Published in final edited form as:

Biochemistry. 2018 February 27; 57(8): 1293–1305. doi:10.1021/acs.biochem.7b01074.

Structures of the L27 Domain of Disc Large Homologue 1 Protein Illustrate a Self-Assembly Module

Agnidipta Ghosh^{*,†}, Udupi A. Ramagopal[‡], Jeffrey B. Bonanno[†], Michael Brenowitz[†], and Steven C. Almo^{*,†}

[†]Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461, United States

[‡]Biological Sciences Division, Poornaprajna Institute of Scientific Research, Sadashivanagar, Bangalore 560080, India

Abstract

Disc large 1 (Dlg1) proteins, members of the MAGUK protein family, are linked to cell polarity via their participation in multiprotein assemblies. At their N-termini, Dlg1 proteins contain a L27 domain. Typically, the L27 domains participate in the formation of obligate hetero-oligomers with the L27 domains from their cognate partners. Among the MAGUKs, Dlg1 proteins exist as homo-oligomers, and the oligomerization is solely dependent on the L27 domain. Here we provide biochemical and structural evidence of homodimerization via the L27 domain of Dlg1 from *Drosophila melanogaster*. The structure reveals that the core of the dimer is formed by a distinctive six-helix assembly, involving all three conserved helices from each subunit (monomer). The homodimer interface is extended by the C-terminal tail of the L27 domain of Dlg1, which forms a two-stranded antiparallel β -sheet. The structure reconciles and provides a structural context for a large body of available mutational data. From our analyses, we conclude that the observed L27 homodimerization is most likely a feature unique to the Dlg1 orthologs within the MAGUK family.

Graphical Abstract

^{*}**Corresponding Authors** agnidipta.ghosh@einstein.yu.edu. steve.almo@einstein.yu.edu.

The authors declare no competing financial interest.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.7b01074.

In silico assembly analyses of the wild-type and mutant L27 domain of *Drosophila* Dlg1 (Table S1), purification of wild-type and mutant DmDlg1-L27 domains (Figure S1), comparison of structures derived from two distinct crystal forms (Figure S2), a close-up view of the polar interactions at the dimer interface (Figure S3), and secondary structure predictions for the wild-type and mutant domains (PDF)

The control of cell polarity in eukaryotes is instrumental for defining epithelial cell surfaces and essential for epithelial cell morphogenesis.^{1,2} The loss of cell polarity is one of the distinctive features of epithelium-derived malignant cells.³ Epithelial cell polarity relies on the asymmetric assembly and distribution of a multitude of protein complexes, including receptors, ion channels, and their binding partners. These multimeric molecular assemblies, in part, requires modular scaffolding proteins, such as membrane-associated guanylate kinases (MAGUKs), a class of multidomain adapter proteins.⁴

Disc large 1 (Dlg1; also known as SAP97) is a prominent member of the MAGUKs and is implicated in cell polarity through the organization of multiprotein assemblies.^{4–6} Dlg1 is primarily present at the membrane–cytoskeleton interface of cell–cell junctions; however, its subcellular localization depends on its interaction partners.^{7,8} Specifically, Dlg1 creates platforms at specific plasma membrane sites for the assembly of a subset of cell adhesion molecules, receptors, or ion channels with cytoskeletal proteins and/or associated signaling components.⁴ Thus, Dlg1 exerts both scaffolding/structural and signaling roles (reviewed in ref 9).

The functional role of Dlg1 in controlling cell polarity was first documented for *Drosophila*, which encodes the *dlg* gene, the lone ortholog of mammalian Dlg1. In flies, *dlg* mutations are embryonic lethal due to disruption of epithelial cellular junctions, culminating in defects in the polarization of apical proteins and in menacing cell overgrowth.^{10,11} Loss of zygotic *dlg* in *Drosophila* causes the onset of tumor formation, which manifests by neoplastic overgrowth of imaginal disc epithelia and hyperplastic growth within larval brains.^{12,13} The ectopic expression of mammalian Dlg1, which acts as a tumor suppressor, reverses these tumorigenic phenotypes in *dlg* mutants.^{14,15} In humans, Dlg1 is a target for several viral oncoproteins, and mislocalization and progressive loss of its expression are associated with epithelial cancer progression.^{16–20} Homozygous deletion of *dlg1* (–/–) in mice causes an alteration of the distribution of polarity and cell–cell adhesion markers and leads to neonatal lethality associated with craniofacial deformities.²¹

Dlg1 includes six protein interaction modules: an N-terminal L27 domain followed by three PDZ domains, an SH3 domain, and a C-terminal inactive guanylate kinase domain.^{4,5,22} The L27 domain, first identified in the *Caenorhabditis elegans* Lin-2 and Lin-7 proteins, is

present in scaffolding proteins, most of which preferentially hetero-oligomerize.^{23,24} L27 domains have been shown to be largely unfolded in isolation and, as obligate heterodimers, form supramolecular assemblies upon interaction with their cognate partners.^{23,25–29} Using heterodimeric building blocks as fundamental assembly units, L27 domains organize into higher-order structures, such as tetramers.^{23,26,29}

Structure-guided sequence alignments of L27 domains reveal two groups: type A and type B.²³ On the basis of insightful interaction studies,²⁸ a working model for the core L27 assembly unit has been postulated where multimerization of L27 domains strongly favors heteromeric A/B assemblies, in which the positively or negatively charged amino acid side chains at specific (or so-called “keystone”) positions within the domain preclude A/A or B/B homodimerization; hence, the assembly unit is considered to be a preferential heterodimer of an A/B pair. However, the type A L27 domains of the Dlg1 clan show variability in the “keystone” positions and have been shown to interact with each other³⁰ and are therefore incompatible with the general rule of preferential A/B heterodimeric pairing. Furthermore, single-particle electron microscopy analyses, together with complementary biochemistry, suggest that Dlg1 self-associates using its type A L27 domain,^{23,31,32} a feature that clearly distinguishes Dlg1 proteins from the closely related MAGUK clan (PSD-95/SAP90, PSD-93/chapsyn-110, and SAP102).

To assess the molecular determinants for the self-assembly of Dlg1 proteins, we probed the L27 domain boundaries of *Drosophila melanogaster* Dlg1 (hereafter DmDlg1) via construct design (C-terminal truncations) and evaluated the solution behavior of the resulting gene products. Our data indicate that the resulting domain boundaries yielded soluble protein that exists as a homodimer in solution. To identify the molecular surface involved in the homotypic interaction, we determined the high-resolution crystal structure of the L27 domain homodimer in two different crystal forms. These structures reveal that the homodimer is formed via a symmetric, antiparallel organization of the two L27 subunits (monomers), each of which adopts a distinct conformation compared to those of the previously reported L27 domains of the MAGUK family members.^{23,26,27,29} All helices present in the individual subunits are involved in dimerization and form a distinctive core six-helix assembly. Additionally, the structure revealed a C-terminal β strand, which is not included in the previously reported structures of L27 heterodimers, also contributes to the homodimer interface. Structure-guided mutational analyses confirmed that the crystallographic contacts observed at the dimer interface are critical for the homomeric assembly of the *Drosophila* Dlg1-L27 domain. From a structure-guided sequence alignment, we predict that the observed dimer interface in the Dlg1-L27 domain would be conserved among the orthologs. Furthermore, the self-assembled module of the Dlg1-L27 domain described herein reconciles a large body of available mutational data and accounts for the chemical and spatial roles of these mutant amino acid residues.^{31,32}

EXPERIMENTAL PROCEDURES

Cloning and Protein Expression

DNAs encoding the L27 domain of *D. melanogaster* Dlg1 (UniProt accession number P31007, amino acid residues 1–97) were inserted in frame into pNIC28 Bsa4 (pSGC-His)

using ligation-independent cloning;³³ the gene was fused to a leader sequence encoding an TEV cleavable N-terminal His₆ tag. The inserted DNA sequence of the corresponding domain was sequenced (GENEWIZ) completely to exclude the acquisition of unwanted coding changes during DNA amplification and cloning.

Mutations were introduced into the wild-type plasmid by polymerase chain reaction with a QuickChange kit (Stratagene). The resultant plasmids were transformed into *Escherichia coli* BL21 (DE3) CodonPlus RIL (Novagen). For protein preparations, 1–2 L bacterial cultures, supplemented with 50 µg/mL kanamycin, 100 µg/mL chloramphenicol, and 100 µL/L anti-foam 204 (Sigma), were grown at 37 °C in baffled flasks to an A₆₀₀ of 2. The temperature was then decreased to 20 °C; isopropyl β-D-thiogalactoside (IPTG) was added to a final concentration of 0.5 mM, and the cultures were incubated overnight.

Protein Purification

Cells were harvested by centrifugation at 7000g and suspended in buffer containing 20 mM HEPES (pH 7.5), 500 mM NaCl, 20 mM imidazole, 0.1% IGEPAL, 20% sucrose, and 1 mM β-mercaptoethanol (BME). Cells were disrupted by sonication, and cell debris was removed by centrifugation at 45000g. The supernatants were applied to chromatography columns packed with 10 mL (5 mL for small-scale preparations) of His60 superflow resin (Clontech) that had been equilibrated with buffer A [20 mM HEPES (pH 7.5), 20 mM imidazole, 500 mM NaCl, and 1 mM BME]. The columns were washed with buffer A, and the His₆-tagged Dlg1-L27 domain was eluted with buffer B [20 mM HEPES (pH 7.5), 350 mM NaCl, 250 mM imidazole, and 1 mM BME]. The N-terminal His₆ tag from the protein was removed by overnight digestion at 4 °C with the TEV protease at a 2000:1 DmDlg1-L27:TEV ratio. The tag-free protein was then separated from the His₆ tag by a Superdex 75 (16/60) gel filtration column in buffer containing 20 mM HEPES (pH 7.5), 350 mM NaCl, and 1 mM BME. Fractions containing the peak of the protein were assessed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), pooled, concentrated to 25 mg/mL, and buffer exchanged with 20 mM HEPES (pH 7.5), 150 mM NaCl, and 5 mM DTT using a 3 kDa Amicon Ultra-15 centrifugal filter device (Millipore).

Analytical Size-Exclusion Chromatography

A Superdex S75 (10/300 GL) column equilibrated in 20 mM HEPES (pH 7.5), 150 mM NaCl, and 5 mM DTT was used to analyze the elution profile of DmDlg-L27. Under this buffer condition, the column was calibrated with Bio-Rad Gel Filtration Standard (catalog no. 151–1901), which contains a mixture of thyroglobulin (670 kDa), bovine γ-globulin (158 kDa), chicken ovalbumin (44 kDa), horse myoglobin (17 kDa), and vitamin B12 (1.35 kDa). The column flow rate was 1 mL/min.

Analytical Ultracentrifugation

The sedimentation velocity experiments were conducted at 58000 rpm and 20 °C using double-sector centerpieces following the absorbance of the protein at either 280 or 230 nm. One hundred scans were taken over the course of the experiment; approximately one-third of them were analyzed by the time-derivative method using DCDT + v2.4.2.^{34,35} The calculated mass of the wild-type DmDlg1-L27 monomer is 11414.1 Da. Values of the partial

specific volume ($\bar{v} = 0.748$), density (1.00556), and viscosity (1.03 cP) were calculated using Sednterp [B. Hayes, T. Laue, and J. Philo, Sedimentation Interpretation Program, University of New Hampshire, Durham, NH (<http://www.jphilo.mailway.com>)] from the amino acid and buffer compositions and used to correct the resolved parameters to standard conditions (20, W). Sedimentation parameters were calculated from the refined structure of the DmDlg1-L27 domain dimer [Protein Data Bank (PDB) entry 4RP5 (Table 1)] using HYDROPRO,³⁶ which takes into account both the size and the shape of the modeled assembly.

Protein Crystallization and Crystal Harvesting

Initial screening of the DmDlg-L27 domain was performed using 800 nL (protein:mother liquor ratio of 1:1) crystallization drops at a concentration of 18 mg/mL with a Crystal Gryphon (Art Robbins Instruments) and utilizing the MCSG sparse matrix crystallization suite (Microlytic). Initial microcrystals obtained from the sparse matrix screening were further optimized with several rounds of grid screening by exploiting a Formulater (Art Robbins Instruments) and/or additive screening and analyzed at home source using a rotating anode X-ray generator (RAXIS-IV, Rigaku) for diffraction quality and identification of an optimal cryo-preservative. Crystals exhibiting a high quality diffraction patterns were eventually obtained in sitting drops by vapor diffusion against well solutions that contained 20% PEG 3350, 0.2 M sodium formate, and 3% 1,8-diaminooctane (or 20% PEG 3350, 0.2 M sodium formate, and 3% 1,5-diaminopentane dihydrochloride). Crystals appeared after 2 days at 19 °C and reached maximum dimensions of $20\ \mu\text{m} \times 50\ \mu\text{m} \times 50\ \mu\text{m}$ in 1 week. For data collection at the synchrotron, crystals were cryo-preserved by addition of 10% ethylene glycol to the mother liquor prior to flash-cooling in liquid nitrogen using an appropriately sized microloop. Crystals obtained in the presence of 1,8-diaminooctane exhibited diffraction consistent with space group $P2_12_12_1$ (orthorhombic crystal form) and in the presence of 1,5-diaminopentane dihydrochloride exhibited diffraction consistent with space group $P2_1$ (monoclinic crystal form) (Table 2). Potassium-derivatized crystals of the DmDlg1-L27 domain (K^+) were obtained by incubating native crystals, which were grown in the presence of 1,8-diaminooctane, in a mother liquor overnight that contained 20% PEG 3350, 0.25 M potassium formate, and 10% ethylene glycol.

Data Collection

Cryo-preserved crystals were screened on beamline X29A at the National Synchrotron Light Source of Brookhaven National Laboratory. The best crystal exhibited diffraction to d spacing of $\sim 1.36\ \text{\AA}$. Data were collected on a ADSC Quantum 315 CCD detector at a wavelength of either $1.075\ \text{\AA}$ (for native) or $0.9798\ \text{\AA}$ (for K^+) with a 1° oscillation and at a 150–200 mm crystal-to-detector distance.

Data Processing, Structure Determination, Model Building, Refinement, and Analysis

Data from single crystals were collected using CBASS,³⁷ and integrated and scaled using HKL-3000³⁸ and AIMLESS.³⁹ The boundaries of highest-resolution shells for refinement were determined on the basis of the $\text{CC}_{1/2}$ values reported by AIMLESS.^{39,40} Initial phases of the DmDlg1-L27 domain were determined from a potassium-derivatized crystal using

SHELX.⁴¹ The phases were further improved by DM⁴² and used as input into ARP/wARP,⁴³ which generated a preliminary model consisting of 153 of the 194 amino acid residues in the dimer present in the asymmetric unit. The model was manually checked and corrected in COOT,⁴⁴ including adding alternate conformations where appropriate and refined using PHENIX⁴⁵ utilizing individual anisotropic temperature factors and occupancy refinement; the final round of refinement was performed in REFMAC5.⁴⁶ Analyses of the structures were performed in COOT⁴⁴ and MolProbity.⁴⁷ The final models of the domain lack the extreme N-terminus (the KCl-soaked orthorhombic crystal form lacks residues 1–4 in protomer 1/chain A and residues 1–5 in protomer 2/chain B; the native orthorhombic crystal form lacks residues 1–4 in both protomers; the native monoclinic crystal form lacks residues 1–4 in protomer 2/chain B) and C-terminus (the KCl-soaked orthorhombic crystal form lacks residue 97 in both protomers; the native orthorhombic crystal form lacks residues 95–97 in protomer 1/chain A and residue 97 in protomer 2/chain B; the native monoclinic crystal form lacks residues 96 and 97 in both protomers) of the protomers due to disorder. All models exhibited excellent geometry with no residues in disallowed regions of the Ramachandran plot.⁴⁸ Crystallographic statistics are listed in Table 2. Structure factors and coordinates have been deposited in the RCSB Protein Data Bank as entries 4RP3 (KCl-soaked in orthorhombic), 4RP4 (native in orthorhombic), and 4RP5 (native in monoclinic).

RESULTS AND DISCUSSION

The DmDlg1-L27 Domain Is a Dimer in Solution

The N-terminal segment of *D. melanogaster* Dlg1 protein (amino acids 1–97, 11.4 kDa) encompassing the L27 domain was expressed in *E. coli* and purified to homogeneity (Figure S1A; see below and Experimental Procedures). The domain was well-behaved throughout purification and eluted as a monodisperse peak from a size-exclusion chromatography column. Neither proteolysis nor aggregation was evident at room temperature (~21 °C) for at least 48 h [judged by centrifugation at 21000*g* and SDS–PAGE (Figure S1B)]. The domain (DmDlg1-L27) elutes from a calibrated analytical size-exclusion column [Superdex S75 (Experimental Procedures)] as a discrete peak with an estimated molecular weight of 23 kDa (Figure 1), which indicates that the domain is a dimer in solution at a physiological salt concentration (150 mM NaCl). The dimeric state of the domain was confirmed by sedimentation velocity–analytical ultracentrifugation (SV–AUC) analyses (Experimental Procedures); the protein sediments as a single homogeneous species with sedimentation and diffusion coefficients that correlate closely with those calculated from the crystal structure and whose ratio, S/D , yields the dimeric molecular weight within experimental error (Table 1). The dependence of S_{20} on protein concentration is consistent with DmDlg1-L27 being a stable dimer over the concentrations that were analyzed (see below). The stability of the dimer is highlighted by the close agreement between the measured sedimentation parameters and the sedimentation parameters calculated from the crystal structure (Table 1, rows 1–4). These data support a model in which the structure of the DmDlg1-L27 dimer shown in Figure 2A represents the oligomeric state of the protein at low micromolar concentrations.

Structural Overview of the Dlg1-L27 Homodimer

The purified L27 domain of DmDlg1 crystallized in two independent forms (Experimental Procedures). Molecular replacement attempts using coordinates available for multiple L27 domains were unsuccessful. Single-wavelength anomalous dispersion (SAD) techniques with a selenomethionine-labeled sample were deemed inappropriate because of the presence of a lone N-terminal methionine (Figure 2, sequence alignment). Experimental phases were obtained from SAD data collected on a potassium-derivatized orthorhombic crystal (Experimental Procedures and Table 2). The structure was refined to 1.36 Å with R_{work} and R_{free} values of 15.0 and 20.6%, respectively. Additionally, two native structures were obtained [orthorhombic crystal form, 1.42 Å and R_{work} and R_{free} values of 15.0 and 19.8%, respectively; monoclinic crystal form, 1.85 Å and R_{work} and R_{free} values of 18.4 and 21.3%, respectively (Table 2)]. The asymmetric units in these two crystal forms are nearly identical and can be superimposed with a root-mean-square deviation (RMSD) of 0.5 Å over 151 aligned C α atoms (Figure S2A). The termini, however, were better defined in the observed electron density for the monoclinic crystal form [PDB entry 4RP5 (Table 2)]; therefore, structural interpretation for this work is based on the coordinates arising from the monoclinic crystal form, unless stated otherwise.

Monoclinic crystals of DmDlg1-L27 contain two molecules in the asymmetric unit (Figure 2A) approximately related by a 2-fold noncrystallographic symmetry axis, forming a dimer. The homo-dimer observed in the crystal structure is in agreement with the observed solution behavior (Figure 1 and Table 1). Each subunit consists of three helical elements [α 1– α 3 (Figure 2A)] connected by loops. The helical elements together with the interhelical segments encompass amino acids 10–84. The helices in each subunit are organized in an antiparallel fashion. The two subunits wrap around each other and form a compact six-helix assembly, shielding an extensive hydrophobic core (Figures 2 and 3). In the structure, helix α 1 (amino acids 7–21) interacts with α 2 helices (amino acids 24–60). In addition, helix α 2 interacts with its α 2 counterpart and with the two α 3 helices (amino acids 66–87). Each subunit also includes a C-terminal β strand (amino acids 88–92). The strands in the dimer interact with each other in an antiparallel manner (Figures 2A and 3).

Structural Symmetry in the L27 Domain Dimer

The overall tertiary structures of the two subunits in the asymmetric unit are similar to each other. All three helices (α 1– α 3) of each subunit overlay remarkably well on their counterparts. Aligning 73 C α atoms spanning amino acids 11–84 (encompassing the core helical region of the subunits) results in an RMSD of 0.51 Å (Figure 2A and Figure S2B). However, superposition of all atoms in the refined structure (PDB entry 4RP5) resulted in an RMSD of 1.40 Å. The major structural differences contributing to the higher overall RMSD between the two subunits are present in the C-terminal tail of the domain including the β strand. The tails of the two subunits show the largest deviation with respect to each other (~6.5 Å) when the core helical regions are aligned (see below). On the basis of the nearly symmetric dimeric arrangement and agreement of the superposition of the helical cores, we surmise that the six-helix assembly observed in the L27 domain dimer of DmDlg1 is a rigid module necessary for self-assembly.

Interface Observed in the Dlg1-L27 Domain Homodimer

Homodimerization of DmDlg1-L27 gives rise to two distinct interfaces, interface 1 and interface 2 (Figure 3), which together bury 2000 Å² of total solvent accessible surface area. Interface 1 utilizes residues from all three helices of both subunits and contributes a total of 1600 Å² of buried area to the interface. These contacts are dominated by hydrophobic interactions, as most of the hydrophobic residues present in the helical assembly are buried in the interface. With the exception of three sets of polar interactions (Figure S3) (i) between the carbonyl oxygen of Gln6 and the side chain nitrogen of Gln53, (ii) between the two side chain nitrogens of Arg36 (NH₁ and NH₂) and the carbonyl oxygen of Leu75 and Thr79 Oγ1, respectively, and (iii) for a salt bridge between Arg44 and Asp51, all interactions between the two subunits are hydrophobic in nature. All residues with bulky hydrophobic side chains present in the helical assembly [Tyr18, Phe41, Phe45, Phe55, Tyr56, and Trp81 (Figures 2B and 3)] are engaged in the formation of the interface. Only seven residues [Leu22, Ile32, Ile38, Leu50, Leu58, Leu61, and Ile77 (Figure 2B)], which are buried in the core of the each subunit, do not participate in the composite interface. A subset of the key residues (Leu15, Tyr18, Leu30, Ile34, Ile40, Phe41, Leu45, Leu49, Ile52, Tyr56, Leu61, Thr71, Thr74, and Trp81) that are involved in interface 1 play dual roles in the assembly by concomitantly participating in dimerization and stabilization of the individual subunit via intramolecular interactions. With the exception of Leu60, which resides in the connecting loop between helices α2 and α3, all participating residues in interface 1 come exclusively from the helical elements of each subunit (Figures 2A and 3).

Secondary structure predictions suggest that the nine C-terminal residues spanning amino acids 86–94 have propensities to form either helical or β strand secondary structure (Figure S4). However, in the crystal structure, these residues form a β strand in both subunits, and the strands pack against each to form an antiparallel β sheet (Figures 2 and 3). This arrangement gives rise to interface 2 (Figure 3), expanding the buried surface area by an additional 400 Å². This second interface is dominated by backbone interactions between the residues present in C-terminal tail of the domain. On the basis of the interactions observed at interface 2, we suggest that homodimerization shifts the secondary structure propensity of the region (i.e., amino acids 88–92) to favor strand, contributing to the stability of the dimer via formation of the sheet. Reinforcing this notion, our attempts to obtain homogeneous, monodisperse protein preparations required the inclusion of the C-terminal tail (Figure S1A; see below). We further predict that the calculated helical propensity of the tail might be important for the heteromeric interaction of the protein with its cognate partners with L27 domains. However, this hypothesis has yet to be validated; to date, the C-terminal tail is absent in all reported structures of the domain bound to its cognate partners (Figure 2B; see below).^{23,27} With the exception of the less conserved C-terminal tail (Figure 2B), the dimer interface in DmDlg1-L27 is highly conserved (58% identical sequence over 82 amino acids) among the Dlg1 orthologs, suggesting that the L27 domains of the Dlg1 clan exploit an analogous surface for homodimerization.

Mutational Analyses Complement the Structure of the Dlg1-L27 Homodimer

Single-amino acid substitutions in the L27 domain of full-length rat Dlg1 identified 11 side chains as being critical for homomeric interaction: Leu14, Leu15, Tyr18, Glu26, Asp27,

Ile34, Val37, Ile38, Ser39, Phe41, and Gln42.³¹ However, substitution of glycine for Leu14, Leu15, and Tyr18 and substitution of alanine for Glu26, Asp27, and Gln42 were reported to result in increased Stokes radii in analytical size-exclusion chromatography compared to that of the wild-type protein, implying that these substitutions promote higher-order oligomerization beyond a dimer in rat Dlg1. Conversely, substitution of glycine for Ile34, Val37, and Ile38 and alanine for Phe41 results in Stokes radii that are decreased to values indicative of monomeric Dlg1.³¹ Eight of these residues are conserved among the Dlg1 orthologs (see alignment in Figure 2B). In the crystal structure of the DmDlg1-L27 dimer, these residues are located at the interface where they are predicted to directly mediate dimerization (Figures 2B and 3).

In the crystal structure reported in this work, the side chains of Leu14, Leu15, and Tyr18 are part of helix $\alpha 1$ and are spatially embedded in the composite hydrophobic environment created by the two subunits (Figure 3). We predict that glycine substitutions at these positions could potentially alter the helical conformation of helix $\alpha 1$, thereby modifying the observed interface for DmDlg1-L27 by exposing the buried hydrophobic residues and allowing interaction with other monomeric subunits or dimers. Secondary structure predictions suggest that L14G and L15G substitutions would induce random coil in helix $\alpha 1$, and substitution of Y18G further extends to predicted coil boundary farther (Figure S4B,C). We also analyzed the effect of these mutations using the Protein Interfaces, Surfaces, and Assemblies (PISA) tool at the European Bioinformatics Institute.⁴⁹ Calculation of the energy of association of the interfaces results in a slightly unfavorable increase in G for these mutants compared to that of the wild-type protein (see column 3 of Table S1), indicative of potential alteration in the observed interface.

To test these predictions, we expressed and purified L14G, Y18G, and L14G/L15G/Y18G *Drosophila* proteins to homogeneity (Figure S1C) and probed their solution behaviors using analytical size-exclusion chromatography (Experimental Procedures). Under a physiological salt condition (150 mM NaCl), both L14G and Y18G substituted L27 domains elute in a manner comparable to that of the wild-type domain except for slight peak broadening toward higher-order oligomers (Figure 4A). In contrast, the L14G/L15G/Y18G triple mutant migrated completely as a higher-order oligomer; SV–AUC analyses show that 85% of this protein sediments as an octamer with an estimated molecular weight of ~80 kDa (Table 1 and Figure 5A).

In *Drosophila*, residue 26 is a glutamine (glutamic acid for RnDlg1); the side chains of Gln26 and Asp27 point away from the hydrophobic face of the interface. Our structural interpretation indicates that disruption of this polar surface by introduction of individual alanine substitutions or by combination (Q26A/D27A) at these positions is likely to introduce an additional hydrophobic patch for a homomeric interaction where another potential dimer (or a monomeric subunit) could potentially interact. While the Q26A substitution in DmDlg1-L27 is not predicted to alter the energy of association during dimer formation, the D27A substitution decreases the calculated G (Table S1, column 3), which is suggestive of stabilizing multimerization via an additional surface of interaction. Analytical size-exclusion chromatography revealed that introduction of the Q26A mutation into the domain initiated oligomerization beyond the dimer [evident from peak broadening

(Figure 4B)], and the Q26A/D27A double mutation resulted in formation of a higher-order species (Figure 4B). This behavior of the double mutant was confirmed by SV–AUC analyses, which revealed that ~14% of the mutant protein forms an assembly that is larger than a dimer (Table 1 and Figure 5A).

In DmDlg1, Ser39 and Gln42 of RnDlg1 are replaced by an Arg and a Lys, respectively. These two side chains are one helical turn from each other and are solvent-exposed. Replacement of this charged surface via alanine substitutions could possibly create a secondary hydrophobic interaction site allowing an additional means of self-assembly. Energy calculations of the interface indicate the R39A mutation indeed favors homotypic interactions, as gauged by a decreased ΔG , while K42A does not elicit any effect (Table S1, column 3). The elution profile during analytical size-exclusion chromatography revealed that introduction of either the R39A mutation or the R39A/K42A double mutation shifted the dimeric nature of the domain toward higher-order oligomers (Figure 4C).

Ile34, Val37, and Ile38 are part of helix $\alpha 2$ and are deeply embedded in the dimer interface (Figures 2B and 3). Glycine substitution at these positions is likely to impose structural distortion on the helix as predicted by the secondary structure prediction algorithms (Figure S4D,E); we surmise that the putative structural distortion poses a significant barrier for $\alpha 2$ to pack into the observed six-helix assembly to facilitate homodimerization. While energy calculations of the dimer interface for I34G and I38G show virtually no change in ΔG when compared to that of the wild type, the V37G substitution resulted in an increase in ΔG , which is indicative of the potential rearrangement in this region of the interface (Table S1). Introduction of I34G and I34G/V37G substitutions into the domain did not alter the elution profile of the dimer (Figure 4D). However, the dimer is destabilized in the I34G/V37G/I38G mutant, which is evident from the extended trailing edge of the elution peak profile in analytical size-exclusion chromatography (Figure 4D). The decreased sedimentation coefficient and the fact that the sedimentation coefficient increases with protein concentration as observed by SV–AUC indicate a destabilized dimer undergoing a reversible monomer–dimer equilibrium at the protein concentrations used in the experiment (Table 1 and Figure 5).

Phe41 (also located in helix $\alpha 2$) from one subunit is packed against Ile52 and Trp81 from the other subunit; we predict that the F41A substitution could yield significant rearrangements in this region by potentially creating a space into which the Trp81 side chain could move into and make the subunits incompatible for dimerization. Analysis of the surface reveals that the F41A substitution predicts no destabilization of the dimer but decreases the total buried surface area (Table S1). However, the behavior of the F41A mutant is comparable to that of the I34G/V37G/I38G mutant [gauged by size-exclusion chromatography and SV–AUC (Figures 4D,E and 5)] as described above, consistent with destabilization of the dimer in solution.

Alanine scanning also identified 11 amino acids within the L27 domain of Dlg1, which are nonessential in the context of homo-oligomerization, despite the fact that most of them are conserved in the Dlg1 clan.³¹ Seven of the nonessential residues lie outside of the observed interface: Lys5, Asp7, Arg10, His13, Glu17 (Asp for DmDlg1), Ser23, and Arg28. Thr8 in

rat Dlg1 is replaced with an alanine in *Drosophila*; the short side chain occupies a hydrophobic patch in the interface created by the second subunit [Phe46, Leu49, and Leu50 (Figure 3)]. We suggest that a threonine at this position could also play a potentially similar role and is compatible with the dimer interface. Consistent with our hypothesis, analyses of the interface indicate that threonine substitution at this position (A8T) has no effect in the observed interface of the DmDlg-L27 dimer (Table S1); this notion was further corroborated by the elution profile during size-exclusion chromatography, which revealed that the A8T substituted domain, similar to the wild type, is a dimer in solution (Figure 4F).

In the dimer, Arg28 does not participate at the interface; however, it interacts with Gln47, Asp51, and Glu54 from a crystallographic symmetry mate. Amino acid Ser32 in rat Dlg1 is replaced by a Leu in *Drosophila*; this side chain is engaged in an interaction with a symmetry mate and is located distal to the dimer interface. Nonessential Asn44 of rat Dlg1 is replaced by an Arg in DmDlg1; although this position is not conserved, in the crystal structure the side chain participates in dimerization [see the description of the dimer interface (Figures 2B and 3)]. It is possible that while this position may not be crucial for homodimerization in some of the Dlg1 orthologs, it might be important for the observed homomeric assembly of *Drosophila* Dlg1. Although this notion is consistent with the fact that the R44A substitution significantly reduces the total buried surface area of the observed dimer interface (Table S1), in solution the mutant showed a propensity to form higher-order oligomers (Figure 4F).

L27 Domain Homodimerization Is Unique to the Dlg1 Clans

On the basis of similarities observed in multiple-sequence alignments, L27 domains are classified into two groups: type A and type B.^{23,26} Generally, proteins with a single L27 domain are grouped under type A, and proteins with multiple (mostly tandem) L27 domains belong to type B. Structural characterization of L27 domain assemblies using X-ray crystallography^{27,29} and nuclear magnetic resonance^{23,26} has revealed that multimerization favors heteromeric assemblies in which a type A class protein preferentially pairs with a type B class. Structure-guided sequence alignment of the cognate partners in these assemblies identified so-called “keystone” positions, which dictate structural asymmetry among the partners and in turn drive hetero- over homo-oligomerization.²⁸ Within the type A class, these keystone positions are typically occupied by conserved positively charged residues [Arg, His, or Lys (Figures 2B and 6)], which interact with conserved negatively charged side chains (Asp or Glu) from their cognate type B partners. Charge reversal substitutions at these positions of individual cognate partners (K29D in *C. elegans* type A Lin-7A or D27K in type B human Lin-2) led to destabilization of the heteromeric assembly and induced homomeric interaction (A/A or B/B pairs). Importantly, heteromerization (A/B pair) was restored when the charge reversal substitutions were simultaneously introduced into the cognate pairs.²⁸

In the type A L27 domain of Dlg1 orthologs, the positively charged keystone position is naturally occupied by an alanine (Ala33 in *Drosophila*) or a serine (Ser33 in human, mouse, and rat) (Figures 2B and 6A). In the crystal structure of the DmDlg1-L27 homodimer, C β of Ala33 from one of the subunits is tightly and reciprocally packed in a hydrophobic

environment created by the Thr71, Thr74, and Leu75 side chains from the other subunit (Figure 3). This intimate packing at the interface suggests that a positively charged extended side chain at this position would likely interfere with the homodimeric packing and thereby interfere with their self-assembly; however, the short side chains (alanine or serine) at this position are compatible with both homodimeric (self) and heterodimeric assemblies of Dlg1 orthologs with their cognate type B partners. Indeed, the S33A substitution in the context of full-length rat Dlg1 protein favors self-assembly, as compared to the case for the wild type.³¹ Our interpretation of homo- and/or heteromeric assembly compatibilities among the Dlg1 orthologs is further strengthened by two structural studies. (i) In the crystal structure of a heterotrimeric assembly of the L27 domains from hDlg1, hMPP7, and hMals3 (Figure 7; PDB entry 3LRA), the Ser33 side chain of hDlg1 sits in a hydrophobic environment created by hMPP7,²⁷ and (ii) in the solution structure of a heterodimeric L27 assembly between rat Dlg1 and rat Lin-2 (Figure 7; PDB entry 1RSO), Ser33 is surrounded by a hydrophobic environment provided by Lin-2,²³ both of which have packing characteristics remarkably similar to those of the analogous Ala33 of the DmDlg1-L27 dimer (Figure 3). Although these reported structures describe heteromeric assemblies of L27 domains, on the basis of our comparative analyses, we surmise that the natural substitution by a small side chain (e.g., alanine or serine) at this keystone position results in an increased propensity for self-assembly in Dlg1-L27 domains. Consonant with our hypothesis, among the L27 domain-containing proteins studied thus far, only Dlg1 clans are shown to self-assemble (i.e., homodimerize) utilizing the L27 module.^{31,32}

The L27 Domain of Dlg1 Compared to L27 Domains from Heteromeric Assemblies

To understand the underlying L27 domain homomeric assembly mechanism, our initial goal was to obtain soluble protein corresponding to a “standalone” self-assembly module of *Drosophila* Dlg1 (i.e., monomer). However, our attempt to isolate the minimal L27 domain (amino acids 1–63) of DmDlg1, by analogy to the available structures of human and rat Dlg1, resulted in aggregated protein [construct 1 (Figure S1A)]. Notably, the structures mentioned above involving the rat Dlg1-L27 domains arose from engineered single-chain construct proteins in which the minimal Dlg1-L27 domains were fused to their cognate partners via flexible linkers; for example, the structure of the rat L27 heterodimer of Dlg1 and Lin-2 (PDB entry 1RSO) was derived by connecting Dlg1-L27 (residues 1–67) with Lin-2 (residues 329–400) by a Gly-Ser linker.^{23,27} We therefore explored revised domain boundaries for self-assembly of DmDlg1.

In addition to construct 1 (amino acids 1–63), we generated three additional constructs (Figure S1A): construct 2 that encompasses amino acids 1–83, construct 3 encompassing amino acids 1–88, and construct 4 that includes amino acids 1–97. Constructs 2 and 3 resulted in insoluble materials. However, construct 4 produced a well-behaved soluble protein, allowing for the evaluation of self-assembly (Figure S1A; see the solution behavior of the domain). In brief, during the course of these studies, we successfully redefined the minimal self-assembly module for Dlg1 orthologs.

The L27 domains characterized to date are primarily composed of three helices, $\alpha 1$ – $\alpha 3$ (Figures 2A and 7). In the previously reported human and rat Dlg1-L27 domain structures,

^{23,27} helix $\alpha 1$ (five turns, 17 residues) is followed by a tight loop (seven residues in the human form and five residues in the rat form), which leads into helix $\alpha 2$ (four turns, 15 residues), positioning it approximately antiparallel to $\alpha 1$, followed by another short linker (four and two residues in the human and rat forms, respectively) that positions helix $\alpha 3$ (three turns, 11 residues) at approximately 60° relative to $\alpha 2$. In the context of the DmDlg1-L27 homodimer, the extents of the helical elements differ strikingly when compared to the known L27 domain structures, including the human and rat Dlg1 structures mentioned above (Figures 2B and 7A–C). While the length and location of helix $\alpha 1$ in the individual subunits of the DmDlg1-L27 domain are comparable to those of the previously reported L27 domain assemblies, helix $\alpha 2$ in the homodimer is much longer, encompassing 37 residues [10 turns (Figures 2A,B and 7A)]. As a result, the residues contributing to $\alpha 2$ and $\alpha 3$ of the human and rat Dlg1-L27 domains are merged into $\alpha 2$ in DmDlg1-L27. In the homodimer, the additional C-terminal DmDlg1 segment (residues 64–96), which is not included in the human and rat Dlg1-L27 domain constructs,^{23,27} forms the additional $\alpha 3$ helix (five turns, 17 residues) and strand $\beta 1$ (five residues). We consider this C-terminal segment to be an important part of the self-assembly-promoting region (see below). The three helices ($\alpha 1$ – $\alpha 3$) of the DmDlg1-L27 domain are organized in an antiparallel fashion and upon dimerization form a six-helix assembly (Figures 2A and 7). On the basis of the extensive interactions observed at the dimer interface (Figure 3), the solvent-exposed surface area ($\sim 2000 \text{ \AA}^2$) buried upon dimerization of the DmDlg1-L27 domain (see description of the interface), and the supporting biochemical evidence of the dimerization presented here [see the mutational analyses of the dimer interface (Figures 4 and 5)], we postulate that the observed extension of $\alpha 2$, the formation of *de novo* helix $\alpha 3$ and strand $\beta 1$, and the concomitant formation of the six-helix assembly by the two DmDlg1-L27 subunits are requirements for self-assembly. This notion is reinforced by the two following observations. (i) The redefined self-assembly competent Dlg1-L27 domain required inclusion of amino acids 1–97 [see the section describing the solution behavior of the domain (Figure S1A)], and (ii) the structurally characterized shorter Dlg1-L27 domain segments for human (residues 1–65; PDB entry 3LRA) and rat (residues 1–67; PDB entry 1RSO) were fused to their cognate partners to obtain soluble material. Moreover, among the Dlg1 orthologs, a high degree of sequence conservation within the six-helix assembly, including the segments analogous to $\alpha 1$ – $\alpha 3$ and $\beta 1$ in DmDlg1-L27 (Figure 2B), suggests that the observed helical assembly is a feature exploited by Dlg1 orthologs during self-assembly.

In all of the heteromeric assemblies of L27 domains currently known, the two N-terminal helices ($\alpha 1$ and $\alpha 2$) of the cognate partners are organized in an antiparallel fashion to interact with their counterparts and form a four-helix bundle^{26,29} (Figures 2B and 5). In these assemblies, the C-terminal $\alpha 3$ helices are oriented parallel to the $\alpha 2$ helices and are involved in antiparallel homomeric interactions (Figure 2B), which brings two four-helix bundles into a tetrameric assembly (Figure 7D,E). In the heteromeric assembly of the rat Dlg1-L27 domain with its partner from rat Lin-2, the C-terminal part of the domain, in addition, folds into a β strand, and both $\alpha 3$ and $\beta 1$ take part in tetramerization²³ (Figure 7C).

The six-helix assembly, which is composed of a highly conserved region of Dlg1 orthologs (Figure 2B), observed in the DmDlg1 homodimer is, thus far, distinctive among the structurally characterized L27 domain assemblies.^{23,26,27,29} A motif search using the Dali

Server (<http://ekhidna.biocenter.helsinki.fi>) did not detect similarities with any of the structurally characterized L27 domain assemblies. Interestingly, the search yielded similarities between the subunits and pneumococcal surface protein A, which inhibits the bactericidal activity of lactoferrins to evade cell-autonomous immunity.⁵⁰

CONCLUSION

We have biochemically redefined the self-assembly module for Dlg1 proteins and presented high-resolution structures of the self-assembled L27 domain from *Drosophila* Dlg1 derived from two distinct crystal forms. These structures reveal that the Dlg1 proteins utilize all structural elements present in the L27 domain to self-assemble and delineate the underlying molecular determinants for homomeric interactions observed for Dlg1 orthologs.

These structures also explain how natural substitutions at “keystone” positions direct Dlg1-L27 domain compatibility for both homomeric and heteromeric interactions with itself and cognate partners, respectively, a feature that distinguishes Dlg1 from other L27 domain-containing MAGUK proteins. These structures further reconcile a large body of biochemical data on the self-assembly of Dlg1 proteins by providing spatial and atomic contexts for the conserved amino acid residues.

The most striking feature observed in the Dlg1-L27 homodimer is the formation of a distinctive six-helix assembly, which shields a large conserved hydrophobic surface during self-assembly. In the context of heteromeric assemblies, L27 domains typically form a four-helix bundle. The structure of the homodimer suggests that Dlg1 proteins employ two distinct modes: one for self-assembly and another for heteromeric assembly.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors thank Dr. William Weis, School of Medicine at Stanford University (Stanford, CA), for his valuable comments on the manuscript. This work was in part made possible by Center for Synchrotron Biosciences Grant P30-EB-009998 from the National Institute of Biomedical Imaging and Bioengineering (NIBIB). Use of the National Synchrotron Light Source, Brookhaven National Laboratory, was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract DE-AC02-98CH10886. The Einstein Crystallographic Core X-ray diffraction facility is supported by National Institutes of Health Shared Instrumentation Grant S10 OD020068, which the authors gratefully acknowledge.

Funding

This work has been partly supported by National Institutes of Health Grant U01 GM094662 and GM094663.

REFERENCES

- (1). Schock F, and Perrimon N (2002) Molecular mechanisms of epithelial morphogenesis. *Annu. Rev. Cell Dev. Biol* 18, 463–493. [PubMed: 12142280]
- (2). Yeaman C, Grindstaff KK, Hansen MD, and Nelson WJ (1999) Cell polarity: Versatile scaffolds keep things in place. *Curr. Biol* 9, R515–517. [PubMed: 10421564]
- (3). Banks L, and Humbert PO (2008) On the guardians of polarity and the disorientation of cancer. *Oncogene* 27, 6876–6877. [PubMed: 19029930]

- (4). Funke L, Dakoji S, and Brecht DS (2005) Membrane-associated guanylate kinases regulate adhesion and plasticity at cell junctions. *Annu. Rev. Biochem* 74, 219–245. [PubMed: 15952887]
- (5). Lue RA, Marfatia SM, Branton D, and Chishti AH (1994) Cloning and characterization of hdlg: the human homologue of the Drosophila discs large tumor suppressor binds to protein 4.1. *Proc. Natl. Acad. Sci. U. S. A* 91, 9818–9822. [PubMed: 7937897]
- (6). Olsen O, Funke L, Long JF, Fukata M, Kazuta T, Trinidad JC, Moore KA, Misawa H, Welling PA, Burlingame AL, Zhang M, and Brecht DS (2007) Renal defects associated with improper polarization of the CRB and DLG polarity complexes in MALS-3 knockout mice. *J. Cell Biol* 179, 151–164. [PubMed: 17923534]
- (7). Muller BM, Kistner U, Veh RW, Cases-Langhoff C, Becker B, Gundelfinger ED, and Garner CC (1995) Molecular characterization and spatial distribution of SAP97, a novel presynaptic protein homologous to SAP90 and the Drosophila discs-large tumor suppressor protein. *J. Neurosci* 15, 2354–2366. [PubMed: 7891172]
- (8). Lue RA, Brandin E, Chan EP, and Branton D (1996) Two independent domains of hDlg are sufficient for subcellular targeting: the PDZ1–2 conformational unit and an alternatively spliced domain. *J. Cell Biol* 135, 1125–1137. [PubMed: 8922391]
- (9). Walch L (2013) Emerging role of the scaffolding protein Dlg1 in vesicle trafficking. *Traffic* 14, 964–973. [PubMed: 23829493]
- (10). Hough CD, Woods DF, Park S, and Bryant PJ (1997) Organizing a functional junctional complex requires specific domains of the Drosophila MAGUK Discs large. *Genes Dev.* 11, 3242–3253. [PubMed: 9389655]
- (11). Bilder D (2001) PDZ proteins and polarity: functions from the fly. *Trends Genet.* 17, 511–519. [PubMed: 11525834]
- (12). Stewart M, Murphy C, and Fristrom JW (1972) The recovery and preliminary characterization of X chromosome mutants affecting imaginal discs of *Drosophila melanogaster*. *Dev. Biol* 27, 71–83. [PubMed: 4621757]
- (13). Gateff E (1978) Malignant neoplasms of genetic origin in *Drosophila melanogaster*. *Science* 200, 1448–1459. [PubMed: 96525]
- (14). Thomas U, Phannavong B, Muller B, Garner CC, and Gundelfinger ED (1997) Functional expression of rat synapse-associated proteins SAP97 and SAP102 in *Drosophila* dlg-1 mutants: effects on tumor suppression and synaptic bouton structure. *Mech. Dev* 62, 161–174. [PubMed: 9152008]
- (15). Javier RT (2008) Cell polarity proteins: common targets for tumorigenic human viruses. *Oncogene* 27, 7031–7046. [PubMed: 19029943]
- (16). Cavatorta AL, Fumero G, Chouhy D, Aguirre R, Nocito AL, Giri AA, Banks L, and Gardiol D (2004) Differential expression of the human homologue of drosophila discs large oncosuppressor in histologic samples from human papillomavirus-associated lesions as a marker for progression to malignancy. *Int. J. Cancer* 111, 373–380. [PubMed: 15221964]
- (17). Gardiol D, Kuhne C, Glaunsinger B, Lee SS, Javier R, and Banks L (1999) Oncogenic human papillomavirus E6 proteins target the discs large tumour suppressor for proteasome-mediated degradation. *Oncogene* 18, 5487–5496. [PubMed: 10523825]
- (18). Hirata A, Higuchi M, Niinuma A, Ohashi M, Fukushi M, Oie M, Akiyama T, Tanaka Y, Gejyo F, and Fujii M (2004) PDZ domain-binding motif of human T-cell leukemia virus type 1 Tax oncoprotein augments the transforming activity in a rat fibroblast cell line. *Virology* 318, 327–336. [PubMed: 14972558]
- (19). Lee SS, Weiss RS, and Javier RT (1997) Binding of human virus oncoproteins to hDlg/SAP97, a mammalian homolog of the Drosophila discs large tumor suppressor protein. *Proc. Natl. Acad. Sci. U. S. A* 94, 6670–6675. [PubMed: 9192623]
- (20). Du M, Fan X, Hanada T, Gao H, Lutchman M, Brandsma JL, Chishti AH, and Chen JJ (2005) Association of cottontail rabbit papillomavirus E6 oncoproteins with the hDlg/SAP97 tumor suppressor. *J. Cell. Biochem* 94, 1038–1045. [PubMed: 15669058]
- (21). Rivera C, Yamben IF, Shatadal S, Waldof M, Robinson ML, and Griep AE (2009) Cell-autonomous requirements for Dlg-1 for lens epithelial cell structure and fiber cell morphogenesis. *Dev. Dyn* 238, 2292–2308. [PubMed: 19623611]

- (22). Kuhlendahl S, Spangenberg O, Konrad M, Kim E, and Garner CC (1998) Functional analysis of the guanylate kinase-like domain in the synapse-associated protein SAP97. *Eur. J. Biochem* 252, 305–313. [PubMed: 9523702]
- (23). Feng W, Long JF, Fan JS, Suetake T, and Zhang M (2004) The tetrameric L27 domain complex as an organization platform for supramolecular assemblies. *Nat. Struct. Mol. Biol* 11, 475–480. [PubMed: 15048107]
- (24). Schultz J, Milpetz F, Bork P, and Ponting CP (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl. Acad. Sci. U. S. A* 95, 5857–5864. [PubMed: 9600884]
- (25). Harris BZ, Venkatasubrahmanyam S, and Lim WA (2002) Coordinated folding and association of the LIN-2, -7 (L27) domain. An obligate heterodimerization involved in assembly of signaling and cell polarity complexes. *J. Biol. Chem* 277, 34902–34908. [PubMed: 12110687]
- (26). Feng W, Long JF, and Zhang M (2005) A unified assembly mode revealed by the structures of tetrameric L27 domain complexes formed by mLin-2/mLin-7 and Patj/Pals1 scaffold proteins. *Proc. Natl. Acad. Sci. U. S. A* 102, 6861–6866. [PubMed: 15863617]
- (27). Yang X, Xie X, Chen L, Zhou H, Wang Z, Zhao W, Tian R, Zhang R, Tian C, Long J, and Shen Y (2010) Structural basis for tandem L27 domain-mediated polymerization. *FASEB J.* 24, 4806–4815. [PubMed: 20702775]
- (28). Petrosky KY, Ou HD, Lohr F, Dotsch V, and Lim WA (2005) A general model for preferential hetero-oligomerization of LIN-2/7 domains: mechanism underlying directed assembly of supramolecular signaling complexes. *J. Biol. Chem* 280, 38528–38536. [PubMed: 16147993]
- (29). Li Y, Karnak D, Demeler B, Margolis B, and Lavie A (2004) Structural basis for L27 domain-mediated assembly of signaling and cell polarity complexes. *EMBO J.* 23, 2723–2733. [PubMed: 15241471]
- (30). Karnak D, Lee S, and Margolis B (2002) Identification of multiple binding partners for the amino-terminal domain of synapse-associated protein 97. *J. Biol. Chem* 277, 46730–46735. [PubMed: 12351654]
- (31). Nakagawa T, Futai K, Lashuel HA, Lo I, Okamoto K, Walz T, Hayashi Y, and Sheng M (2004) Quaternary structure, protein dynamics, and synaptic function of SAP97 controlled by L27 domain interactions. *Neuron* 44, 453–467. [PubMed: 15504326]
- (32). Marfatia SM, Byron O, Campbell G, Liu SC, and Chishti AH (2000) Human homologue of the *Drosophila* discs large tumor suppressor protein forms an oligomer in solution. Identification of the self-association site. *J. Biol. Chem* 275, 13759–13770. [PubMed: 10788497]
- (33). Gileadi O, Burgess-Brown NA, Colebrook SM, Berridge G, Savitsky P, Smee CE, Loppnau P, Johansson C, Salah E, and Pantic NH (2008) High throughput production of recombinant human proteins for crystallography. *Methods Mol. Biol* 426, 221–246. [PubMed: 18542867]
- (34). Philo JS (2006) Improved methods for fitting sedimentation coefficient distributions derived by time-derivative techniques. *Anal. Biochem* 354, 238–246. [PubMed: 16730633]
- (35). Stafford WF 3rd (1992) Boundary analysis in sedimentation transport experiments: a procedure for obtaining sedimentation coefficient distributions using the time derivative of the concentration profile. *Anal. Biochem* 203, 295–301. [PubMed: 1416025]
- (36). Ortega A, Amoros D, and Garcia de la Torre J (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J* 101, 892–898. [PubMed: 21843480]
- (37). Skinner JM, Cowan M, Buono R, Nolan W, Bosshard H, Robinson HH, Heroux A, Soares AS, Schneider DK, and Sweet RM (2006) Integrated software for macromolecular crystallography synchrotron beamlines II: revision, robots and a database. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 62, 1340–1347. [PubMed: 17057337]
- (38). Minor W, Cymborowski M, Otwinowski Z, and Chruszcz M (2006) HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 62, 859–866. [PubMed: 16855301]
- (39). Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR,

- Read RJ, Vagin A, and Wilson KS (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 67, 235–242. [PubMed: 21460441]
- (40). Evans PR (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 67, 282–292. [PubMed: 21460446]
- (41). Sheldrick GM (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 66, 479–485. [PubMed: 20383001]
- (42). Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 50, 760–763. [PubMed: 15299374]
- (43). Cohen SX, Ben Jelloul M, Long F, Vagin A, Knipscheer P, Lebbink J, Sixma TK, Lamzin VS, Murshudov GN, and Perrakis A (2008) ARP/wARP and molecular replacement: the next generation. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 64, 49–60. [PubMed: 18094467]
- (44). Emsley P, and Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 60, 2126–2132. [PubMed: 15572765]
- (45). Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, and Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 66, 213–221. [PubMed: 20124702]
- (46). Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, and Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 67, 355–367. [PubMed: 21460454]
- (47). Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, and Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 66, 12–21. [PubMed: 20057044]
- (48). Ramachandran GN, Ramakrishnan C, and Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J. Mol. Biol* 7, 95–99. [PubMed: 13990617]
- (49). Krissinel E, and Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol* 372, 774–797. [PubMed: 17681537]
- (50). Senkovich O, Cook WJ, Mirza S, Hollingshead SK, Protasevich II, Briles DE, and Chattopadhyay D (2007) Structure of a complex of human lactoferrin N-lobe with pneumococcal surface protein a provides insight into microbial defense mechanism. *J. Mol. Biol* 370, 701–713. [PubMed: 17543335]

Figure 1. L27 domain of Dlg1 that is a dimer in solution. The elution profile of the domain from *D. melanogaster* Dlg1 measured at 280 nm via analytical size-exclusion chromatography using a calibrated Superdex S75 column at a physiological salt concentration (Experimental Procedures) is shown as a solid black line. The dotted gray line denotes the elution profile of a 10 kDa monomeric protein standard under a similar condition.

Figure 2.

Structure of the Dlg1-L27 domain dimer. (A) Orthogonal views of the *D. melanogaster* Dlg1-L27 domain dimer with solid cylinders for α helices and arrows for β strands. The subunits (two monomers) are colored light pink and green. The helices and strands are labeled. N and C denote the amino and carboxy termini, respectively. (B) Structure-based aligned amino acid sequences of the type A L27 domain of Dlg1 from *D. melanogaster* (Dm), *Homo sapiens* (Hs), *Mus musculus* (Mm), *Danio rario* (Dr), and *Rattus rattus* (Rr). Gaps in alignment are denoted with dashes. Side chain identity/similarity denoted by shading (dark gray conserved in all; light gray conserved in most). Experimentally characterized secondary structure elements with respective PDB entries are either above (for the DmDlg1-L27 domain) or below the corresponding sequence. Amino acid residues involved in the homo- or heterodimer interface are specified by red filled circles. Amino acid residues packing in the interface between two heterodimers are marked with black filled circles. (C) Views (as in panel A) of surface representations of one of the subunits (light pink) onto which sequence conservation (from panel B) was mapped. The second subunit of the dimer is represented in semitransparent cylinders and an arrow. All structural models were generated by the PyMOL Molecular Graphics System (version 1.7.4, Schrödinger, LLC), unless stated otherwise.

Figure 3. Contacts at the dimer interface of the L27 domain of Dlg1. Ribbon trace of the DmDlg1-L27 domain dimer colored as in Figure 1A. Boxes outline three regions selected for close-up views. Black and blue boxes denote interface 1, and the red box denotes interface 2. Side chains involved at the interface are depicted as sticks. Atomic contacts are indicated by dashed lines.

Figure 4.

Elution profiles of wild-type and mutant Dlg1-L27 domains during size-exclusion chromatography (Experimental Procedures). Overlaid traces for (A) the wild type (black), L14G (blue), Y18G (green), and L14G/L15G/Y18G (red), (B) the wild type (black), Q26A (magenta), and Q26A/D27A (aquamarine), (C) the wild type (black), R39A (orange), and R39A/K42A (purple), (D) the wild type (black), I34G (purple), I34G/V37G (orange), and I34G/V37G/I38G (green), (E) the wild type (black) and F41A (green), and (F) the wild type (black), A8T (orange), and R44A (purple). The void volume for the column, determined from the elution peak of Dextran Blue, is indicated in each panel.

Figure 5.

SV-AUC analyses of wild-type and mutant Dlg1-L27 domains. (A) Sedimentation coefficient distributions, $g(s^*)$, obtained from time-derivative analysis of sedimentation velocity boundary analyses of wild-type (black) and L14G/L15G/Y18G (red), Q26A/D27A (magenta), I34G/V37G/I38G (green), and F41A (orange) substituted domains. For the sake of clarity, every fourth data point is shown and the distributions are highlighted by the solid smoothed curves. (B) Sedimentation rate, $S_{(20,w)}$, plotted as a function of the initial protein concentration (Table 1): wild type (black), I34G/V37G/I38G (green), and F41A (orange). The error bars represent the joint confidence intervals at one standard deviation.

Figure 6.

Degeneration of the electrostatic “keystone” position in Dlg1. Molecular surfaces with electrostatic potential are shown for aligned type A L27 domains: (A) *D. melanogaster* (Dm) Dlg1-L27, (B) *M. musculus* (Mm) Lin-7, (C) *H. sapiens* (Hs) PatJ, and (D) *Rattus norvegicus* (Rn) PatJ. On the surface, the position of alanine 33 (for Dm) or analogous positively charged keystone positions (for Mm, Hs, and Rn) are circled and labeled.

Figure 7. L27 domain assemblies and comparison of individual type A L27 domains. Ribbon diagram for L27 domain assemblies aligned to their type A L27 domain (left panel). An aligned individual type A L27 domain is shown in the right panel in each case. (A) Homomeric assembly of *D. melanogaster* (Dm) Dlg1-L27 (PDB entry 4RP5). Subunits are colored according to Figure 2. (B) Heteromeric assembly of *H. sapiens* (Hs) Dlg1 (yellow), MPP7 (light gray), and Mals (dark gray) (PDB entry 3LRA). (C) Heterotetrameric assembly of *R. rattus* (Rr) Dlg1 (pink) and *M. musculus* (Mm) Lin-2 (light gray) (PDB entry 1RSO). (D) Assembly of *H. sapiens* (Hs) PatJ (cyan) and *M. musculus* (Mm) PalS1 (light gray) (PDB entry 1VF6). (E) L27 assembly of *M. musculus* (Mm) Lin-2 (light gray) and Lin-7 (magenta) (PDB entry 1Y74).

Table 1.

Summary of Sedimentation Velocity Data

DmDlg1-L27	<i>S</i>₂₀	<i>D</i>_{20,w}	<i>S/D</i> (kDa)
wild type			
13 μ M	2.149 (2.137, 2.162)	10.10 (9.71, 10.48)	20.50 (19.63, 21.36)
40 μ M	2.099 (2.096, 2.103)	8.81 (8.71, 8.921)	22.94 (22.64, 23.21)
70 μ M	2.111 (2.109, 2.113)	8.76 (8.70, 8.83)	23.18 (22.99, 23.37)
HYDROPRO ^a	[1.988]	[8.43]	[22.80]
based on amino acid sequence			
L14G/L15G/Y18G			[22.24]
94 μ M (85%)	6.483 (6.459, 6.510)	7.57 (7.07, 7.94)	82.92 (78.57, 87.67)
94 μ M (15%)	3.969 (3.618, 4.351)	15.99 (11.23, 22.37)	23.30 (17.70, 28.45)
Q26A/D27A			
76 μ M (14%)	5.978 (5.893, 6.067)	17.52 (12.30, 24.28)	32.97 (23.21, 42.23)
76 μ M (86%)	2.160 (2.144, 2.174)	11.60 (11.10, 12.01)	17.91 (17.29, 18.53)
I34G/V37G/I38G			
30 μ M	2.000 (1.987, 2.011)	12.47 (12.07, 12.85)	15.47 (14.88, 16.00)
50 μ M	1.980 (1.963, 1.997)	10.92 (10.39, 11.45)	17.47 (16.49, 18.42)
80 μ M	1.906 (1.889, 1.920)	11.59 (11.08, 12.12)	15.80 (14.98, 16.61)
F41A			
7 μ M	1.823 (1.782, 1.862)	10.33 (9.05, 11.76)	17.06 (14.54, 19.47)
16 μ M	1.970 (1.953, 1.985)	13.01 (12.49, 13.53)	13.54 (12.89, 14.10)
35 μ M	1.987 (1.973, 2.000)	12.79 (12.29, 13.29)	15.02 (14.32, 15.63)

^aSedimentation parameters were calculated from the refined structure of the DmDlg1-L27 domain dimer (PDB entry 4RP5) using HYDROPRO.³⁶

Table 2.

Crystallographic Data Statistics

K ⁺ -soaked		native (orthorhombic)		native (monoclinic)
Data Collection ^a				
source	NSLS X29A	NSLS X29A	NSLS X29A	NSLS X29A
wavelength (Å)	0.9798	1.075	1.075	1.075
no. of crystals	1	1	1	1
space group cell dimensions	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁	<i>P</i> 2 ₁
<i>a</i> (Å)	48	48	28.5	28.5
<i>b</i> (Å)	57.6	57.7	47.8	47.8
<i>c</i> (Å)	64.5	64.6	64.4	64.4
β (deg)			95.5	95.5
resolution (Å)	30–1.36 (1.38–1.36)	30–1.42 (1.44–1.42)	30–1.65 (1.68–1.65)	30–1.65 (1.68–1.65)
completeness (%)	99.9 (99.9)	95.7 (93.9)	99.7 (99.8)	99.7 (99.8)
total no. of reflections	280814 (27623)	260817 (25632)	85203 (8371)	85203 (8371)
no. of unique reflections	39086 (1918)	33288 (1623)	20902 (3027)	20902 (3027)
Wilson <i>B</i> factor	15.7	17.1	21.5	21.5
redundancy	7.2 (5.1)	7.8 (8.1)	4.1 (3.9)	4.1 (3.9)
<i>R</i> _{merge} (%)	5.9 (91.9)	9.5 (93.4)	4.7 (54.4)	4.7 (54.4)
CC _{1/2} (%)	99.9 (84.6)	99.9 (84.1)	99.9 (84.1)	99.9 (84.1)
CC* (%)	100.0 (95.7)	100.0 (95.6)	100.0 (95.6)	100.0 (95.6)
$\langle I \rangle / \sigma(I)$	14.6 (2.3)	15.9 (3.1)	14.2 (2.9)	14.2 (2.9)
Phasing				
FOM (SHELX)	0.46			
FOM (DM)	0.81			
no. of K ⁺ per asymmetric unit	1			
Refinement ^a				
resolution (Å)	30–1.36 (1.39–1.35)	30–1.42 (1.45–1.42)	30–1.65 (1.72–1.65)	30–1.65 (1.72–1.65)
no. of reflections (work/free)	37070/1959	31575/1671	20868/1071	20868/1071
<i>R</i> _{work} , <i>R</i> _{free} (%)	15.0 (24.2), 20.6 (29.6)	15.0 (19.1), 19.8 (27.4)	18.4 (24.1), 21.3 (27.0)	18.4 (24.1), 21.3 (27.0)

	K ⁺ -soaked	native (orthorhombic)	native (monoclinic)
no. of atoms	1715	1720	1691
protein	1539	1551	1561
ligand	8	0	0
water	168	169	130
average <i>B</i> factor (Å ²)	28	32.1	39.2
protein	27.2	27.0	37.6
ligand	44.0	0	0
water	37.6	37.2	40.7
root-mean-square deviation			
bond lengths (Å)	0.02	0.022	0.010
bond angles (deg)	2.01	2.058	1.00
MolProbity ^b			
favored	98.8% (166 amino acids)	98.8% (163 amino acids)	98.3% (168 amino acids)
allowed	100% (168 amino acids)	100% (165 amino acids)	100% (171 amino acids)
outliers	none	none	none
Clash score	100th percentile	100th percentile	100th percentile
Molprobability score	100th percentile	100th percentile	100th percentile
PDB entry	4RP3	4RP4	4RP5

^a Statistics calculated using PHENIX;⁴⁵ data for highest-resolution shells are in parentheses.

^b Calculated with the program MolProbity.⁴⁷