proteins, CG1504 and CG5819, are connected to DIP-γ.

(b) Expression of Dpr11 (GFP; green) and its partner DIP-γ (dsRed; magenta) in the 3rd instar larval ventral nerve cord (VNC).

(c) Expression of Dpr6 (Ci) and its partner DIP-α (Cii) (green) in the VNC. Magenta, motor neurons labeled by the OK6 driver. Arrows in B and C: cells expressing both reporters at high levels. Caret: cells expressing GFP but only low levels of dsRed. Double carets: cells expressing only dsRed.

(d)–(g). Expression of Dpr and DIP partners in the medulla of the pupal optic lobe. Each Dpr and DIP (green) is expressed in neurons projecting to different layers of the 10-layered medulla neuropil, which is labeled by the synaptic marker Brp (magenta). Layer numbers are marked on the sides of the panels; asterisks mark prominent labeled layers. Dpr11 is expressed by ‘yellow’ R7 photoreceptors, which synapse on DIP-γ-expressing Dm8 neurons in layer M6. For further information see Ref. [41**].

Develop a more complete understanding of the functions of the Dpr-ome.

Dprs and DIPs also bind to proteins with other XC domains. The ‘common DIP’ (cDIP) is a secreted LRR protein that binds to most Dprs and DIPs. Two other LRR proteins interact directly or indirectly with DIP-γ, and a cell-surface IgSF protein called Klingon binds to cDIP (Figure 2a) [1**]. An interactome that includes proteins with other types of XC domains might uncover additional interactions that would link the Dpr-ome to other adhesion and signal transduction pathways.

The Beat-Side network

A screen for mutations causing motor axon defects identified two genes, beaten path (beat) and sidestep (side), that had very similar phenotypes [44,45]. In these mutants, motor axons fail to innervate muscles because they are unable to leave their axon bundles and grow onto muscle surfaces. Beat and Side are cell-surface IgSF proteins in different subfamilies. It was later discovered that Beat is a neuronal receptor for Side on muscles, providing a satisfying explanation for their similar phenotypes [46]. Beat protein is localized to motor axon growth cones and CNS axons [44].

There are 14 members of the Beat subfamily, and 8 members of the Side subfamily. The interactome showed that 6 other Beats also bind to Side paralogs, so these two subfamilies also define an IgSF interaction network [1**]. There are 7 Beats and 4 Sides that remain orphans, however. It is possible that their binding partners were not identified in the interactome screen because their affinities are too low and/or they require coreceptors for binding.

Conclusions

The Drosophila interactome studies show that the largest IgSF networks span subfamily boundaries, and that interactions within these networks regulate synapse formation and cell fate. This suggests that a global IgSF interactome for a mammalian species would be likely to identify new...
networks that are important for nervous system development and function. It might also define pathways for modulation of immune system function that could be targeted by new therapies.

Conflict of interest
Nothing declared.

Acknowledgements
We thank Christopher Garcia, Robert Carrillo and Kazushiki Menon for discussions and R.C. and K.M. for figures in Figure 2. Work in the Zinn group was supported by National Institutes of Health (NIH) grants R37 NS28182 and U01 MH09147. Work in the Ozkan group was supported by NIH R01 NS097161.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:
● of special interest
◆ of outstanding interest
A comprehensive mass spectrometry analysis of the mouse brain proteome, analyzing the protein composition of cortical neurons, astrocytes, oligodendrocytes, and microglia. The paper also shows that the IgLON protein Lsamp is a negative regulator of myelination.


This paper shows that TIM-3 forms a heterodimeric complex with CEACAM1, and that this complex mediates T cell inhibition.


41. Carrillo RA, Ozkan E, Menon KP, Nagarkar-Jaiswal S, Lee PT, Jeon M, Birnbaum ME, Bellen HJ, Garcia KC, Zinn K: Control of synaptic connectivity by a network of Drosophila IgSF cell surface proteins. Cell 2015, 163:1770-1782. This paper describes the completion of the Dpr-ome network and the structure of Dpr-DIP complexes. It shows that each of several Dprs and DIPs is expressed by unique subsets of neurons in the larval ventral nerve cord and pupal brain. Dpr11 is expressed by a subtype of R7 photoreceptors, and its binding partner DIP-Y is expressed on a subset of Dm8 amacrine neurons, which are the primary synaptic targets for R7s. The paper also describes dpr1 and DIP-Y phenotypes in the larval neuromuscular system and pupal optic lobe.


43. Tan L, Zhang KX, Pecot MY, Nagarkar-Jaiswal S, Lee PT, Takemura SY, McEwen JM, Nern A, Xu S, Tadros W et al.: Ig superfamily ligand and receptor pairs expressed in synaptic partners in Drosophila. Cell 2015, 163:1756-1769. An RNA-seq analysis of lamina neurons in the optic lobe. It demonstrates that each of the five types of lamina neurons expresses a unique complement of Dprs. In many cases, lamina neurons expressing a particular Dpr are presynaptic to medulla neurons expressing a DIP to which that Dpr binds in vitro.

