The molecular architecture of the desmosomal outer dense plaque by integrative structural modeling

Satwik Pasani[1], Kavya S Menon[1], Shruthi Viswanath[1,*]

1: National Center for Biological Sciences, Tata Institute of Fundamental Research, Bengaluru 560065, India
*Corresponding Author: shruthiv@ncbs.res.in (SV)

Abstract

Desmosomes are protein assemblies that mediate cell-cell adhesion and are prevalent in tissues under mechanical stress, such as heart and epithelial tissues. However, their detailed structural characterization is not yet available. Here, we characterized the molecular architecture of the desmosomal outer dense plaque (ODP) using Bayesian integrative structural modeling *via* IMP (Integrative Modeling Platform; https://integrativemodeling.org). We integrated information from X-ray crystallography, electron cryo-tomography, immuno-electron microscopy, yeast two-hybrid experiments, co-immunoprecipitation, *in vitro* overlay, *in vivo* co-localization assays, *in-silico* sequence-based predictions of transmembrane and disordered regions, homology modeling, and stereochemistry information to generate an integrative structure of the ODP. The structure was validated by additional information from biochemical assays that was not used in modeling. The ODP resembles a densely packed cylinder with two layers: a PKP layer and a PG layer; the desmosomal cadherins and PKP span the two layers. We identified previously unknown protein-protein interfaces between DP and Dsc, DP and PG, and PKP and the desmosomal cadherins. The integrative structure sheds light on the function of disordered regions, such as the N-terminus of PKP (N-PKP) and C-terminus of PG in desmosome assembly. In our structure, N-PKP interacts with several proteins in the PG layer, alluding to its importance in desmosome assembly, and implying that it is not merely a structural filler as previously posited. Further, we identified the structural basis for defective cell-cell adhesion in Naxos disease, Carvajal Syndrome, Skin Fragility/Woolly Hair Syndrome, and cancers *via* mapping of disease-related mutations on the structure. Finally, we point to features of the structure that could confer resilience to mechanical stress, such as the PG-DP interaction and the embedding of cadherins amidst the other proteins. Taken together, we contribute the most complete and robustly validated model of the desmosomal ODP so far, providing mechanistic insight into the function and assembly of desmosomes in normal and disease states.

Introduction

Desmosomes are large, 300nm-long protein assemblies that connect the keratin intermediate filaments of adjacent cells. They mediate cell-cell adhesion and play a crucial role in maintaining tissue integrity for tissues under mechanical stress, such as heart and epithelial tissues. They also play critical roles in cell signalling and tissue differentiation. Dysfunction of desmosomes has been implicated in skin and heart diseases, auto-immune diseases, and cancers (Garrod and Chidgey, 2008; Green and Simpson, 2007; Kowalczyk and Green, 2013).

The ultra-structure of desmosomes shows its organization in three areas: the extracellular core region (EC), the outer dense plaque (ODP), and the inner dense plaque (IDP)(Delva et al., 2009). The EC is made up of the desmosomal cadherins (DCs), desmoglein (Dsg) and desmocollin (Dsc), which interact with similar molecules in adjacent cells to achieve cell-cell adhesion. The ODP, which spans 15-20nm, is a protein dense region between the EC and IDP. Here, members of the armadillo family - plakoglobin (PG) and plakophilin (PKP), members of the plakin family - desmoplakin (DP), and the cytoplasmic tails of the desmosomal cadherins interact. The ODP functions to regulate cadherins, since it contains several phosphorylation sites and binding sites for regulatory proteins. Desmoplakin links to the keratin intermediate filaments in the IDP at the cytoplasmic end of the desmosome (Garrod and Chidgey, 2008; Kowalczyk and Green, 2013).

A detailed structural characterization of the ODP is not yet available. A molecular map based on immuno-electron microscopy is known (North et al., 1999). However, this map provides the distances of plaque protein termini from the plasma membrane; it does not provide information on the three-dimensional arrangement of the proteins. A 32Å cryo-electron tomogram of the ODP, which shows its organization in two layers, has been determined by (Al-Amoudi et al., 2011). This is also the most comprehensive structural study on the ODP so far. However, the resolution of the tomogram is too low to unambiguously fit the known structures of plaque proteins and protein complexes. Moreover, domains of unknown structure, comprising a significant portion of the ODP, were not modeled. These domains make up about 40% of the stratified epithelial desmosomal ODP (Fig. 1A, Table S1). In this study, we built a more complete model of the ODP, including domains of unknown structure, by combining the data from electron cryo-tomography and immuno-EM experiments with an array of known biophysical, biochemical, and cell biological experimental data, bioinformatics predictions, and physical principles (Fig. 1, Table S2-S3) (Bonné et al., 2003; Bornslaeger et al., 2001; Hatzfeld et al., 2000; Kowalczyk et al., 1999; Smith and Fuchs, 1998).

Structures of large protein assemblies such as desmosomes are challenging to characterize using a single experimental method such as X-ray crystallography or cryo-electron microscopy. Purifying the component proteins is difficult since several of these are membrane proteins. Here we applied integrative structural modeling *via* IMP (Integrative Modeling Platform; https://integrativemodeling.org) to characterize the molecular architecture of the ODP (Alber et al., 2007; Rout and Sali, 2019; Russel et al., 2012). In this approach, we combined information from experiments along with physical principles, statistical inference, and prior models for structure determination. Several assemblies have been determined using this approach, including the yeast nuclear pore complex (Alber et al., 2007; Kim et al., 2018), 26S proteasome (Lasker et al., 2012), yeast centrosome (Viswanath et al., 2017a), and chromatin-modifying assemblies (Arvindekar et al., 2022; Robinson et al., 2015). Importantly, the Bayesian inference framework allowed us to rigorously and objectively combine multiple sources of experimental data at different spatial resolutions by accounting for the data uncertainty. It also facilitated the modeling of full-length proteins, including regions of unknown structure and/or disorder along with regions of known and/or readily modeled atomic structure.

We integrated information from X-ray crystallography, electron cryo-tomography, immuno-electron microscopy, yeast two-hybrid experiments, co-immunoprecipitation, *in vitro* overlay, *in vivo* co-localization assays, *in-silico* sequence-based predictions of transmembrane and disordered regions, homology modeling, and stereochemistry information to obtain the integrative structure of the desmosomal ODP (Fig. 1-2, Table S1-S3). Our structure was validated by additional information from biochemical assays not used in modeling (Choi et al., 2009; Kowalczyk et al., 1997; Smith and Fuchs, 1998; Troyanovsky et al., 1994b, 1994a; Wahl et al., 1996)(Table S3). The structure allowed us to

identify novel protein-protein interfaces in the ODP and reveal the structural basis for defective cell-cell adhesion associated with pathogenic mutations seen in skin diseases and cancers. Our analysis highlights the role of disordered domains such as the N-terminus of plakophilin and the C-terminus of plakoglobin in desmosome assembly. We point to aspects of the desmosome structure that could confer robustness to mechanical stress, such as the PG-DP interaction, and the embedding of cadherins amidst other proteins. Our integrative structure not only provides insights into desmosome function and assembly but also provides concrete hypotheses for future experiments.

Results and Discussion

Summary of the integrative modeling workflow

The expression of isoforms of the ODP subunits is tissue-dependent (Delva et al., 2009; Green and Simpson, 2007). Below, we detail the integrative structure of the desmosomal ODP corresponding to the upper epidermis, comprising of plakoglobin (PG), desmoplakin (DP), plakophilin (PKP1), desmocollin (Dsc1), and desmoglein (Dsg1) (Fig. 1A, Table S1).

The protein domains constituting the desmosomal ODP and the corresponding terminology used henceforth are shown (Fig. 1A, Table S1). The stoichiometry of these proteins was determined using a previously published cryo-electron tomogram (Methods) (Al-Amoudi et al., 2011). Integrative modeling proceeded in four stages (Fig. 2, Methods). Data from X-ray crystallography, electron cryo-tomography, immuno-electron microscopy, and biochemical assays was integrated with *in-silico* sequence-based predictions of transmembrane and disordered regions, homology modeling, and stereochemistry information (Fig. 1B-1C, Table S2, Table S3).

Each protein was represented by a series of spherical beads along the backbone, each bead denoting a fixed number of residues. Protein domains with X-ray structures or homology models (such as the PKP1 armadillo repeat domain) were represented at 30 residues per bead and modeled as rigid bodies, whereas domains without known atomic structure (such as the PKP1-N) were coarse-grained at 20 residues per bead and modeled as flexible strings of beads (Fig. 1A, 1C, Table S1). Data from immuno-EM was used to restrain the distance of protein termini from the plasma membrane, cryo-electron tomograms were used to restrain the localization of ODP proteins, and the data from biochemical assays restrained the distance between interacting protein domains (Fig 1B, Methods). Starting with random initial configurations for the rigid bodies and flexible beads, 180 million models were sampled using Markov Chain Monte Carlo (Replica Exchange Gibbs Sampling MCMC). At each step, models were scored based on agreement with the immuno-EM, tomography, and biochemical data, together with additional stereochemical restraints such as cylinder restraints, connectivity, and excluded volume (see Methods).

About 24866 good-scoring models were selected for further analysis. These models were clustered based on structural similarity and the precision of the clusters was estimated (Arvindekar et al., 2022; Saltzberg et al., 2021; Viswanath et al., 2017b) (Fig. S2). The quality of the models was assessed by the fit to input data, as well as to data not used in modeling (Fig. S3-S4, Table S2-S3, Methods). Further analysis included identification of previously undescribed protein-protein interfaces *via* contact maps and rationalizing skin and cancer-related diseases involving ODP proteins *via* mapping of known missense, pathogenic mutations on the integrative structure (Fig. 4-5, Fig. S5, Table S4-S5, Methods).

Additionally, we determined two other integrative structures of the desmosome ODP, based on two other sets of ODP isoforms, a second one corresponding to the upper layers of the epidermis, and one corresponding to the basal layer of the epidermis (Table S1). Given the

resolution of the input information, no significant differences were found among the three modeled structures (Fig. 3, Fig. S6).

Integrative structure of the desmosomal ODP in the upper epidermis

Integrative modeling of the desmosomal ODP in the upper epidermis resulted in a single cluster of 24016 models (97% of 24866 models), with a model precision of 67Å. Model precision is the variability of models in this cluster and is computed as the average RMSD of the cluster models to the cluster centroid (Fig 3, Fig. S2, Methods). These models fit well the input information used in modeling (Fig. S3, Table S2, Methods). They were further corroborated by their excellent agreement with information not used for modeling (Fig. S4, Table S3). The resulting integrative structures were visualized in two ways: a bead model representing the centroid of the major cluster (Fig. 3A), and a localization probability density map, representing the localization of protein domains by specifying the probability of a voxel (3D volume) being occupied by a domain in the set of structurally superposed cluster models (Fig. 3B-3G).

Overall, the desmosomal ODP resembles a densely packed cylinder with two layers, the PG layer on top of the PKP layer (Fig. 3A-3B). A striking feature of the ODP model is that the two layers are not distinct and well-separated. Rather, the desmosomal cadherins and PKP1 span both the layers. The N-terminus of PKP1 penetrates into the PG layer while the rest of it is in the PKP layer (Fig. 3B).

PKP layer

PKP1-C is the region of the ODP closest to the plasma membrane. This region has low precision in the integrative model as shown by the spread of the localization densities (Fig. 3B-3C). PKP1-S, the armadillo repeat domain of PKP1, is juxtaposed between PKP1-C and PKP1-N, at high precision (Fig. 3B-3C). This is consistent with PKP1-S localization in tomograms (Al-Amoudi et al., 2011). PKP1-N extends from PKP1-S in the PKP layer to the middle of the PG layer, forming interfaces with several proteins in the PG layer (see also Protein-protein interfaces) (Fig. 3B-3C). Its density is spread out, *i.e.*, it has a low precision, consistent with the idea that it is a disordered domain (Fig. 3B-3C) (Al-Amoudi et al., 2011).

PG layer

PKP1-N, DP-N, and PG-C form the approximate boundary between the PKP and PG layers (Fig. 3B-3D). The last two are approximately equidistant from the plasma membrane, consistent with previous immuno-EM studies (North et al., 1999)(Table S2B).

PG-S and DP-S, the armadillo repeat and plakin domains of PG and DP respectively, seem to localize in approximately the same region and physically interact (see also Protein-protein interfaces) (Fig. 3B-3E). Previously, PG-S and DP-S were hypothesized to form a regular zigzag arrangement, with both domains approximately equi-distant to the plasma membrane (Al-Amoudi et al., 2011). In contrast, in our integrative structure, the centers of PG-S and DP-S are at slightly different distances from the membrane (Fig. 3D-3E). On average, PG-S is slightly closer to the plasma membrane and DP-S is slightly closer to the cytoplasmic end. Also, there is no regular orientation to either PG-S or DP-S, although based on the localization densities, these domains appear to prefer an orientation where their long axis is approximately perpendicular to the membrane (Fig. 3D-3E).

The cytoplasmic end of the desmosomal ODP is occupied by PG-N. The PG layer protein termini having unstructured domains, PG-N, DP-N, and PG-C, are localized at low precision (Fig. 3D).

Desmosomal cadherins

The desmosomal cadherins extend from the membrane end of the ODP, through the space in the PKP layer, towards the PG layer, interacting with PG, DP, as well as PKP1 (see also Protein-protein interfaces) (Fig. 3B, 3F-3G). Dsg1 being longer, extends towards the cytoplasmic end of the PG layer, close to PG-N, where it is localized at low precision. Whereas, Dsc1 extends until PG-S in the middle of the PG layer (Fig. 3B, 3F-3G).

Protein-protein interfaces

To enable the discovery of previously unknown protein-protein interfaces in the desmosomal ODP, we computed contact maps and predicted interfaces between protein pairs (Fig. 4, Fig. S5, Table S4, Methods). Our contact maps denote the percentage of models in the cluster in which the corresponding bead surfaces are within contact distance (10Å). The contact maps are consistent with the localization of PG and PKP1 in separate layers and with the structures of known ODP sub-complexes, *e.g.*, the PG-desmosomal cadherin complexes (Fig. S5, Table S4). Analysis of the set of top 2-5% contacts, which likely excludes contacts made randomly, enabled us to identify previously unknown interfaces (Fig. 4, Fig. S5, Table S4, Methods). In general, the newly predicted interfaces are consistent with input biochemical binding information and refine the latter, providing higher-resolution information due to the integration of additional sources of information in the modeling. They form an extensive set of concrete hypotheses for future experiments (Fig. 4, Fig. S5, Table S4). Below, we discuss some of these novel interfaces in light of the role of desmosomal subunits in maintaining robust cell-cell adhesion, assembly of desmosomes, and desmosome-related diseases.

Insights into the molecular basis of desmosome-related diseases

Next, we determined the structural basis of desmosomal defects in skin diseases and cancer by mapping disease-associated mutations on our integrative structure. Specifically, we mapped known pathogenic missense mutations on desmosomal subunits that are associated with Naxos disease, Carvajal syndrome, or cancers (Fig. 5, Table S5, Methods). Both Naxos disease and Carvajal syndrome are characterized by abnormalities in epithelial tissue including palmoplantar keratoderma (thickened skin) and woolly hair (Boulé et al., 2012; Den Haan et al., 2009; Erken et al., 2011; Keller et al., 2012; Marino et al., 2017; McKoy et al., 2000; Pigors et al., 2015; Whittock et al., 2002).

PG mutations in Naxos Disease

The missense mutations PG R265H and PG E301G seen in Naxos disease are in the armadillo repeat domain of PG (Fig. 5A-5B, Table S5). These mutations are in the newly predicted PG-DP interface and known PG-Dsg1 interface, and may result in disruption of these interfaces (Fig. 4, Fig. 5A-5B). Additionally, since they are in the armadillo domain, these mutations may also affect the folding and stability of this domain, and therefore desmosome assembly.

On the other hand, the truncation mutation PG Δ690-745 is in the disordered PG C-terminus (Fig. 5A). The latter is known to regulate the size of the desmosome; deletion of PG-C results in desmosomes that are larger than usual (Palka and Green, 1997). This truncation mutation may therefore affect desmosome assembly by altering the mechanism by which PG-C regulates desmosome size, *e.g.*, by modifying interactions with regulatory proteins.

DP mutations in Carvajal Syndrome and Skin Fragility/Woolly Hair (SF/WH) Syndrome

The DP missense mutations N287K (SF/WH syndrome) and T356K, T564I, and L583P (Carvajal Syndrome) are in the spectrin homology domain of DP (Fig. 5A-5B, Table S5) as well as the newly predicted PG-DP interface (Fig, 4, Fig. 5A, Table S4). These mutations

may alter the integrity of the DP-PG interface as well as the folding and stability of the spectrin domain.

Cancers

The PKP1 R502H missense mutation is in the armadillo repeat domain of PKP1 and might affect the folding and stability of PKP1 in the ODP (Fig. 5C, Table S5). It is noteworthy that this residue is missing in the PDB structure of PKP1, PDB: 1XM9 (Choi and Weis, 2005)), indicating that it could be heterogeneous.

The other mutations associated with these diseases could not be readily rationalized by our structure (Table S5). In summary, three reasons can be identified for the pathogenicity of these mutations. They alter the folding and/or stability of ODP proteins, they disrupt protein-protein interfaces in the ODP, or they modify the binding properties of functionally important disordered protein domains in the ODP. All three types of mutations may disrupt the assembly and stability of the ODP, thereby affecting cell-cell adhesion. However, these mutations could also be pathogenic due to their effects on other functions such as cell signalling (Garrod and Chidgey, 2008).

Additionally, the structure of cardiac desmosomes is likely similar to that of the modeled epithelial desmosome. Therefore, our model could also be used to determine the structural basis of the numerous mutations related to cardiac diseases (*e.g.*, ARVC). However, we restricted the mutation analysis to epithelial diseases since our integrative structure is based on epithelial tissue isoforms (cardiac tissues consist of a slightly different set of isoforms (Delva et al., 2009; Green and Simpson, 2007)) (Fig. 1, Table S1).

Penetration of PKP-N to the PG layer and the role of PKP-N

Our models indicate that PKP1-N penetrates to the PG layer and a conserved forty-residue segment in PKP1-N interacts with several ODP proteins. In our integrative structure, the N-terminus of PKP1 (PKP1-N) penetrates from the PKP layer to the PG layer and the two layers are not well-separated (Fig. 3A-3C, Fig. 6). In contrast, PG and PKP were seen in two distinct layers in cryo-electron tomograms (Al-Amoudi et al., 2011). The densities in these tomograms were likely contributed by regions of known structure (*e.g.*, PG-S and PKP1-S). PKP1-N, being disordered, is possibly flexible and heterogenous, leading to smoothing out of its densities upon averaging (Fig. S7). In integrative modeling *via* IMP, regions of unknown structure can be modeled alongside regions of known structure. By combining biochemical binding data along with structural (electron cryo-tomography) data, our approach allowed us to localize disordered domains like PKP1-N.

In particular, PKP is posited as a structural filler in the ODP (Al-Amoudi et al., 2011; Kowalczyk et al., 1999; Smith and Fuchs, 1998). But, in our structure, N-PKP mediates interactions with several ODP proteins (Fig. 4, Fig. 6), implying that PKP plays a more integral role in desmosome function and assembly. Specifically, PKP1$^{181-220}$ interacts with Dsc1$^{775-814}$, and PKP1$^{161-220}$ with Dsg1$^{650-689}$; notably, both desmosomal cadherins share binding sites on PKP1 (Fig. 4). Also, PKP1$^{141-180}$ interacts with DP$^{1-60}$ at a slightly lower, but still stringent, contact map cutoff (top 5% of all contacts, Fig. S5, Table S3). Interestingly, this forty residue stretch in PKP1-N, PKP1$^{161-200}$, interacts with DP as well as the cadherins, and sequence analysis suggests that this sequence in PKP1-N is conserved (Fig. S7).

This is consistent with studies that show that PKP1 enhances recruitment of other desmosomal proteins, increasing desmosome size, and promoting desmosome assembly. For example, (Bornslaeger et al., 2001; Kowalczyk et al., 1999; Sobolik-Delmaire et al., 2006) showed that PKP1 clusters DP, (Hatzfeld et al., 2000) showed that PKP1 interacts with DP as well as desmosomal cadherins, and keratins, and (Tucker et al., 2014) showed

that PKP1 interacts with DP and Dsg3. Two of the above studies mention that the amino tip of PKP1, $PKP1^{1-70}$ (Hatzfeld et al., 2000) and $PKP1^{1-34}$ (Sobolik-Delmaire et al., 2006), recruits DP. In our models, although $PKP1^{141-180}$ (middle of PKP1-N), is the most probable PKP1-binding region (*i.e.,* highest confidence contact) for DP, the amino-tip of PKP1 is also proximal to DP and is among the top 5% of PKP1-DP contacts (Table S4). This region is also predicted to bind to DP based on Alphafold2-Multimer (See Comparison to Alphafold2-Multimer). In summary, our models indicate that PKP1-N, specifically the conserved region $PKP1^{161-200}$, could be involved in the recruitment and/or subsequent stabilization of other ODP proteins, suggesting that PKP1 is not just a structural filler in the desmosome.

PG-C as a regulator of desmosome size

In our integrative structure, the C-terminus of PG, $PG^{674-745}$, extends outwards, suggesting that it can form lateral connections with other proteins (Fig. 3D, Fig. 6). It is known to play a role in regulating the size of the desmosome. Deletion of the PG C-terminus resulted in larger desmosomes due to lateral association (Palka and Green, 1997). Moreover, this deletion was also associated with Naxos disease and defects in tissue integrity, highlighting the importance of PG-C (McKoy et al., 2000).

The mechanism by which PG-C regulates the size of desmosomes remains to be elucidated. It is predicted to be intrinsically disordered (IDR) (Fig. S7). $PG^{683-687}$ in this region is predicted to be a MoRF (molecular recognition feature), which is a motif in a disordered protein sequence that recognizes and binds to another protein (Disfani et al., 2012). The presence of the MoRF may allow PG-C to bind to itself, *i.e.,* PG-S, or to other proteins to enable regulation of desmosome size. In particular, the former mechanism, *i.e.*, IDR tails competitively binding to domains of the same protein to inhibit their function, is well-known for several enzymes and single-stranded DNA-binding proteins (Uversky, 2013). Finally, this region also contains a phosphosite (PG S730), suggesting that phosphorylation could potentially be another mechanism by which the desmosome size is regulated (Bian et al., 2014).

PG-DP interaction strengthens the connection between the cadherins and the IF

Our integrative structure identifies an interaction between the plakin domain of DP and the armadillo repeat domain of PG, $DP^{178-267}$ and $PG^{276-335}$ (Fig. 4). DP-S appears to encapsulate PG-S in the densities (Fig. 3E, Fig. 6). This interaction could provide a robust mechanism for desmosomes to anchor intermediate filaments (IF) and withstand mechanical stress. In fact, PG-DP binding is shown to be required for effective IF anchoring in desmosomes. PG knockout cells showed defective anchoring of IF (Acehan et al., 2008). Both the DP plakin domain and the PG arm domain are conserved across vertebrates, suggesting this interaction could also be conserved (Green et al., 2020; Smith and Fuchs, 1998). Further, a mutation in this region, PG E301G, was associated with Naxos disease and defects in epithelial tissue, further alluding to the importance of this interaction (Fig. 5, Table S5).

Moreover, this interaction could be important for desmosome assembly. In transient expression experiments in COS cells, PG was shown to be required for DP recruitment to cell borders (Kowalczyk et al., 1999). In our models, the DP binding region of PG overlaps with its cadherin-binding region, consistent with the fact that these three ODP proteins cluster together in desmosome assembly (Fig. 4) (Kowalczyk et al., 1999). Given their proximity, DP could also regulate the signaling functions of PG and PG-mediated crosstalk between desmosomes and adherens junction (Garrod and Chidgey, 2008).

Localization and interactions of desmosomal cadherins

The desmosomal cadherins wind their way through the other proteins in the PG and PKP layers, making several interactions (Fig. 3B, F, G, Fig. 6). They appear to be embedded in the thick of the other proteins, instead of circumnavigating the other proteins. This embedding in the midst of other ODP proteins provides a stronger anchoring for the cadherins and their extracellular domains in the cytoplasm. In turn, this feature could buffer the desmosomes from mechanical stress.

Notably, Dsc and Dsg different in their interactions with the other proteins. Dsc1$^{795-833}$ (the Dsc1 region N-terminal to its PG binding site) interacts with DP$^{1-60}$ (Fig. 4). Whereas, an interaction with DP is not seen for Dsg1 (Fig. S5). This is consistent with the input information (Fig. 1, Table S2) (Smith 1998). It is also consistent with experiments that showed that Dsg requires PG to recruit DP, while Dsc can recruit DP independently (Kowalczyk et al., 1999).

Comparison to Alphafold2-Multimer

We also attempted to model sub-complexes of the ODP using the recent AI-based protein structure prediction method, Alphafold2-Multimer (Evans et al., 2021)(Methods Stage 4).

PG-Dsc1 and PG-Dsc1 complexes

AF2-multimer correctly reproduced the PG-Dsc1 and PG-Dsc1 complexes which were homology modeled based on PDB 3IFQ in this study (Fig. 1, Fig. 4, Fig. S8A-S8B, Fig. S9A-S9B). The template was likely part of the AF2 training set (all pre-2019 PDB structures).

PKP1-Dsc1 and PKP1-Dsc1 complexes

AF2-Multimer produced confident predictions for the PKP1-Dsg1 and PKP1-Dsc1 complexes (Fig. S8C-S8D, Fig. S9C-S9D). In these predictions, the PKP1 binding region for Dsg1 and Dsc1 is approximately similar to the PG binding region for cadherins, which is a reasonable prediction based on structural similarity, since PKP1 has an armadillo domain like PG (Fig, 4, Fig. S8C-S8D, Fig. S9C-S9D).

The Dsc1 region binding to PKP1 is distinct from that binding to PG and is consistent with the region of Dsc1 located in the PKP layer in our model. However, the Dsg1 region that binds PKP1 in the Alphafold2 structure overlaps with its PG-binding region in our model. Additionally, our contact predictions from integrative modeling identify interfaces between the disordered PKP1-N and desmosomal cadherins, which are not captured in AF2-Multimer.

PKP1-DP and PG-DP complexes

Interestingly, AF2 predicted an interface between a part of the disordered N-terminus of PKP1 (approximately PKP1$^{20-51}$) and DP (Fig. S8E, Fig. S9E), predicting a potential disordered-to-ordered transition on binding for PKP1. The predicted interface overlaps with our contact map predictions of interfaces between DP and PKP1 (Fig. 4, Fig. S5, Table S4) and is also consistent with studies that show that the tip of PKP1-N binds to DP (Hatzfeld et al., 2000; Sobolik-Delmaire et al., 2006). However, there was no predicted interface between PG and DP.

AF2-multimer presumably predicts the structure of a complex if it is similar to known complexes, or if it involves disordered regions that become ordered upon binding to a partner. However, the predicted interface information is incomplete. For example, no interface was detected for the PG-DP complex. AF2-multimer predicts a single model or a small number of candidate models, while our integrative modeling method predicts a larger probability-weighted ensemble of models consistent with input information. Furthermore, it is a deep-learning method based on general patterns in existing protein structures. It does not account for information that is specific to a system, such as the membrane topology, layered arrangement of proteins or the oligomeric states of proteins. Lastly, these are only dimeric

predictions, and the error in AF2-Multimer predictions would get amplified for larger multimeric assemblies, such as the full desmosomal ODP, leading to a potentially inaccurate prediction. In summary, tools like AF2-Multimer are not currently sufficient to model the complete desmosome at high-resolution, presumably due to low sequence similarity to existing structures and the presence of disordered regions.

Here, we obtained an integrative structure of the desmosomal ODP by combining X-ray crystal structures, cryo-electron tomograms, distances from immuno-EM data, interacting protein domains from biochemical binding assays, bioinformatics predictions of transmembrane and disordered regions, homology modeling, and stereochemistry information. Our model is the most complete and robustly validated model of the desmosomal ODP so far, providing mechanistic insight into the function and assembly of desmosomes in normal and disease states. Importantly, the model also provides concrete hypotheses for future experiments. Notably, the termini of several modeled subunits (PG, DP, PKP) have disordered regions (Fig S7), and our structures shed light on their functions. High-resolution structural data, *e.g.*, higher-resolution cryo-EM maps would improve the structural characterization of the desmosome and our knowledge of the mechanistic details of cell-cell adhesion.

Methods

Integrative structure determination of the desmosomal ODP proceeded through four stages (Alber et al., 2007; Rout and Sali, 2019)(Fig. 1-2). Our modeling procedure used the Python Modeling Interface of the Integrative Modeling Platform (IMP 2.17.0; https://integrativemodeling.org), an open-source library for modeling macromolecular complexes (Russel et al., 2012), and is primarily based on previously described protocols (Arvindekar et al., 2022; Saltzberg et al., 2021; Viswanath et al., 2017b). Python libraries scipy (Virtanen et al., 2020) and matplotlib (Hunter, 2007) were used for analysis, GNU Parallel (Tange, Ole, 2020) was used for parallelization, UCSF Chimera v1.15 (Pettersen et al., 2004) and UCSF ChimeraX v1.5 (Pettersen et al., 2021) were used for visualization. Input data, scripts, and results are publicly available at https://github.com/isblab/desmosome and ZENODO. Integrative structures will be deposited in the PDB-DEV (https://pdb-dev.wwpdb.org) .

Stage 1: Gathering data

Isoforms The ODP comprises of PG (plakoglobin), PKP (plakophilin), DP (desmoplakin), and Desmosomal Cadherins (DC of two types, Desmoglein, Dsg, and Desmocollin, Dsc). Desmosomes from different tissues vary in the isoforms of these constituent proteins (Garrod and Chidgey, 2008; Green and Simpson, 2007). Here, we modeled three desmosomal ODPs: two corresponding to the stratified epithelium, varying in the PKP isoform used (PKP1 and PKP3), and the third based on isoforms in the basal epithelium (Garrod and Chidgey, 2008; Green and Simpson, 2007)(Fig. 1A, Table S1). The modeling protocol and resulting structures are similar for all three ODPs (Fig. 3, Fig. S6). The description in the paper corresponds to the first ODP, *i.e.,* stratified epithelium with PKP1. Epithelial desmosomes were chosen for modeling as there was more information (*e.g.*, from protein-protein binding experiments) on epithelial isoforms than desmosomes in Heart tissue. The extracellular regions of the Desmosomal Cadherins were not modeled, based on sequence annotations in Uniprot (see also (Choi et al., 2009)). Further, we do not model Dsg1$^{843\text{-}1049}$ and DP$^{585\text{-}2871}$ as they are known to be outside the ODP (Al-Amoudi et al., 2011; Garrod and Chidgey, 2008; Nilles et al., 1991)(Table S1).

Stoichiometry and number of copies The stoichiometry of the desmosomal proteins was based on previous studies using modeling and density analysis on cryo-electron microscopy data (Al-Amoudi et al., 2011)(See Stage 2).

Atomic structures The plakin domain of DP and armadillo domains of PG and PKP1 were modeled by their X-ray structures (PDB: 1XM9 (PKP) (Choi and Weis, 2005), 3R6N (DP) (Choi and Weis, 2011), 3IFQ (PG-DC) (Choi et al., 2009), while the PG-Dsc and PG-Dsg complexes were obtained by homology modeling using MODELLER (Šali and Blundell, 1993) and HHPRED (Gabler et al., 2020) for sequence alignment (Fig. 1C, Table S1).

Cryo-electron tomogram We used a 32 Å cryo-electron tomogram (EMD-1703, denoised mask without symmetrization) of the ODP (Al-Amoudi et al., 2011). The map was segmented using UCSF Chimera Segger (Pintilie et al., 2010) and the densities corresponding to the PKP and PG layers were used for modeling. (Fig. 1C).

Immuno-EM The distance of the N and C termini of the desmosomal proteins from the plasma membrane was informed by immuno-electron microscopy gold-staining experiments (Fig. 1B, Table S2) (North et al., 1999). Using Clustal-Omega (Sievers et al., 2011), the alignment between Bovine/Xenopus PG and DP (used in the experiments) and the Human PG and DP (used in modeling) is almost 1-to-1, and therefore, the residue ranges for the antibody-binding regions are taken to be the same. Values for PKP1 were used for PKP3 after alignment.

Protein-protein binding assays The relative distance between ODP protein domains was informed by biochemical data from multiple biochemical studies, including yeast-2-hybrid (Bonné et al., 2003; Hatzfeld et al., 2000; Kowalczyk et al., 1999), co-immunoprecipitation (Bonné et al., 2003; Kowalczyk et al., 1999), *in-vitro* overlay assays (Smith and Fuchs, 1998), and *in-vivo* co-localisation assays (Bonné et al., 2003; Bornslaeger et al., 2001; Kowalczyk et al., 1999) (Fig.1B, Table S2-S3). Due to experimental issues, the information pertaining to Dsc3a binding is not usable from (Bonné et al., 2003) and we therefore use the corresponding information from Dsc3b binding.

Stage 2: Representing the system and translating data into spatial restraints

Stoichiometry and number of copies, PG layer and the Desmosomal Cadherins

The stoichiometry of the desmosome ODP was 1:1:1:1 for DP:PG:PKP:DC based on previous studies (Al-Amoudi et al., 2011). The number of copies of each protein was based on fitting an equal number of PG and DP molecules to the PG layer of the cryo-electron tomogram. However, the number of PG and DP proteins that correspond to the tomogram was unknown and computed to be four each by fitting different numbers of PG and DP molecules to the PG layer density in independent modeling runs (Supplementary Section 1.1). We model 4 DC molecules, two each of Dsc1 and Dsg1.

Stoichiometry and number of copies, PKP layer

The PKP layer has seven distinct densities. These correspond well (average EM cross-correlation around mean in UCSF Chimera = 0.91) to the structured ARM repeats of

seven PKP molecules (Al-Amoudi et al., 2011)(Fig. S1 inset). To keep a 1:1:1:1 stoichiometry for PG:DP:PKP:DC, we selected four of these seven PKP molecules to represent in full; the central PKP and three symmetrically surrounding PKPs (Fig. S1 inset).

We also represented the remaining three PKP molecules ("non-interacting" PKPs) by their structured ARM repeats alone. These PKPs participate only to satisfy the cryo-electron tomogram and to exclude other proteins from these locations in space. The locations and orientations of each of these PKPs were fitted based on cross-correlation to the PKP densities in the tomogram; subsequently they were fixed during sampling.

Multi-scale coarse-grained bead representation

We use a coarse-grained representation of the proteins where a set of contiguous amino acids in a protein is represented by a spherical bead (Fig. 1A, Table S1). Domains with known atomic structures were represented by 30-residue beads to maximize computational efficiency and modeled as rigid bodies where the relative configuration between the beads is fixed during sampling. In contrast, domains without known structure were coarse-grained at 20 residues per bead and modeled as flexible strings of beads which can move relative to one another.

Next, we encoded the information gathered in stage 1 into spatial restraints that constitute a scoring function which allows scoring each model in proportion to its posterior probability. This score allows sampling high-probability models that best satisfy the data.

EM restraints

A Bayesian EM restraint was used to incorporate the information from the cryo-electron tomogram (Bonomi et al., 2019). PKP-S, the structured region of PKP, was restrained by the PKP-layer density; PG and DP molecules were restrained by the PG layer density. The EM restraint was not applied to regions such as PKP-N, PKP-C, and the desmosomal cadherins as they are either disordered and/or extended and therefore considered to be averaged out or contribute negligibly to the density in the tomogram (Al-Amoudi et al., 2011). The part of DC complexed with PG was included in this restraint.

Immuno-EM restraints

The distances of ODP protein termini to the plasma membrane were restrained by a Gaussian restraint with the mean and standard deviation equal to the mean distance and standard error measured in immuno-EM gold-staining experiments (North et al., 1999). The set of restrained beads for each protein terminus corresponded to the antibody-binding region in the experiments. The restraint score was based on the bead in the terminus that was closest to the mean distance obtained from the experiment, for each protein copy. Desmosomal Cadherins were not restrained by immuno-EM since they form a complex with PG, which is restrained by immuno-EM data. The complexed region is more specific than the antibody-binding region for Dsg1. Further, immuno-EM measurements were not available for specific Dsc isoforms.

Binding restraints

The distances between interacting protein domains were restrained by a harmonic upper bound on the minimum distance among the pairs of beads representing the two interacting domains, (the score is zero for distances less than or equal to 0, and quadratically rises above zero). For two interacting proteins A and B, ambiguity, *i.e.,* multiple copies of a

protein, was factored in by adding multiple such distance restraints. For each copy of protein A, the minimum distance among all pairs of beads across all copies of B was restrained. Similarly, for each copy of protein B, the minimum distance among all pairs of beads across all copies of A was restrained. This formulation allows a protein copy to find a binding partner from any of the available copies of the other protein, potentially allowing multiple protein A copies to bind to the same protein B copy.

Different experiments provide different levels of evidence as to whether their results can be extended to *in-vivo* conditions and whether the results preclude indirect binding via an intermediary protein. Restraints were therefore weighed in the order Overlay Assays = Co-Immunoprecipitation > Yeast-2-Hybrid. However, the results we obtain are fairly robust to this weighting scheme and all the experimental data are individually satisfied in the final set of models (Fig. S3). If multiple experiments provided data on the binding of two proteins, the highest-resolution data (*i.e.,* more specific binding site) was chosen.

Cylindrical restraints

To keep the modeled proteins close to the tomogram, beads were restrained to lie within a cylinder of radius 150 Å that encloses the map. The restraint was implemented using a harmonic upper bound on the distance of each bead from the cylinder surface.

Excluded volume restraints

The excluded volume restraints were applied to each bead to represent the steric hindrance of the protein residues that disallow other residues to come in physical proximity. The bead radius was calculated assuming standard protein density (Alber et al., 2007), with beads penalized based on the extent of their overlap.

Sequence connectivity restraints

We applied sequence connectivity restraints on the distance between consecutive beads in a protein molecule. The restraint was encoded as a harmonic upper bound score that penalizes beads that are greater than a threshold distance apart. The threshold distance is different for each protein and the calculation is inspired by models from statistical physics (Teraoka, 2002)(Supplementary Section 1.2). As a summary, we predict what proportion of each protein's predicted secondary structure is disordered using PSIPRED (Buchan and Jones, 2019), and compute the threshold based on this proportion, the known radii of gyration for disordered regions, and bead radii for globular proteins estimated from their density (Alber et al., 2007). For regions with known structures, the inter-bead distances were fixed during sampling and their contribution to the restraint score was fixed across models.

Stage 3: Structural sampling to produce an ensemble of structures that satisfies the restraints

We employed Gibbs sampling Replica Exchange Monte Carlo sampling (Arvindekar et al., 2022; Saltzberg et al., 2021; Viswanath et al., 2017b). The positions of the rigid bodies and flexible beads were sampled as in previous protocols, with a few customizations.

First, we implemented an Anchoring Constraint wherein the membrane-proximal beads of the desmosomal cadherins were initialized adjacent to the membrane and were constrained to move only along the membrane plane during sampling (Fig. 1B).

Second, a custom random initialization was used for the PG layer. The PG and DP rigid bodies and beads were randomized within a bounding box that tightly enclosed the PG layer density. The orientation of the long axis of the structured region of PG and DP molecules

with respect to the membrane determines the polarity of each PG/DP molecule (N-to-C along the normal to the membrane). After the random initialization, if it was opposite of the polarity observed from immuno-EM (North et al., 1999), this polarity was corrected by flipping the structured region along a random axis in the plane of the membrane by 180 degrees; in effect, reversing the polarity along the normal to the membrane while keeping its orientation random. For example, if a PG molecule was initialized with its N-terminus closer to the membrane than its C-terminus, its orientation would be flipped. This is because, owing to the high protein density of the PG layer, molecules with the incorrect polarity might not have the freedom to flip polarity during sampling.

Finally, a custom random initialization was used for the PKP layer. Each PKP was initialized around one of the molecule-wise PKP densities with a random orientation.

The Monte Carlo moves included random translations of individual beads in the flexible segments, random translations and rotations of rigid bodies and super-rigid bodies, *i.e.*, groups of rigid bodies and beads of the same protein or complex. The size of these moves and the replica exchange temperature for the replicas was optimized using StOP (Pasani and Viswanath, 2021). A model was saved every 10 Gibbs sampling steps, each consisting of a cycle of Monte Carlo steps that proposed a move for every bead and rigid body once. We sampled a total of 180 million integrative models.

Stage 4: Analyzing and validating the ensemble of structures

The sampled models were analyzed to assess sampling exhaustiveness and estimate the precision of the structure, its consistency with input data and consistency with data not used in modeling. We based our analysis on the protocols published earlier (Arvindekar et al., 2022; Saltzberg et al., 2021; Viswanath et al., 2017b; Webb et al., 2018).

Filtering the models into a good-scoring set

To make analysis computationally tractable and to select models that have a good score, *i.e.* higher probability, we first selected the models to create a good-scoring set which involved the following steps. Models were first filtered based on score equilibration and auto-correlation decay along the MCMC runs (Supplementary Section 2.1). Filtered models were clustered based on their restraint scores using HDBSCAN (McInnes et al., 2017), resulting in a single cluster of 37145 models (Saltzberg et al., 2021). Subsequently, these models were filtered to choose models for which each restraint score as well as the total score is better than the corresponding mean plus 1.46 standard deviations, leading to a good-scoring set of 24866 models for the next stage of analysis (Arvindekar et al., 2022).

Clustering, Precision, and Localization Densities

We next assessed if the sampling was exhaustive by previously established protocols which randomly divide the models into two independent sets and assess *via* statistical tests whether the two sets had similar scores and structures (Arvindekar et al., 2022; Saltzberg et al., 2021; Viswanath et al., 2017b)(Fig S2). We perform structural clustering of the models to find the minimum clustering threshold for which the sampling is exhaustive (sampling precision) as well as the mean RMSD between a cluster model and its cluster centroid (model precision) (Fig. S2). The bead-wise RMSD calculation in the protocol was extended to consider ambiguity, *i.e.*, multiple protein copies, and the resulting code was parallelized. The RMSD between two models is the minimum RMSD among all combinations of protein copy pairings between the models. For example, two models containing four copies of PG have 4! possible bipartite pairings of PG copies among them for which the RMSD needs to be computed. The consideration of ambiguity was applied to all proteins except PKP. Each PKP copy was initialized to the same molecule-wise EM density in every simulation and usually remained close to it throughout the simulation. PKP copies could be considered as

non-interchangeable because of the presence of fixed, non-interacting PKPs in their midst. The latter also precludes the need of alignment to a common frame of reference during RMSD calculations.

The result of integrative modeling was a single major cluster corresponding to 24016 (96.6% of 24866) models. The model precision, which quantifies the variability of models in the cluster, and is defined as the average RMSD of a cluster model from the cluster centroid, was 67Å . The cluster is visualized *via* localization probability density maps, which specify the probability of a volume element being occupied by beads of a given domain in the set of superposed models from the cluster (Fig 3).

Fit to input information and fit to information not used in modeling
To calculate the fit to data from protein-protein binding assays, we calculated the minimum distance among all bipartite pairs of beads representing all copies of interacting domains for each model in the cluster and visualized the distribution (Fig. S3A, Fig. S4, Table S2A, Table S3). With the exception of two experiments, the information not used in modeling was in complete agreement with our model (Fig. S4, Table S3). The two exceptions correspond to data that mentions that the first few cytoplasmic residues of Dsc1, in the PKP layer in our model, bind to proteins in the PG layer.

To calculate the fit to immuno-EM data (North et al., 1999), for each restrained protein terminus, we calculated the difference between the model-predicted distance of the terminus to the plasma membrane and the corresponding mean distance from experiment. The model-predicted distance for a terminus was equal to the distance of the terminus bead closest to the experimental mean. The distribution of the difference for each copy of a protein for each model in the cluster was visualized (Fig S3B).

To calculate the fit to the tomogram, we computed cross-correlation between the localization probability densities for the cluster and the segmented tomogram for the PG and PKP layers separately (Supplementary Section 2.2, Fig S3C).

Contact Maps
A contact between beads is defined as a surface-to-surface distance of 10 Å or lower. For each protein pair, we obtained the proportion of models in the cluster that have at least one contact for each bead pair across all copies of the two proteins. To filter out the significant contacts from those that might occur by chance, we identified significant contacts as those present in at least 20-25% of the models. A 25% cutoff corresponds to approximately the top ≤2% of all possible contacts for each protein-pair while a 20% cutoff corresponds to ~8% of all possible contacts for PG-DP and ≤5% for the rest of the protein pairs.

Mapping disease mutations
We considered two kinds of mutations to map to the integrative structure. First, disease mutations associated with defects in epithelial tissue that could be mapped to ODP protein domains and/or residues were obtained by a literature search and using databases such as OMIM and Uniprot ("Online Mendelian Inheritance in Man," 2023; The UniProt Consortium et al., 2023). These mutations corresponded to those seen in Naxos disease (ARVC with palmoplantar keratoderma and woolly hair) and Carvajal syndrome (Left ventricular cardiomyopathy with palmoplantar keratoderma and woolly hair) (Boulé et al., 2012; Den Haan et al., 2009; Erken et al., 2011; Keller et al., 2012; Marino et al., 2017; McKoy et al., 2000; Pigors et al., 2015; Whittock et al., 2002). Second, cancer-associated somatic, missense, confirmed pathogenic mutations on ODP proteins that occurred in five or more samples were extracted from the COSMIC database (Tate et al., 2019).

We did not consider mutations involved in cardiac disease as we model an epithelial ODP. In general, we refrained from mapping mutations across isoforms. We also did not consider

mutations that could not be mapped to the protein domains, although a large number of these are known, for example, pathological differential expression of proteins.

Comparison to Alphafold Multimer

We ran Alphafold2-Multimer (Evans et al., 2021) for pairs of proteins: PG-DP, PKP-DC, PG-DC and PKP-DP. For each pair, we chose the best ranked prediction, based on the PTM + IPTM score, to discover confidently predicted interfaces between the proteins. Confidently predicted interfaces were identified as residue pairs, with one residue from each protein, in which each residue was confidently predicted (pLDDT > 70), the residue pair had an accurate relative prediction (PAE < 5), and the pair was at an interface (Cα-Cα distance < 10Å).

Data availability

Files containing input data, scripts, and results are at https://github.com/isblab/desmosome. Integrative structures will be deposited in the PDB-Dev (https://pdb-dev.wwpdb.org).

Acknowledgements

We thank lab members Shreyas Arvindekar, Kartik Majila, and Muskaan Jindal for helpful comments on the manuscript. We thank Aditi Pathak for help with Alphafold-multimer. We also thank Karan Gandhi, Sarika Tilwani, and Sorab Dalal of ACTREC, India and Swadhin Jana of NCBS for their help. Molecular graphics images were produced using the UCSF Chimera and UCSF ChimeraX packages from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR001081, NIH R01-GM129325, and National Institute of Allergy and Infectious Diseases).

Funding

This work has been supported by the Department of Atomic Energy (DAE) TIFR grant RTI 4006 and Department of Science and Technology (DST) SERB grant SPG/2020/000475 from the Government of India to SV.

Figures

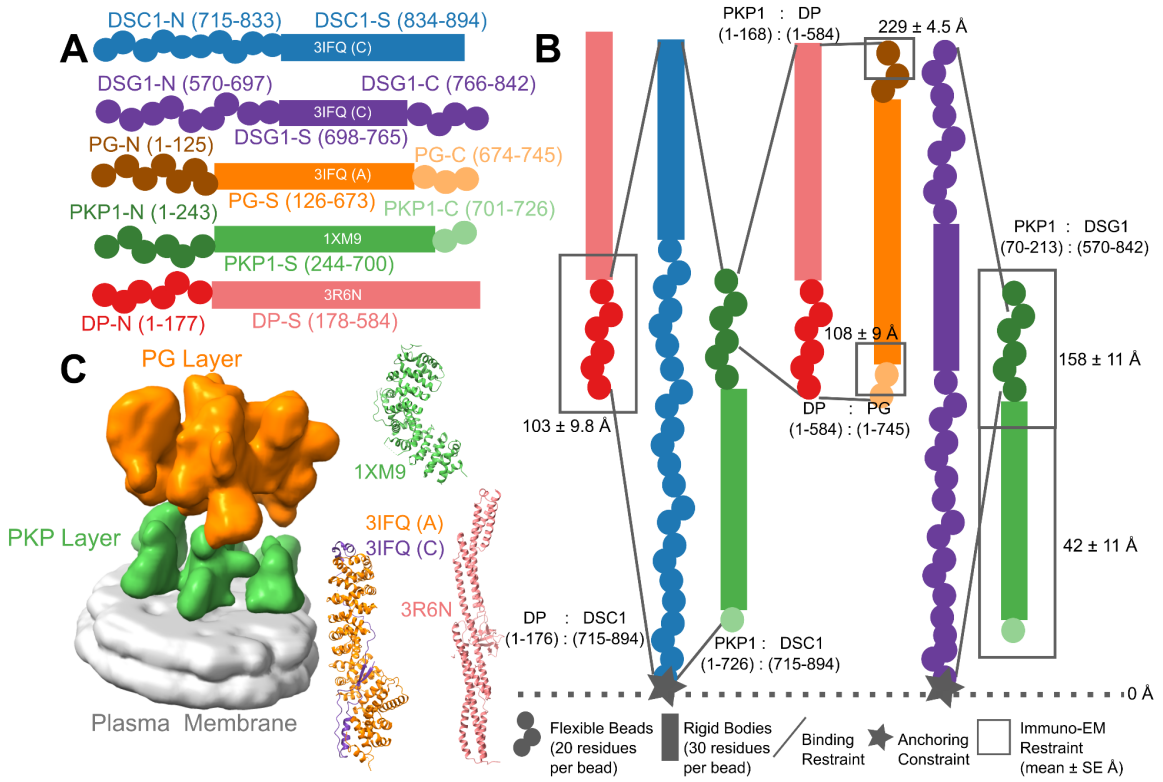

Figure 1 Representation and restraints used for integrative modeling of the desmosomal ODP A) The isoforms used in modeling the desmosomal ODP of stratified epithelia and the representation of the different protein domains as rigid bodies with known structure (rectangles with PDB ID and chain name) or flexible beads (circles). The domains with known structure are usually denoted by a suffix -S after the protein (*e.g.,* DP-S), while the termini are denoted by -N or -C suffixes after the protein (*e.g.,* DP-N). **B)** Three types of restraints are shown. **1.** Binding restraints between interacting protein domains depicted by a pair of lines connecting the boundaries of each interacting domain pair. **2.** Immuno-EM restraint for localizing protein termini depicted by rectangles around the restrained protein terminus, and **3.** Anchoring constraint for localizing the transmembrane region of the cadherins depicted by star. The color scheme follows that in Panel A. **C) (Left)** The cryo-electron tomogram (EMD-1703) used for modeling with the PKP and the PG layers segmented. The density corresponding to the plasma membrane was not used for modeling. **(Right)** The PDB structures used, colored according to panel A. See also Methods, Fig. S1, Tables S1-S2.

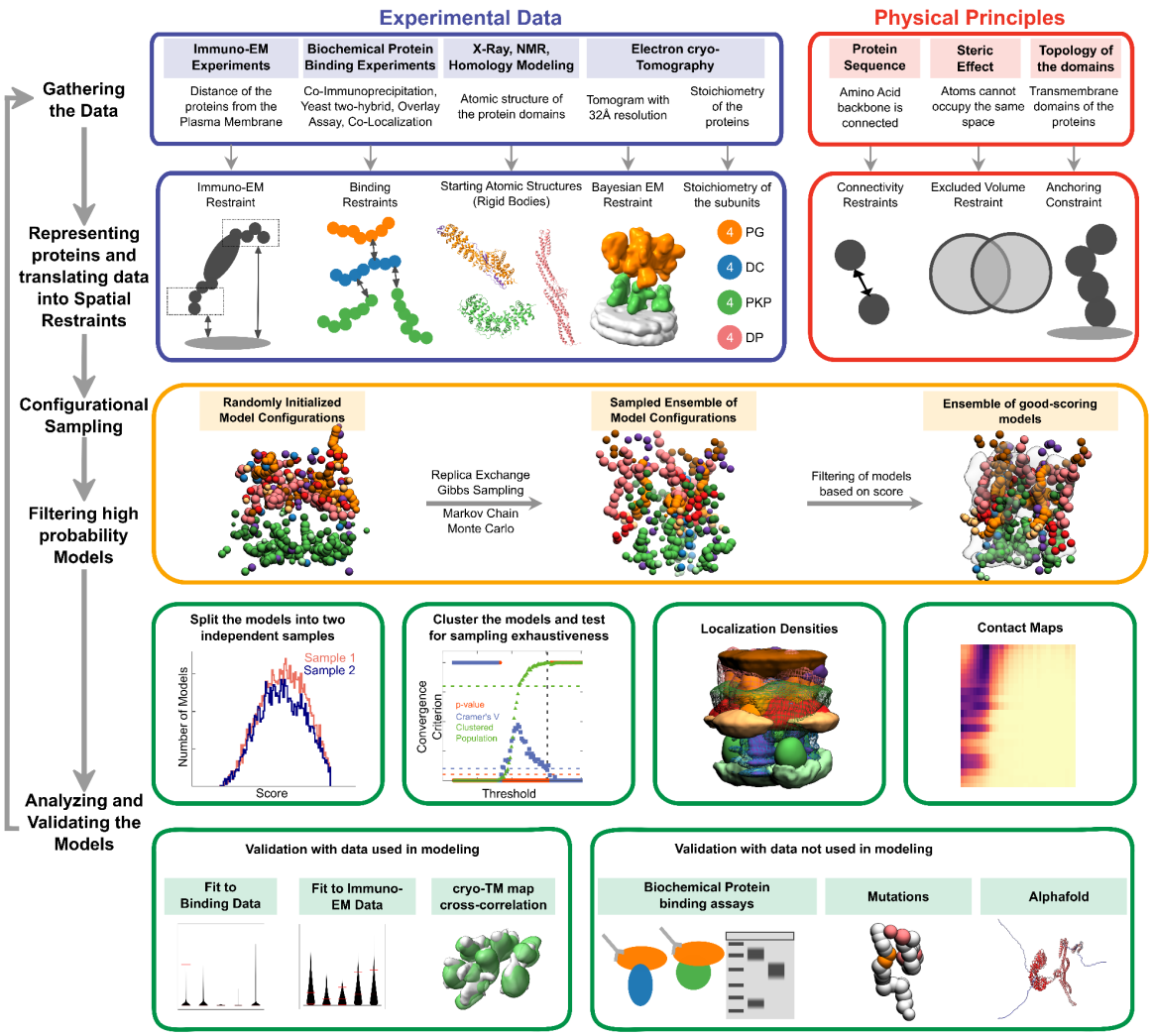

Figure 2 Integrative modeling of the desmosomal ODP. From top to bottom, the rows describe the input information (first), how the input information is encoded into spatial restraints (second), the sampling procedure (third), analysis (fourth), and validation of the results (fifth).

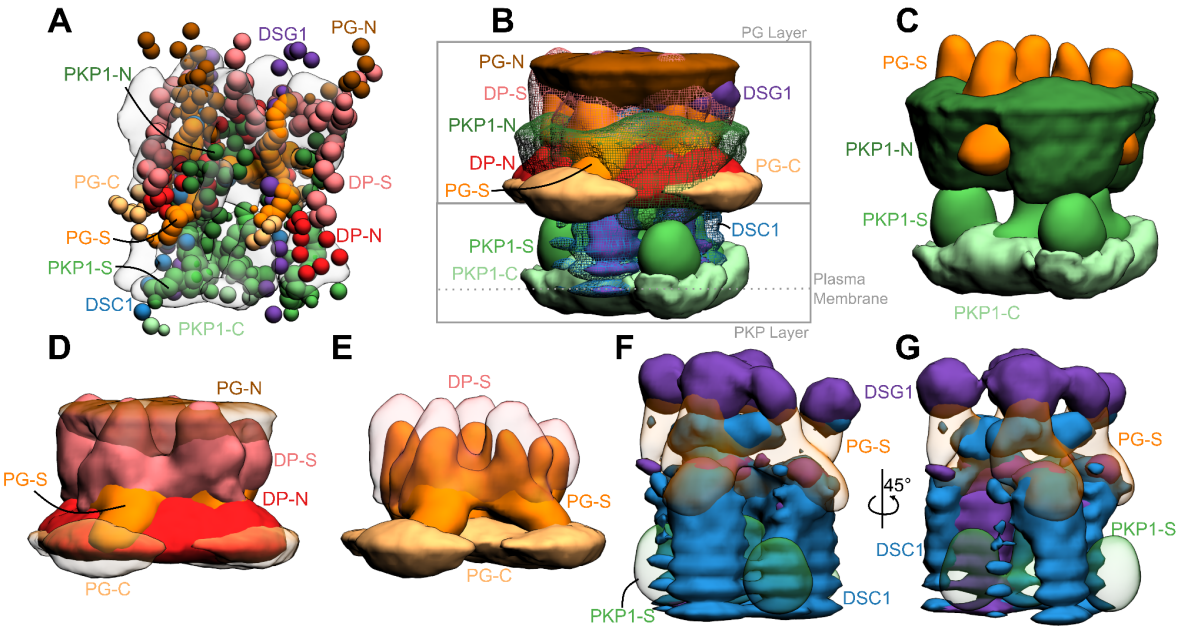

Figure 3 Integrative structure of the desmosomal ODP A) The cluster center bead model for the major structural cluster with the cryo-tomogram (EMD-1703) superimposed in translucent gray. **B)** Localization densities of the major cluster. The densities are at a cutoff of approximately 15% for PKP1-C, PKP1-S, PG-S, DP-S, Dsc1, Dsg1 and around 30% for disordered termini regions (PKP1-N, PG-N, DP-N, PG-C). **C)** Localization densities for PKP1 layer (PG-S density is shown for reference). **D)** Localization densities for PG-layer. **E)** The densities for PG-S and DP-S with PG-C as a reference. **F-G)** Localization densities for the cadherins. Panel G is a rotated view of Panel F. See also Fig. S2-S4.

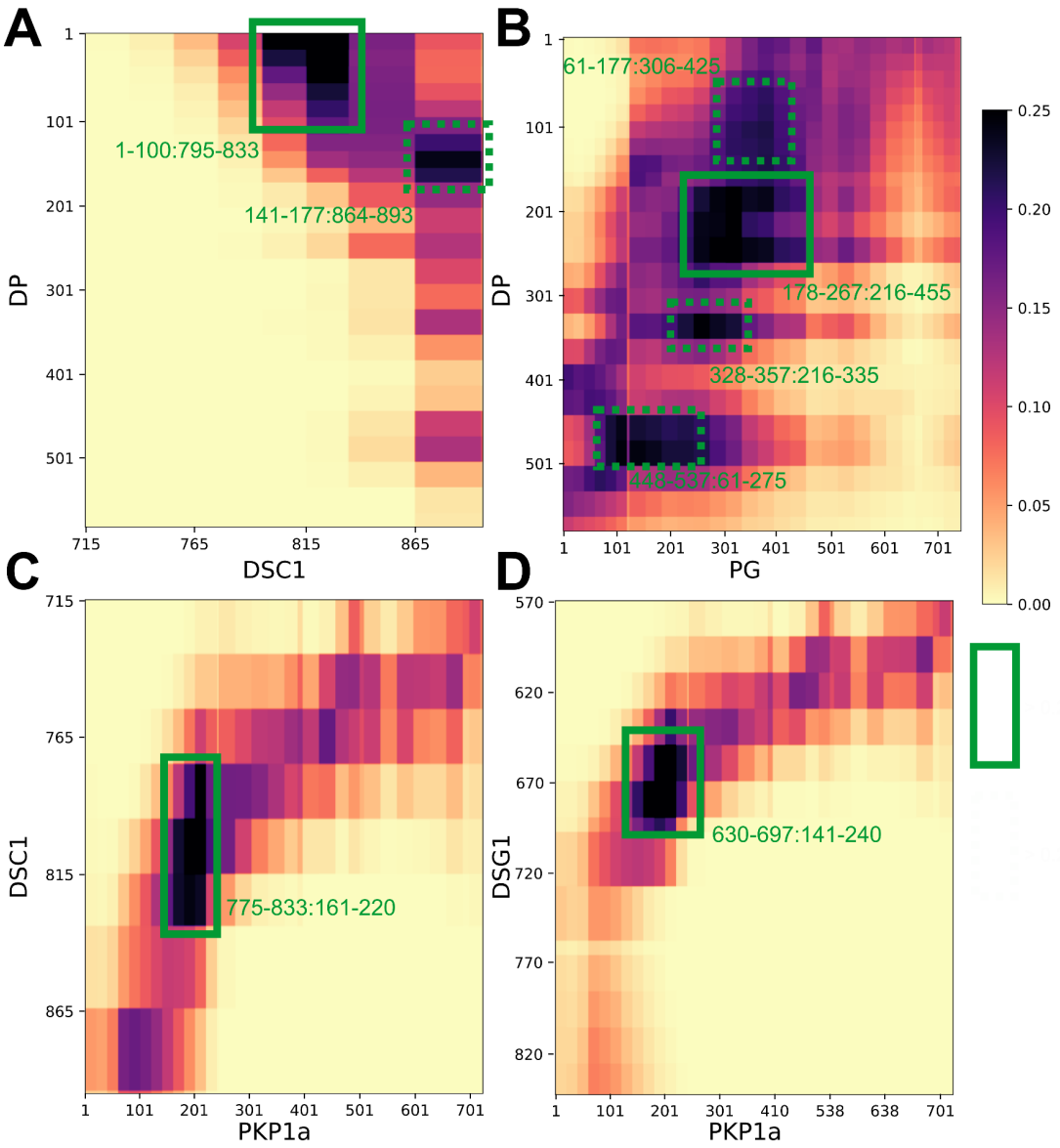

Figure 4 Novel ODP protein-protein interfaces Protein-protein contact maps for DP-Dsc1 (**A**), DP-PG (**B**), Dsc1-PKP1 (**C**), and Dsg1-PKP1 (**D**) pairs. Maps are colored by the proportion of the models in the major cluster where the corresponding bead surfaces are within contact distance (10 Å). Rectangles with solid green (broken green) lines outline novel contacts present in >25% (>20%) of the models. See also Fig. S5, Table S4.

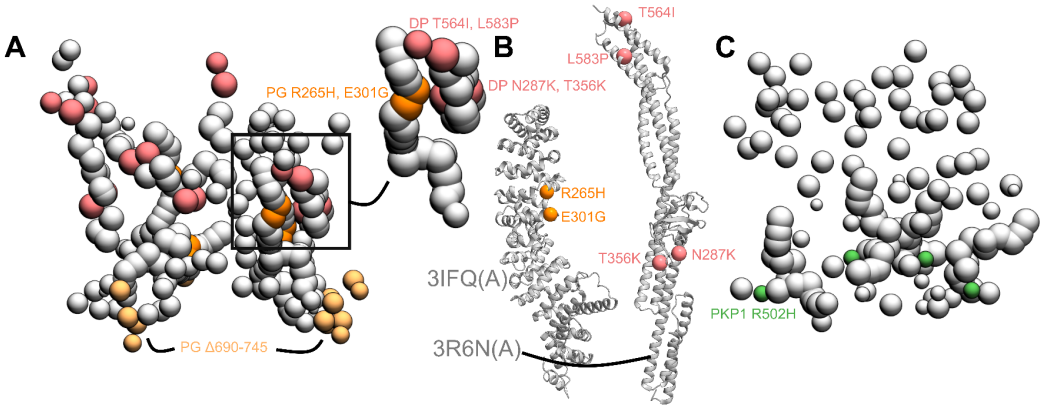

Figure 5 Disease-associated mutations mapped onto the integrative structure A) Cluster center bead model showing mutations in PG and DP. Mutations in DP-S (pink), PG-S (orange) and PG-C (light orange) are colored as per Fig. 1. Remaining beads of DP and PG are shown in gray. Top right shows a zoomed-in version of a novel predicted PG-DP interface harboring disease mutations. **B)** PG-S and DP-S mutations mapped onto the corresponding structures 3IFQ(A) and 3R6N(A) (Choi et al., 2009; Choi and Weis, 2011) **C)** Bead model showing mutations in PKP1-S (green). See also Table S5.

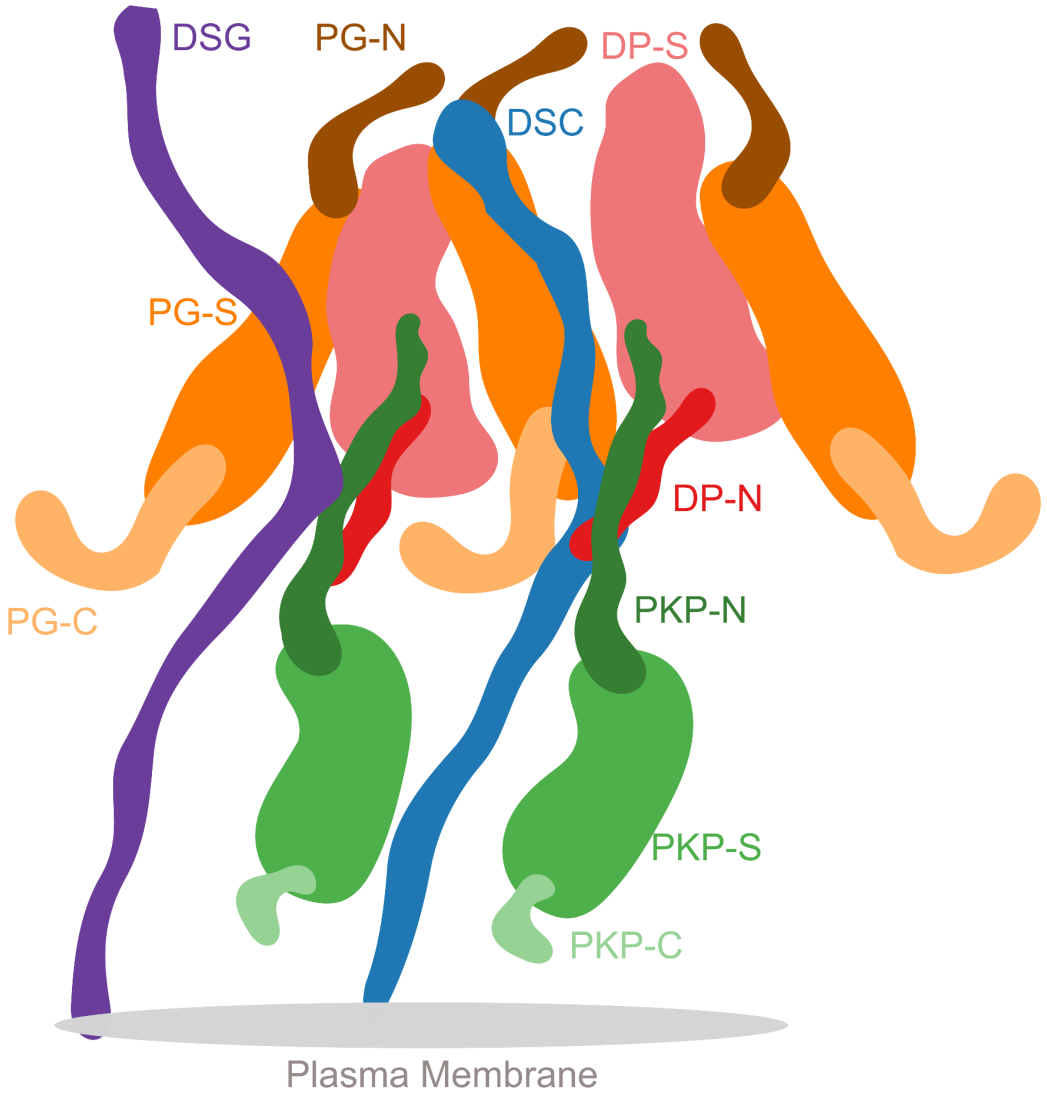

Figure 6 Schematic of the desmosome ODP Schematic showing the salient features of the protein organization and protein-protein interfaces in the ODP. Wavy thick lines represent potentially disordered regions without known structure (DC, PKP-N, PG-N, DP-N, PKP-C, PG-C). Larger shapes represent regions with known structure (PG-S, DP-S, PKP-S).

Supplementary Information

Section 1: Modeling

Section: 1.1 Stoichiometry Runs

The number of molecules that could be accommodated in the PG layer was unclear since it was not possible to unambiguously dock the molecules manually in this region. A higher number would result in an overcrowded model or push molecules outside the tomogram space and be spatially unconstrained by the tomogram, while a smaller number would not be enough to explain the whole EM map. To determine this, we ran independent modeling runs that only included the PG and DP molecules and the PG layer EM density. We ran six simulations, each with an equal number of PG and DP ranging from 2 to 7. For each of these runs, the representation and sampling followed the IMP modeling protocol (Saltzberg et al., 2021; Viswanath et al., 2017b). The restraints applied included the EM restraint, the immuno-EM restraint, the excluded volume restraint, connectivity restraint, and the cylinder restraint to ensure that the molecules are not too far away from the tomogram. 3 million models were simulated per run. For each run, after filtering the models sampled before equilibration, the top 10% models were determined and the cross-correlation coefficient of these models to the cryo-electron tomogram was computed (Bonomi et al., 2019) (Fig. S1). The number of PG/DP molecules was taken to be four. It was only slightly lower in average cross-correlation than the value for three PG/DP molecules. Selecting four PG/DP copies allowed selecting four PKPs in the PKP layer including the central PKP without introducing any asymmetries in the selection. It also allowed us to maintain 1:1:1:1 stoichiometry for PG:DP:PKP:DC with equal numbers of Dsc and Dsg (2 each). Finally, it was consistent with previous studies which showed that no more than four PG and DP molecules each could fit in the PG layer (Al-Amoudi et al., 2011).

Section: 1.2 Restraints

Distance threshold for sequence connectivity restraint

To set up the connectivity restraint, we need to scale the inter-bead distance to allow the more disordered N/C termini as well as the DC proteins to span a greater end-to-end distance compared to the globular protein domains. For each protein domain with at least a partial disorder (for example, PKP-N, DP-N, etc), we first find the radius of gyration of this fragment assuming the fragment to be completely disordered, using $R_g = 1.92N^{0.6}$(Kohn et al., 2004) where *N* is the number of residues. We model the fragment as a chain of monomers composed of *n* statistically independent segments each of length *a*. We assume *n* to be approximately equal to the number of beads in our representation of the fragment as two adjacent beads are free to be in any relative orientation without any other restraints; *a* is then the inter-bead distance. We then use a relation between the RMS Distance between the two ends of the chain ($R_F = an^{0.6}$) and the radius of gyration (R_G) to estimate *a* for our fragment: ${R_G}^2/{R_F}^2 = 25/176$ (Teraoka, 2002). Another estimate for *a* is calculated internally in IMP and comes from the assumption that the fragment is globular (Alber et al., 2007). The final scaling depends on the weighted sum of these two estimates, the weights corresponding to the portion of the fragment predicted to be disordered by PSIPRED (Buchan and Jones, 2019). Given an estimate of *a*, we can calculate the surface-to-surface distance for adjacent beads to create an harmonic upper bound restraint such that the beads are only penalized when they are farther than this distance apart. We use the maximum end-to-end distance (*an*) and find the bead surface-to-surface distance needed to achieve this end-to-end distance. This is approximated by the following relation where *r* is the typical radius of a bead in our model: $d = (an - 2r) / n - 2r$. The calculated scale matches the

scale calculated using a more accurate measure for $R_G^2/R_F^2 \approx 0.95/6$ given by renormalization theory (Teraoka, 2002) up to rounding. However, the scale is only a heuristic parameter and the results obtained are relatively robust to its exact value.

Section 2: Analysis

Section: 2.1 Filtering based on Autocorrelation

In order to filter a computationally feasible subset of models from the large set of sampled models, we first remove the initial few models based on statistical testing (Chodera, 2016; Saltzberg et al., 2021), to consider only the models after equilibration assumed to be in the stationary distribution. Next, we only take every 20th model in the MCMC sampling run (PMI analysis parameter nskip=2, writing every 10th frame to disk). To identify an appropriate number of models to skip, we ran eight independent single-replica runs with all the restraints. We analyzed the spatial autocorrelation of the XYZ coordinates of each bead along the sampling trajectory. We chose as our cutoff the smallest number of sampling steps after which the autocorrelation of all the beads had fallen to at most 85-90%. This allows us to remove the highly correlated models to obtain an independent set of models to analyze downstream.

Section: 2.2 Cross-Correlation of Localization Densities and cryo-electron tomograms

We first computed the predicted localization density by combining the densities from all modeled proteins for the major cluster separately for the PKP layer (PKP-S) and PG-Layer (PG-N,S,C, DP-N,S). We then calculated the cross-correlation between the predicted density and the reference cryo-electron tomogram by calculating the Pearson correlation between the voxel-wise values in the two maps. This is calculated at all grid points at a voxel spacing of 5Å spread over the volume enclosing both the predicted localization density and the cryo-TM map. The values of the maps at these grid points were found by interpolation (RegularGridInterpolator in scipy (Virtanen et al., 2020)). This is similar to calculating Correlation around mean in UCSF Chimera (fitmap) except that the Chimera calculation only involves the non-zero grid points of the reference map, causing the correlation value to change depending on the order of the two maps.

Supplementary Figures

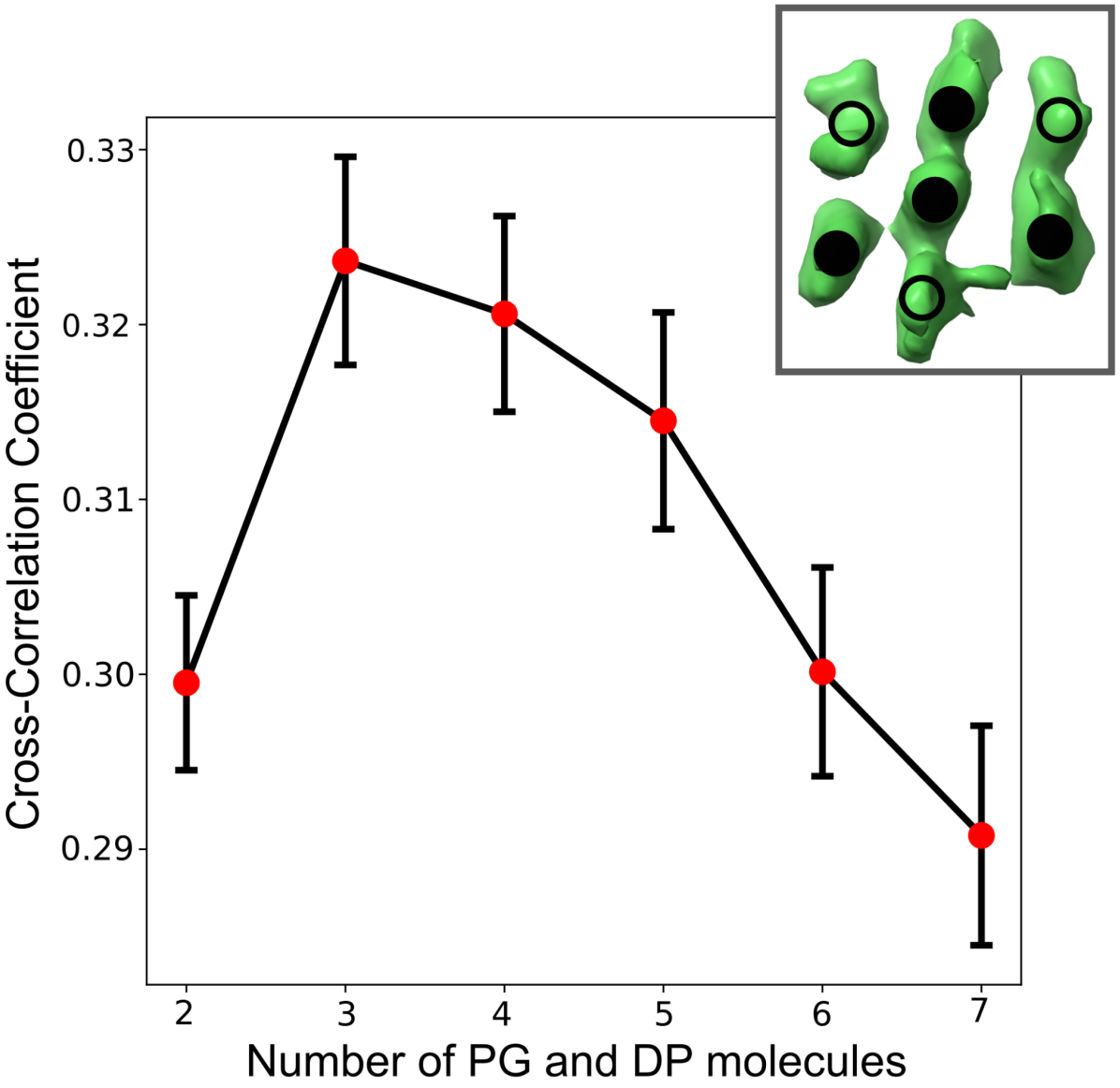

Figure S1 Estimating the number of PG and DP copies and selecting the layout for PKP copies The graph shows the results for independent stoichiometry runs (see Methods, Stage 2 and Supplementary Section 1.1) with the number of PG and DP molecules ranging from 2 to 7. The boxplot marks the mean (red dot) and the standard deviation (black error bars) of the cross-correlation coefficient (Bonomi et al., 2019) between the top 10% best-scoring models (based on their cross-correlation coefficient) and the cryo-tomogram in each of the runs. **(Inset)** The seven densities in the PKP layer of the tomogram (top view) of which four were full-length PKPs in our model (filled circles) and three were fixed, non-interacting PKP structured regions (empty circles).

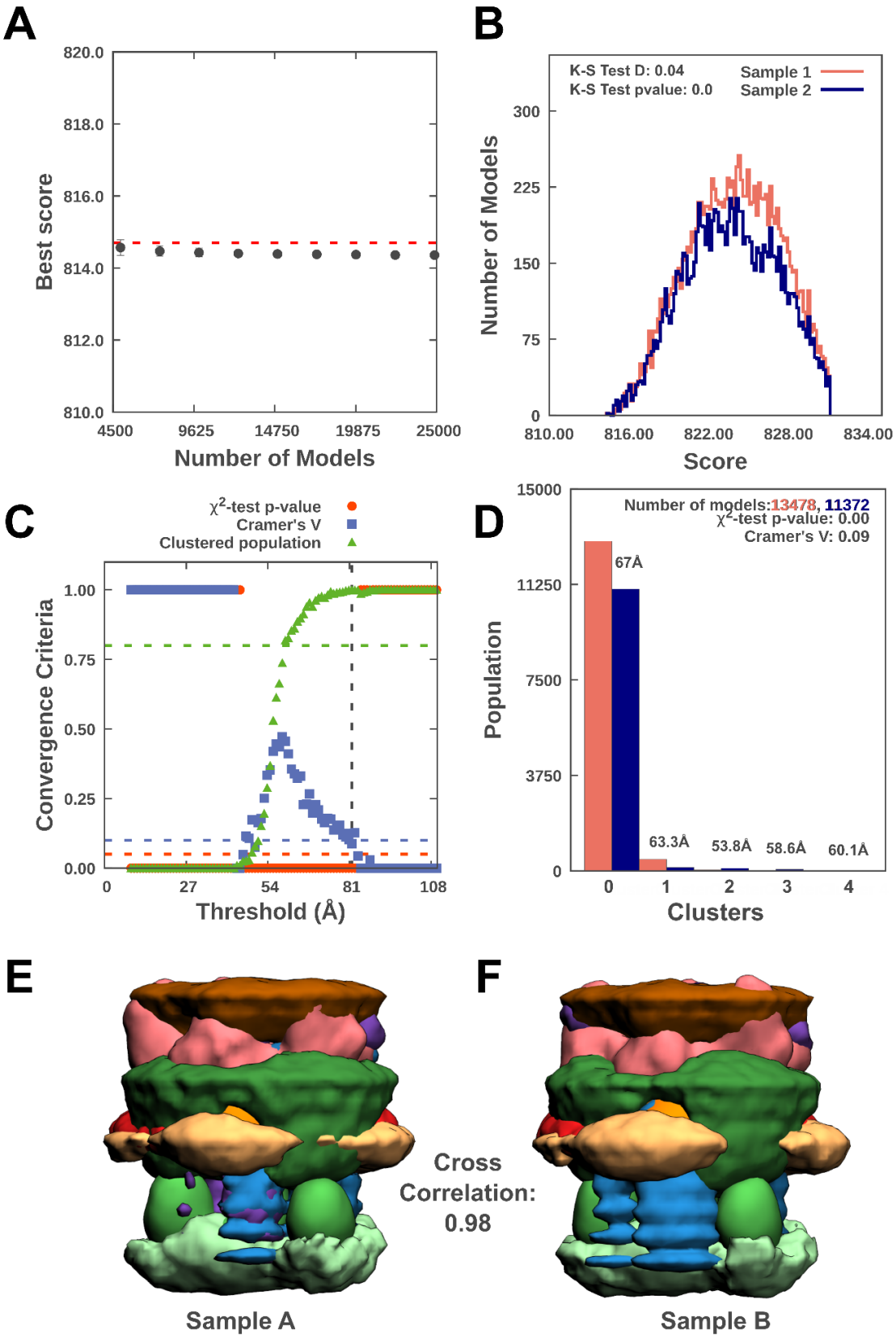

Figure S2 Sampling Exhaustiveness protocol for Desmosomal ODP (Viswanath et al., 2017b) **A)** Test for the convergence of the model score for the 24016 good-scoring models. The scores do not continue to improve as more models are added independently. The error

bar represents the standard deviations of the best scores, estimated by repeating sampling of models 10 times. The red dotted line indicates a lower bound reference on the total score. **B)** Testing the similarity of model score distributions between samples 1 (red) and 2 (blue). The difference in the distribution of scores is significant (Kolmogorov-Smirnov two-sample test p-value less than 0.05) but the magnitude of the difference is small (the Kolmogorov-Smirnov two-sample test statistic D is 0.04); thus, the two score distributions are effectively equal. **C)** Three criteria for determining the sampling precision (Y-axis), evaluated as a function of the RMSD clustering threshold (X-axis). First, the p-value is computed using the χ^2-test for homogeneity of proportions (red dots). Second, an effect size for the χ^2-test is quantified by the Cramer's V value (blue squares). Third, the population of models in sufficiently large clusters (containing at least 10 models from each sample) is shown as green triangles. The vertical dotted grey line indicates the RMSD clustering threshold at which three conditions are satisfied (p-value > 0.05 [dotted red line], Cramer's V < 0.10 [dotted blue line], and the population of clustered models > 0.80 [dotted green line]), thus defining the sampling precision of 82 Å. **D)** Populations of sample 1 and 2 models in the clusters obtained by threshold-based clustering using the RMSD threshold of 82 Å. Cluster precision is shown for each cluster **E-F)** Comparison of localization probability densities of models from sample A and sample B for the major cluster. The cross-correlation of the density maps (see Supplementary section 2.2) of the two samples is greater than 0.98.

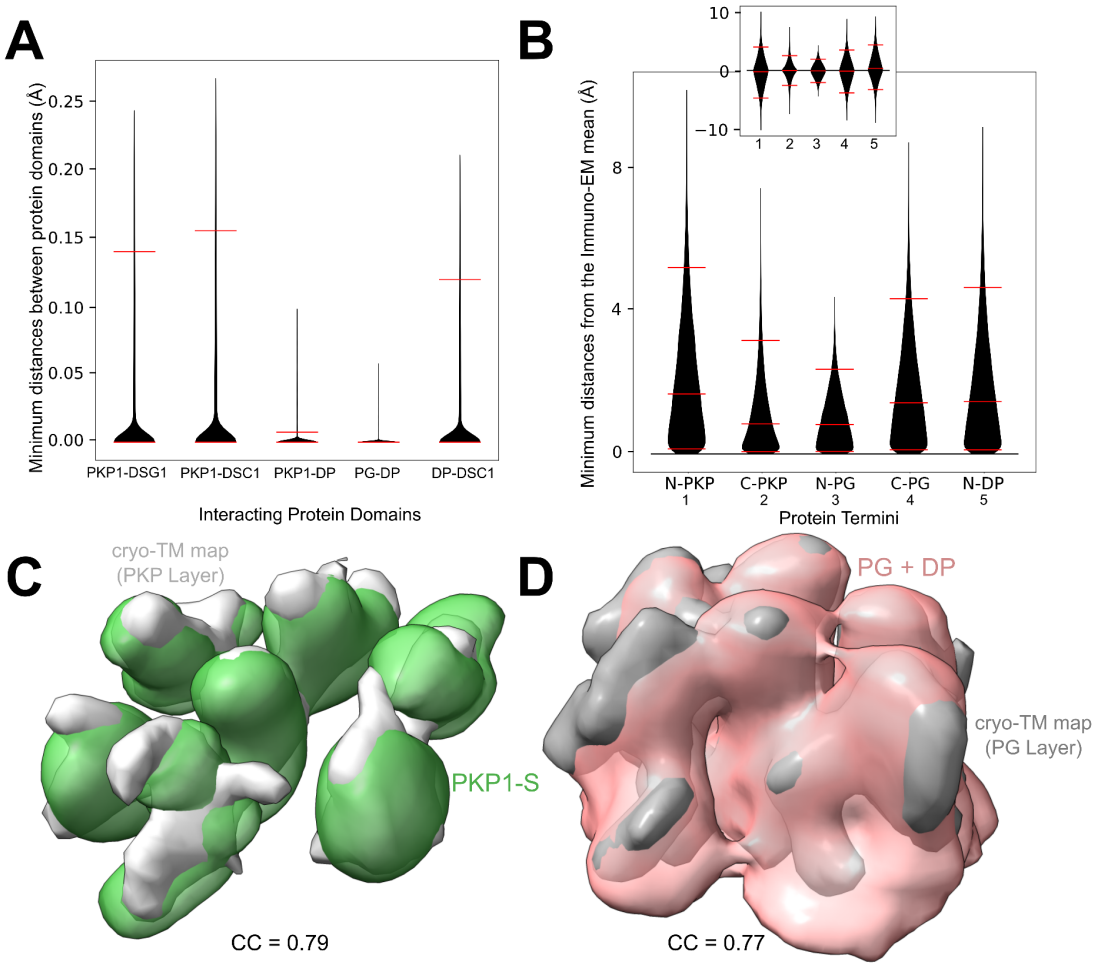

Figure S3 Fit to data used in modeling A) Fit to the data from biochemical experiments formulated as binding restraints (Methods). Each violin corresponds to the absolute closest distance between the two interacting domains across all protein copies for a model in the

major cluster (Methods). Each distribution corresponds to a restraint in Table S2A. Red horizontal lines correspond to 5th, 50th and 95th percentile (in **A** and **B** both) after outlier removal. **B)** Fit to immuno-EM data (North et al., 1999). Each violin corresponds to the absolute difference between the experimental mean and the model-predicted distance from the membrane (Methods). The inset shows the same information without the absolute value (i.e. signed difference). **C-D)** Fit to the cryo-tomogram for the PKP Layer (**C**) and the PG Layer (**D**). Densities from the model (colored) are shown along with the segmented densities from the tomogram, EMD-1703 (Al-Amoudi et al., 2011)(Methods). The cross-correlation (CC) (Methods, Supplementary Section 2.2) is mentioned for each of the fits. PKP1-S density (including the non-interacting PKP1 molecules) and the PG + DP density are visualized at a ~10% threshold and the tomogram is visualized at the recommended threshold. See also Fig. 3, Table S2.

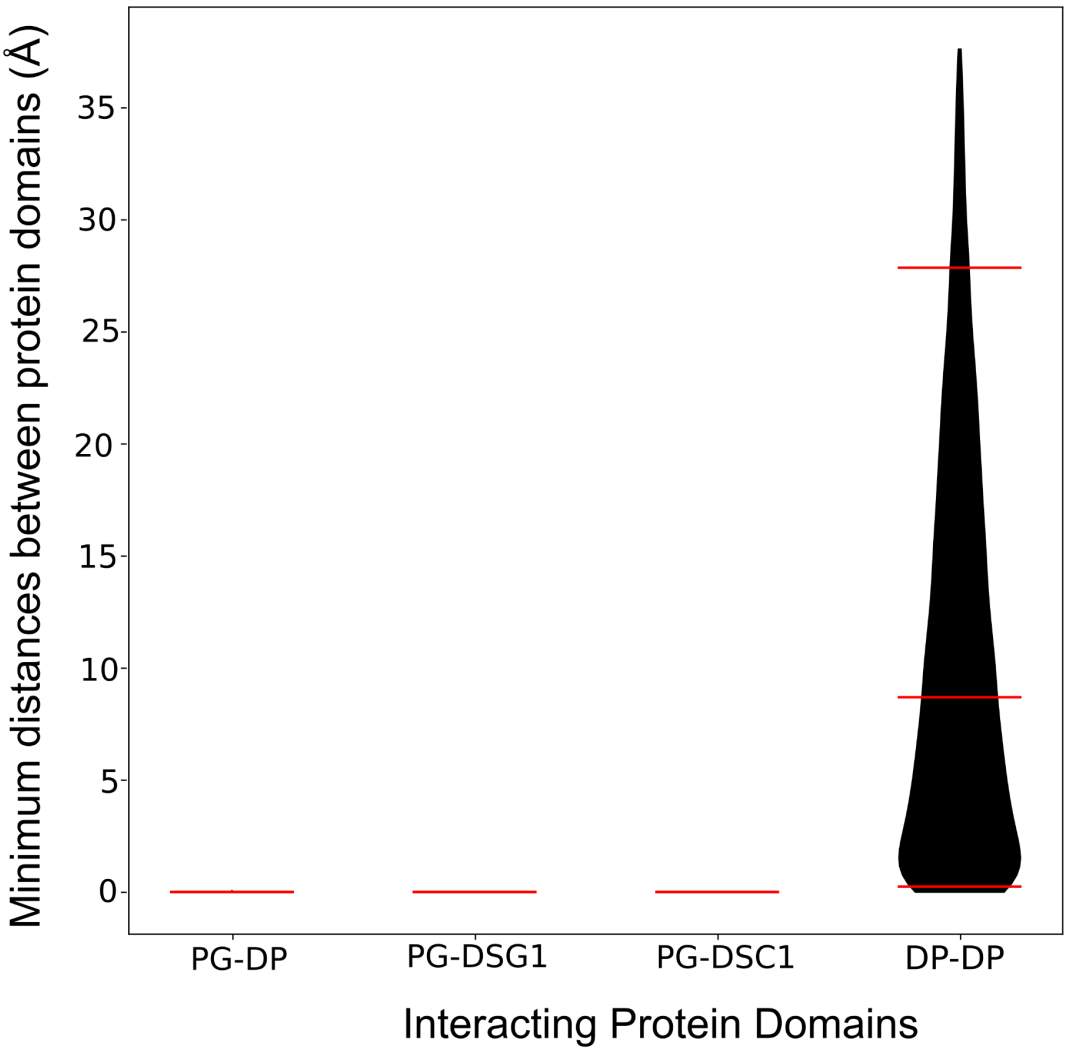

Figure S4 Fit to data not used in modeling Fit to the data from biochemical experiments not used for modeling (Methods). Each violin corresponds to the closest distance between the two interacting domains across all copies for a model in the major cluster (Methods). Each distribution corresponds to a row in Table S3. Red horizontal lines correspond to 5th, 50th and 95th percentile after outlier removal. See also Table S3.

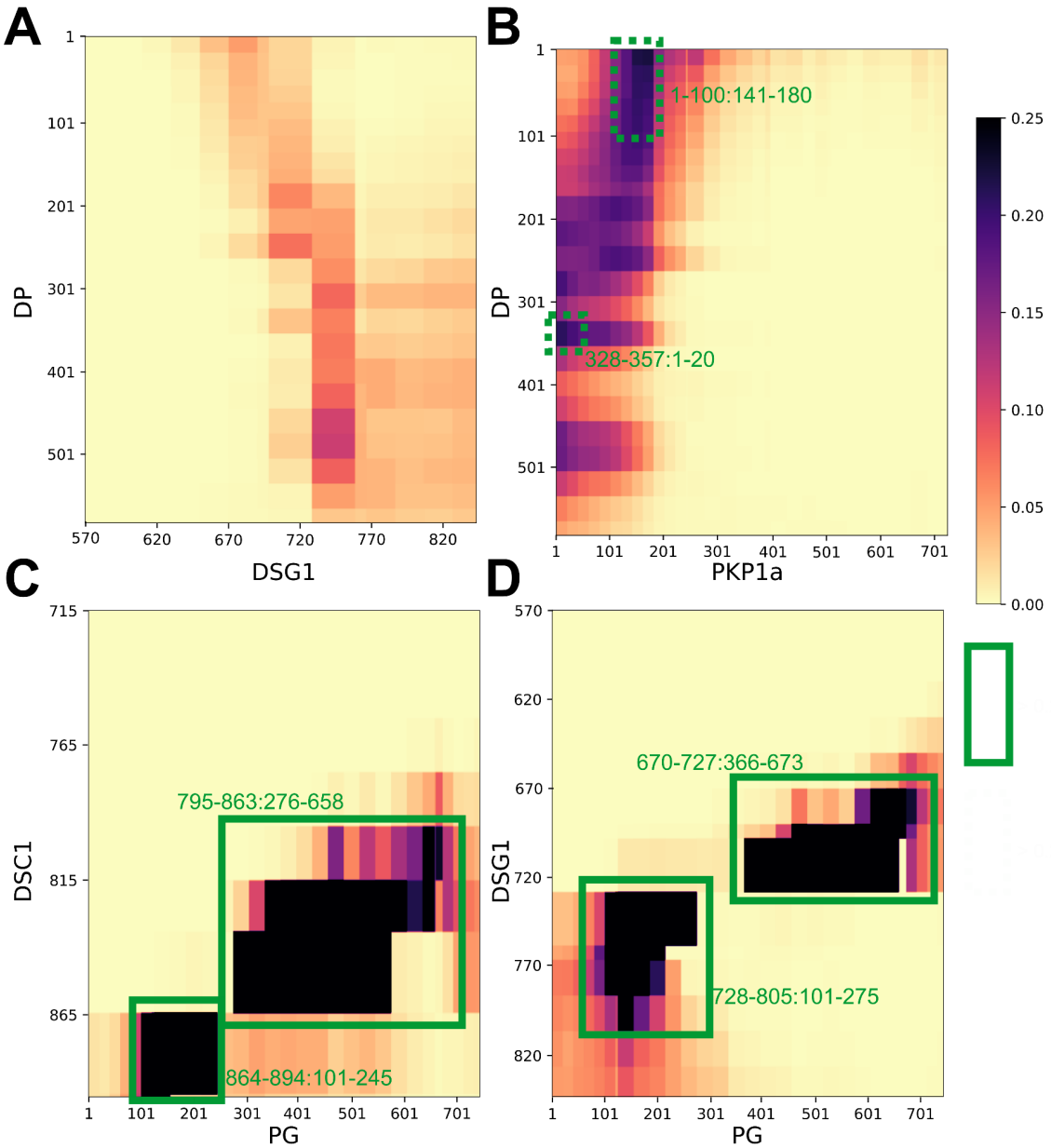

Figure S5 Additional contact maps Protein-protein contact maps for DP-Dsg1 (**A**), DP-PKP1 (**B**), PG-Dsc1 (**C**), and PG-Dsg1 (**D**) pairs. Maps are colored by the proportion of the models in the major cluster where the corresponding two bead surfaces are within contact distance (10 Å). Rectangles with solid green (broken green) lines outline novel contacts present in >25% (>20%) of the models. See also Fig. 4 and Table S4.

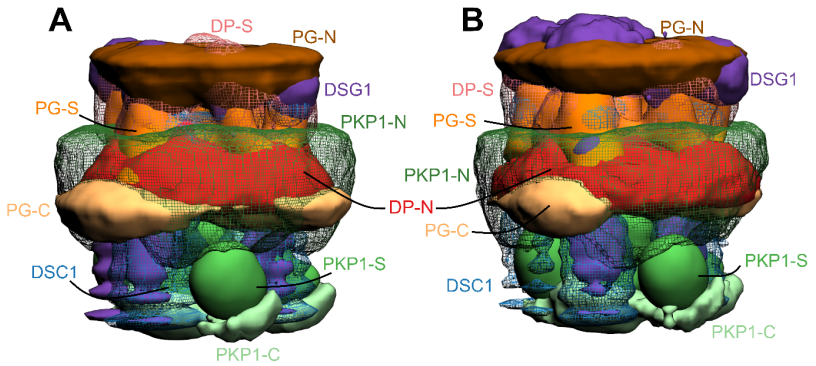

Figure S6 Integrative structures of other ODPs The localization densities for the stratified epithelium ODP with PKP3 (**A**) and basal epithelium ODP with PKP3, Dsg3, Dsc3 (**B**). Colors are in accordance with Fig 1. For more details of modeled proteins in these ODPs, see Table S1. Density thresholds are the same as in Fig 3.

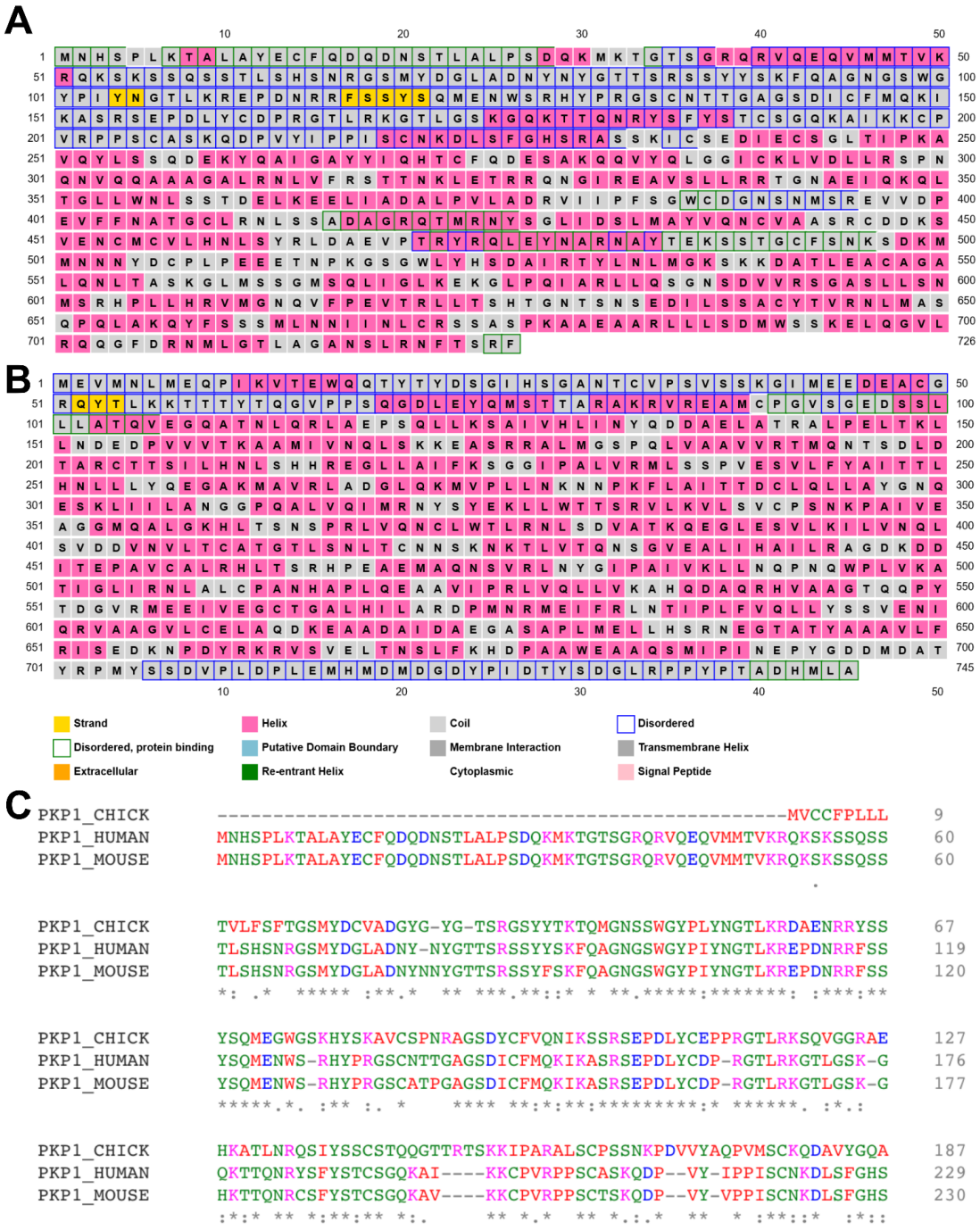

Figure S7 Prediction of disorder in PG and PKP1 and sequence conservation for PKP1
The PSIPRED (Buchan and Jones, 2019) output for PG (**A**) and PKP1 (**B**) are shown with the color scheme displayed below panel B. **C)** Sequence conservation output using Clustal Omega (Goujon et al., 2010; Sievers et al., 2011) for Chick, Human and Mouse PKP1$^{1\text{-}229}$. The colors represent different kinds of amino acids (acidic: blue, basic: magenta, hydrophobic/small: red, Hydroxyl/sulfhydryl/amine/G: green). An asterisk (*) represents an exact residue match, a colon (:) represents a strongly similar match and a period (.) represents a weakly similar match.

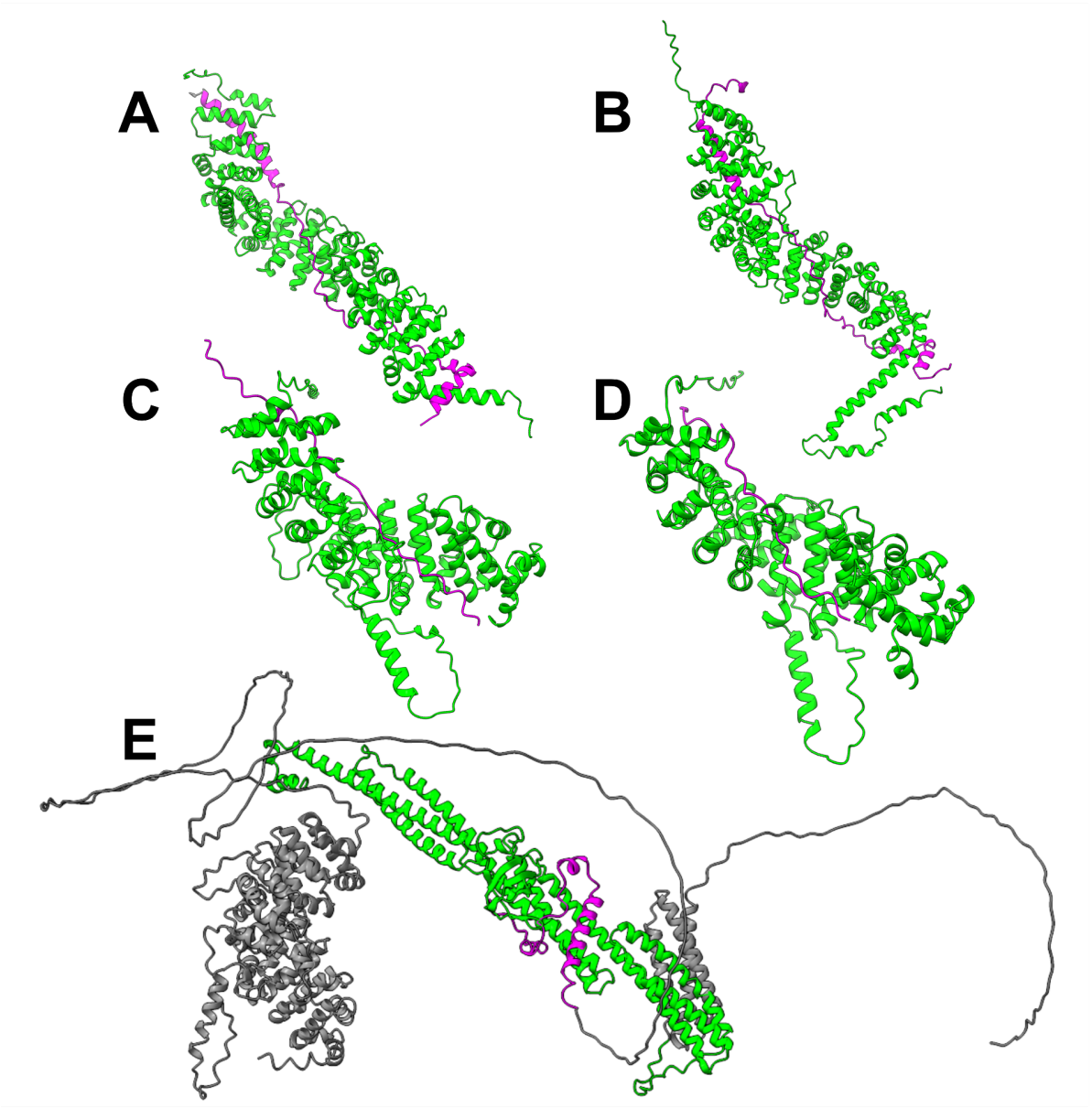

Figure S8 Alphafold2-Multimer predictions The best ranked model (based on PTM + IPTM score) is displayed for each protein pair. **A)** PG$^{107-697}$ (lime) and Dsc1$^{812-894}$ (magenta) **B)** PG$^{75-699}$ (lime) and Dsg1$^{680-768}$ (magenta) **C)** PKP1$^{229-747}$ (lime) and Dsc1$^{722-759}$ (magenta) **D)** PKP1$^{230-744}$ (limer) and Dsg1$^{712-737}$ (magenta) **E)** DP$^{160-584}$ (lime) and PKP1$^{1-60}$ (magenta). Rest of the residues of DP and PKP1 that are away from the interface are in gray. The PKP1 isoform used in **C, D, E** is PKP1b (in contrast to PKP1a used for integrative modeling in the paper) which has an extra 20-residue segment (PKP1b$^{412-432}$). This however does not affect any conclusions in the paper or the figure.

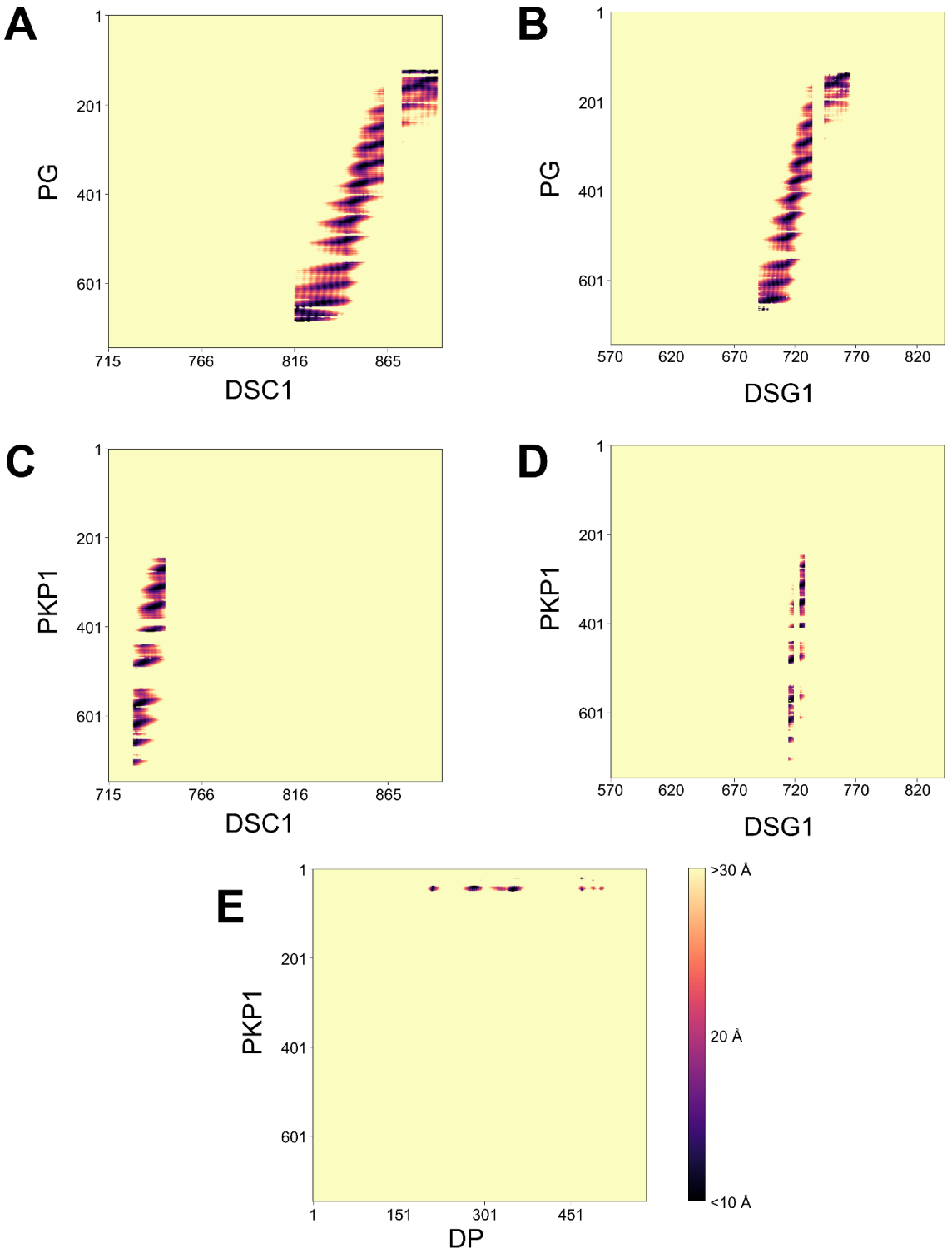

Figure S9 Contacts predicted by Alphafold2-Multimer Contact maps (Cα-Cα distance between residues) for PG-Dsc1 (**A**), PG-Dsg1 (**B**), PKP1-Dsc1(**C**), PKP1-Dsg1(**D**), and PKP1-DP (**E**) based on the best ranked AF2-Multimer model for each protein pair. The map is colored by the distance between Cα atoms of a residue pair. Regions in contact distance (Cα-Cα distance <10Å) are colored black. Distances are shown only for residue pairs with reliable AF2 prediction (PAE <5 and pLDDT >70 for each residue in the pair). Residue pairs without reliable AF2 prediction are colored as >30Å irrespective of their actual Cα-Cα distance. The PKP1 isoform used in **C, D, E** is PKP1b as in Fig. S8.

Supplementary Tables

Protein	Residue Ranges	Structure: PDB (chain) or unknown	Type of ODP	Uniprot ID
DP1	1-177	Unknown	BASAL, UPPER 1, UPPER 3	P15924
	178-584	3R6N (A)		
PKP1a	1-243	Unknown	UPPER 1	Q13835-2
	244-700	1XM9 (A)		
	701-726	Unknown		
PG	1-125	Unknown	BASAL, UPPER 1, UPPER 3	P14923
	126-673	3IFQ (A)		
	674-745	Unknown		
DSG1	570-697	Unknown	UPPER 1, UPPER 3	Q02413
	698-765	3IFQ (C)		
	766-842	Unknown		
DSC1a	715-833	Unknown	UPPER 1, UPPER 3	Q08554-1
	834-894	3IFQ (C)		
PKP3a	1-318	Unknown	UPPER 3, BASAL	Q9Y446-1
	319-781	1XM9 (A)		
	782-797	Unknown		
DSG3	641-769	Unknown	BASAL	P32926
	770-863	3IFQ (C)		
	864-999	Unknown		
DSC3a	712-819	Unknown	BASAL	Q14574-1
	820-896	3IFQ (C)		
	897	Unknown		

Table S1 Modeled protein domains The different domains of the modeled proteins are shown along with their residue ranges, structure, and the ODP models in which they were used. BASAL refers to basal epithelium, UPPER 1 and UPPER 3 refer to stratified epithelium with PKP1 and PKP3 respectively. UPPER 1 is the ODP displayed in the main text. Regions of unknown structure were represented as flexible 20-residue beads, while regions of known structure were represented as rigid bodies consisting of 30-residue beads. The colors refer to domains without a known structure (red), domains which are homology modeled on a structure template of an isoform or homolog (yellow), and domains which have the structure in the PDB (green). Only the domains in the ODP were modeled; extracellular and transmembrane domains ($Dsg3^{1-640}$, $Dsg1^{1-569}$, $Dsc3^{1-711}$, $Dsc1^{1-714}$) or domains outside the ODP ($DP^{585-2871}$, $Dsg1^{843-1049}$) were not modeled (Al-Amoudi et al., 2011; Garrod and Chidgey, 2008; Nilles et al., 1991).

Protein 1	Protein 2	Domain 1	Domain 2	Experiment	Reference	Data for Restraint Protein 1 residues: Protein 2 residues	Type of ODP
PKP1a	DSG1	70-213	570-1049	Y2H	(Hatzfeld et al., 2000)	70-213:570-842	UPPER 1
		1-286	568-1049	Y2H	(Kowalczyk et al., 1999)		
		1-726	499-1049	OA	(Smith and Fuchs, 1998)		
PKP1a	DSC1	1-726	715-894	OA	(Smith and Fuchs, 1998)	1-726:715-894	UPPER 1
DP	DSC1	1-176	715-894	OA	(Smith and Fuchs, 1998)	1-176:715-894	UPPER 1, UPPER 3
PKP1a	DP	1-168	1-584	Y2H	(Hatzfeld et al., 2000)	1-168:1-584	UPPER 1
		1-286	1-584	Y2H, co-IP, Loc	(Kowalczyk et al., 1999)		
		1-726	1-1014	OA	(Smith and Fuchs, 1998)		
		1-726	1-2871	Loc	(Bornslaeger et al., 2001)		
PG	DP	1-745	1-2871	OA	(Smith and Fuchs, 1998)	1-745:1-584	UPPER 1, UPPER 3, BASAL
PKP3a	DSC1	1-18 & 51-797	715-894	Y2H, co-IP, Loc	(Bonné et al., 2003)	(1-18 + 51-797):715-894	UPPER 3
PKP3a	DSG1	1-18 & 51-293	519-715 Or 715-1k	Y2H, co-IP	(Bonné et al., 2003)	(1-18 + 51-293):570-842	UPPER 3
PKP3a	DP	19-50	1-2871	Loc	(Bonné et al., 2003)	(1-18 + 51-293):1-584 & 19-50:1-63	UPPER 3, BASAL
		1-18 & 51-293	1-584	Y2H, co-IP, Loc	(Bonné et al., 2003)		
		19-50	1-63	Y2H	(Bonné et al., 2003)		
PKP3a	DSC3	1-18 & 51-797	712-897	Y2H, co-IP	(Bonné et al., 2003)	(1-18 + 51-797):712-897	BASAL
PKP3a	DSG3	1-18 & 51-797	641-999	Y2H, co-IP	(Bonné et al., 2003)	(1-18 + 51-797):641-999	BASAL

Table S2A Binding Restraints The protein-protein binding restraints are shown along with the experimental data they are based on. The restraint is formulated by including only the residues that are modeled in our ODP model. If multiple experiments provide data for a protein pair, the data from the experiment with the highest resolution is used. In Columns 3, 4 (and 5), green background (or text) represents the highest resolution information that was used to formulate the restraint and Gray represents other information that is at a lower resolution than the restraint. Experiment abbreviation are as follows: Y2H: Yeast 2 Hybrid, OA: Overlay Assay (in-vitro), Loc: Co-Localization assays, Co-IP: Co-Immunoprecipitation. The ODPs they are used in are mentioned in the last column: BASAL refers to the basal epithelium, UPPER 1 and UPPER 2 refer to the stratified epithelium with PKP1 and PKP3 respectively. UPPER 1 is the ODP model shown in the main paper.

Protein Termini	Residues	Mean dist. (Å)	SE (Å)
DP-N	1-189	103	9.8
PG-N	1-106	229	4.5
PG-C	666-738	108	9
PKP1-N	1-285	158	11
PKP1-C	286-726	42	11
PKP3-N	1-359	158	11
PKP3-C	360-797	42	11

Table S2B Immuno-EM Restraints The antibody binding domains for the different termini, mean distance of the termini from the plasma membrane, and the respective standard errors are shown (North et al., 1999). Domain names are in accordance with Fig. 1.

Protein 1	Protein 2	Domain 1	Domain 2	Experiment	Reference	Data for validation Protein 1 residues: Protein 2 residues
DP	DP	1-1014	1-176	OA	(Smith and Fuchs, 1998)	1-584:1-176
PG	DP	FULL	FULL	Y2H, co-IP	(Kowalczyk et al., 1997)	1-745:1-584
PG	Dsg1	FULL	701-768	co-IP, Loc	(Troyanovsky et al., 1994a)	123-632:701-768
PG	Dsg1	123-632	FULL	co-IP	(Wahl et al., 1996)	
PG	Dsg1	FULL	663-958	ITC	(Choi et al., 2009)	
PG	Dsc1	FULL	858-894	co-IP, Loc	(Troyanovsky et al., 1994b)	123-632: 858-894
PG	Dsc1	123-632	FULL	co-IP	(Wahl et al., 1996)	
PG	Dsc1	FULL	795-894	ITC	(Choi et al., 2009)	
DP	Dsc1	FULL	728-740	co-IP, Loc	(Troyanovsky et al., 1994b)	1-584:728-740
PG	Dsc1	FULL	728-740	co-IP, Loc	(Troyanovsky et al., 1994b)	123-632: 728-740

Table S3 Validation protein-protein binding data not used in modeling The protein-protein binding data and the corresponding references not used in modeling are shown. Data for validation was obtained using the same reasoning as in Table S2A. Experiment abbreviation are as follows: Y2H: Yeast 2 Hybrid, OA: Overlay Assay (in-vitro), Loc: Co-Localization assays, Co-IP: Co-Immunoprecipitation, ITC: Isothermal Calorimetry. Gray text represents the information that was not validated by the ensemble of models.

Protein 1	Residues in protein 1	Protein 2	Residues in protein 2
DP	1-20	Dsc1	795-833
DP	21-60	Dsc1	815-833
DP	21-40	Dsc1	795-814
DP	61-100	Dsc1	815-833
DP	141-177	Dsc1	864-893
DP	178-237	PG	306-335
DP	238-267	PG	276-335
DP	61-80	PG	366-395
DP	81-100	PG	336-395
DP	101-140, 161-177	PG	306-395
DP	141-160	PG	306-425
DP	178-207	PG	246-305, 336-425
DP	208-237	PG	216-305, 336-395
DP	238-267	PG	216-275, 336-455
DP	268-297	PG	246-275
DP	328-357	PG	216-335
DP	448-507	PG	81-120, 126-275
DP	508-537	PG	61-100
DP	1-60	PKP1	141-180
DP	61-100	PKP1	141-160
DP	328-357	PKP1	1-20
Dsc1	795-814	PG	636-658
Dsc1	815-833	PG	336-605, 636-658
Dsc1*	834-863	PG	276-575
Dsc1*	864-894	PG	101-245
Dsc1*	894	PG	101-155
Dsc1	795-814	PG	661-673
Dsc1	815-833	PG	606-635
Dsg1	670-689	PG	606-673
Dsg1	690-697	PG	456-673
Dsg1*	698-727	PG	366-658
Dsg1*	728-757	PG	101-275
Dsg1*	758-765	PG	101-215
Dsg1	766-785	PG	101-185
Dsg1	786-805	PG	126-155
Dsg1	670-689	PG	674-693
Dsg1*	698-727	PG	659-660
Dsg1	766-785	PG	186-215
Dsc1	775-794	PKP1	201-220
Dsc1	795-814	PKP1	181-220
Dsc1	795-814	PKP1	161-180
Dsc1	815-833	PKP1	161-220
Dsg1	650-669	PKP1	181-220
Dsg1	670-689	PKP1	161-220
Dsg1	630-649	PKP1	201-220
Dsg1	650-669	PKP1	221-240
Dsg1	670-689	PKP1	141-160
Dsg1	690-697	PKP1	181-200

Table S4 Protein-Protein contacts in the ODP All the contacts (bead surface-to-surface distance of less than 10 Å) identified in at least 20% (yellow) to 25% (green) of the models in the ensemble are shown (Methods). Contacts consistent with sub-complexes of known structure are marked with an asterisk in column 1.

Protein Domain	Residues	Mutation	Reference	Disease
PG-N	19	T → I	(Den Haan et al., 2009)	Naxos disease
PG-S	265	R → H	(Erken et al., 2011)	Naxos disease
PG-S	301	E → G	(Marino et al., 2017)	Naxos disease
PG-C	680-745	WEAAQSMIPI → GGCPEHDSHQ + Δ690-745	(McKoy et al., 2000)	Naxos disease
DP-S	287	N → K	(Whittock et al., 2002)	Skin Fragility - Woolly Hair Syndrome
DP-S	356	T → K	(Pigors et al., 2015)	Carvajal syndrome
DP-S	564	T → I	(Boulé et al., 2012; Keller et al., 2012)	Carvajal syndrome
DP-S	583	L → P	(Keller et al., 2012)	Carvajal syndrome
PKP1-S	502	R → H	COSMIC	Cancer Mutation
PG-N	4	M → V	COSMIC	Cancer Mutation
Dsg1	788	E → K	COSMIC	Cancer Mutation
Dsc1	841	Y → F	COSMIC	Cancer Mutation

Table S5 Mutations The mutations of interest in the different protein domains are shown along with the pathology associated with the mutation (Fig. 5, Results). Domain names are in accordance with Fig. 1.

References

Acehan, D., Petzold, C., Gumper, I., Sabatini, D.D., Müller, E.J., Cowin, P., Stokes, D.L., 2008. Plakoglobin Is Required for Effective Intermediate Filament Anchorage to Desmosomes. J. Invest. Dermatol. 128, 2665–2675. https://doi.org/10.1038/jid.2008.141

Al-Amoudi, A., Castaño-Diez, D., Devos, D.P., Russell, R.B., Johnson, G.T., Frangakis, A.S., 2011. The three-dimensional molecular structure of the desmosomal plaque. Proc. Natl. Acad. Sci. 108, 6480–6485. https://doi.org/10.1073/pnas.1019469108

Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., Rout, M.P., Sali, A., 2007. Determining the architectures of macromolecular assemblies. Nature 450, 683–694. https://doi.org/10.1038/nature06404

Arvindekar, S., Jackman, M.J., Low, J.K.K., Landsberg, M.J., Mackay, J.P., Viswanath, S., 2022. Molecular architecture of nucleosome remodeling and deacetylase sub-complexes by integrative structure determination. Protein Sci. 31. https://doi.org/10.1002/pro.4387

Bian, Y., Song, C., Cheng, K., Dong, M., Wang, F., Huang, J., Sun, D., Wang, L., Ye, M.,

Zou, H., 2014. An enzyme assisted RP-RPLC approach for in-depth analysis of human liver phosphoproteome. J. Proteomics 96, 253–262. https://doi.org/10.1016/j.jprot.2013.11.014
Bonné, S., Gilbert, B., Hatzfeld, M., Chen, X., Green, K.J., Van Roy, F., 2003. Defining desmosomal plakophilin-3 interactions. J. Cell Biol. 161, 403–416. https://doi.org/10.1083/jcb.200303036
Bonomi, M., Hanot, S., Greenberg, C.H., Sali, A., Nilges, M., Vendruscolo, M., Pellarin, R., 2019. Bayesian Weighing of Electron Cryo-Microscopy Data for Integrative Structural Modeling. Structure 27, 175-188.e6. https://doi.org/10.1016/j.str.2018.09.011
Bornslaeger, E.A., Godsel, L.M., Corcoran, C.M., Park, J.K., Hatzfeld, M., Kowalczyk, A.P., Green, K.J., 2001. Plakophilin 1 interferes with plakoglobin binding to desmoplakin, yet together with plakoglobin promotes clustering of desmosomal plaque complexes at cell-cell borders. J. Cell Sci. 114, 727–738. https://doi.org/10.1242/jcs.114.4.727
Boulé, S., Fressart, V., Laux, D., Mallet, A., Simon, F., De Groote, P., Bonnet, D., Klug, D., Charron, P., 2012. Expanding the phenotype associated with a desmoplakin dominant mutation: Carvajal/Naxos syndrome associated with leukonychia and oligodontia. Int. J. Cardiol. 161, 50–52. https://doi.org/10.1016/j.ijcard.2012.06.068
Buchan, D.W.A., Jones, D.T., 2019. The PSIPRED Protein Analysis Workbench: 20 years on. Nucleic Acids Res. 47, W402–W407. https://doi.org/10.1093/nar/gkz297
Chodera, J.D., 2016. A Simple Method for Automated Equilibration Detection in Molecular Simulations. J. Chem. Theory Comput. 12, 1799–1805. https://doi.org/10.1021/acs.jctc.5b00784
Choi, H.-J., Gross, J.C., Pokutta, S., Weis, W.I., 2009. Interactions of Plakoglobin and β-Catenin with Desmosomal Cadherins. J. Biol. Chem. 284, 31776–31788. https://doi.org/10.1074/jbc.M109.047928
Choi, H.-J., Weis, W.I., 2011. Crystal Structure of a Rigid Four-Spectrin-Repeat Fragment of the Human Desmoplakin Plakin Domain. J. Mol. Biol. 409, 800–812. https://doi.org/10.1016/j.jmb.2011.04.046
Choi, H.-J., Weis, W.I., 2005. Structure of the Armadillo Repeat Domain of Plakophilin 1. J. Mol. Biol. 346, 367–376. https://doi.org/10.1016/j.jmb.2004.11.048
Delva, E., Tucker, D.K., Kowalczyk, A.P., 2009. The Desmosome. Cold Spring Harb. Perspect. Biol. 1, a002543–a002543. https://doi.org/10.1101/cshperspect.a002543
Den Haan, A.D., Tan, B.Y., Zikusoka, M.N., Lladó, L.I., Jain, R., Daly, A., Tichnell, C., James, C., Amat-Alarcon, N., Abraham, T., Russell, S.D., Bluemke, D.A., Calkins, H., Dalal, D., Judge, D.P., 2009. Comprehensive Desmosome Mutation Analysis in North Americans With Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy. Circ. Cardiovasc. Genet. 2, 428–435. https://doi.org/10.1161/CIRCGENETICS.109.858217
Disfani, F.M., Hsu, W.-L., Mizianty, M.J., Oldfield, C.J., Xue, B., Dunker, A.K., Uversky, V.N., Kurgan, L., 2012. MoRFpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order transitioning binding regions in proteins. Bioinformatics 28, i75–i83. https://doi.org/10.1093/bioinformatics/bts209
Erken, H., Yariz, K.O., Duman, D., Kaya, C.T., Sayin, T., Heper, A.O., Tekin, M., 2011. Cardiomyopathy with alopecia and palmoplantar keratoderma (CAPK) is caused by a *JUP* mutation. Br. J. Dermatol. 165, 917–921. https://doi.org/10.1111/j.1365-2133.2011.10455.x
Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., Ronneberger, O., Bodenstein, S., Zielinski, M., Bridgland, A., Potapenko, A., Cowie, A., Tunyasuvunakool, K., Jain, R., Clancy, E., Kohli, P., Jumper, J., Hassabis, D., 2021. Protein complex prediction with AlphaFold-Multimer (preprint). Bioinformatics. https://doi.org/10.1101/2021.10.04.463034
Gabler, F., Nam, S., Till, S., Mirdita, M., Steinegger, M., Söding, J., Lupas, A.N., Alva, V., 2020. Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. Curr. Protoc. Bioinforma. 72. https://doi.org/10.1002/cpbi.108
Garrod, D., Chidgey, M., 2008. Desmosome structure, composition and function. Biochim. Biophys. Acta BBA - Biomembr. 1778, 572–587.

https://doi.org/10.1016/j.bbamem.2007.07.014
Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., Lopez, R., 2010. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res. 38, W695–W699. https://doi.org/10.1093/nar/gkq313
Green, K.J., Roth-Carter, Q., Niessen, C.M., Nichols, S.A., 2020. Tracing the Evolutionary Origin of Desmosomes. Curr. Biol. 30, R535–R543. https://doi.org/10.1016/j.cub.2020.03.047
Green, K.J., Simpson, C.L., 2007. Desmosomes: New Perspectives on a Classic. J. Invest. Dermatol. 127, 2499–2515. https://doi.org/10.1038/sj.jid.5701015
Hatzfeld, M., Haffner, C., Schulze, K., Vinzens, U., 2000. The Function of Plakophilin 1 in Desmosome Assembly and Actin Filament Organization. J. Cell Biol. 149, 209–222. https://doi.org/10.1083/jcb.149.1.209
Hunter, J.D., 2007. Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng. 9, 90–95. https://doi.org/10.1109/MCSE.2007.55
Keller, D., Stepowski, D., Balmer, C., Simon, F., Guenthard, J., Bauer, F., Itin, P., David, N., Drouin-Garraud, V., Fressart, V., 2012. De novo heterozygous desmoplakin mutations leading to Naxos-Carvajal disease. Swiss Med. Wkly. https://doi.org/10.4414/smw.2012.13670
Kim, S.J., Fernandez-Martinez, J., Nudelman, I., Shi, Y., Zhang, W., Raveh, B., Herricks, T., Slaughter, B.D., Hogan, J.A., Upla, P., Chemmama, I.E., Pellarin, R., Echeverria, I., Shivaraju, M., Chaudhury, A.S., Wang, J., Williams, R., Unruh, J.R., Greenberg, C.H., Jacobs, E.Y., Yu, Z., de la Cruz, M.J., Mironska, R., Stokes, D.L., Aitchison, J.D., Jarrold, M.F., Gerton, J.L., Ludtke, S.J., Akey, C.W., Chait, B.T., Sali, A., Rout, M.P., 2018. Integrative structure and functional anatomy of a nuclear pore complex. Nature 555, 475–482. https://doi.org/10.1038/nature26003
Kohn, J.E., Millett, I.S., Jacob, J., Zagrovic, B., Dillon, T.M., Cingel, N., Dothager, R.S., Seifert, S., Thiyagarajan, P., Sosnick, T.R., Hasan, M.Z., Pande, V.S., Ruczinski, I., Doniach, S., Plaxco, K.W., 2004. Random-coil behavior and the dimensions of chemically unfolded proteins. Proc. Natl. Acad. Sci. 101, 12491–12496. https://doi.org/10.1073/pnas.0403643101
Kowalczyk, A.P., Bornslaeger, E.A., Borgwardt, J.E., Palka, H.L., Dhaliwal, A.S., Corcoran, C.M., Denning, M.F., Green, K.J., 1997. The Amino-terminal Domain of Desmoplakin Binds to Plakoglobin and Clusters Desmosomal Cadherin–Plakoglobin Complexes. J. Cell Biol. 139, 773–784. https://doi.org/10.1083/jcb.139.3.773
Kowalczyk, A.P., Green, K.J., 2013. Structure, Function, and Regulation of Desmosomes, in: Progress in Molecular Biology and Translational Science. Elsevier, pp. 95–118. https://doi.org/10.1016/B978-0-12-394311-8.00005-4
Kowalczyk, A.P., Hatzfeld, M., Bornslaeger, E.A., Kopp, D.S., Borgwardt, J.E., Corcoran, C.M., Settler, A., Green, K.J., 1999. The Head Domain of Plakophilin-1 Binds to Desmoplakin and Enhances Its Recruitment to Desmosomes. J. Biol. Chem. 274, 18145–18148. https://doi.org/10.1074/jbc.274.26.18145
Lasker, K., Förster, F., Bohn, S., Walzthoeni, T., Villa, E., Unverdorben, P., Beck, F., Aebersold, R., Sali, A., Baumeister, W., 2012. Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. Proc. Natl. Acad. Sci. 109, 1380–1387. https://doi.org/10.1073/pnas.1120559109
Marino, T.C., Maranda, B., Leblanc, J., Pratte, A., Barabas, M., Dupéré, A., Lévesque, S., 2017. Novel founder mutation in French-Canadian families with Naxos disease: Letter to the Editor. Clin. Genet. 92, 451–453. https://doi.org/10.1111/cge.12971
McInnes, L., Healy, J., Astels, S., 2017. hdbscan: Hierarchical density based clustering. J. Open Source Softw. 2, 205. https://doi.org/10.21105/joss.00205
McKoy, G., Protonotarios, N., Crosby, A., Tsatsopoulou, A., Anastasakis, A., Coonar, A., Norman, M., Baboonian, C., Jeffery, S., McKenna, W.J., 2000. Identification of a deletion in plakoglobin in arrhythmogenic right ventricular cardiomyopathy with palmoplantar keratoderma and woolly hair (Naxos disease). The Lancet 355, 2119–2124. https://doi.org/10.1016/S0140-6736(00)02379-5

Nilles, L.A., Parry, D.A.D., Powers, E.E., Angst, B.D., Wagner, R.M., Green, K.J., 1991. Structural analysis and expression of human desmoglein: a cadherin-like component of the desmosome. J. Cell Sci. 99, 809–821. https://doi.org/10.1242/jcs.99.4.809
North, A.J., Bardsley, W.G., Hyam, J., Bornslaeger, E.A., Cordingley, H.C., Trinnaman, B., Hatzfeld, M., Green, K.J., Magee, A.I., Garrod, D.R., 1999. Molecular map of the desmosomal plaque. J. Cell Sci. 112, 4325–4336. https://doi.org/10.1242/jcs.112.23.4325
Online Mendelian Inheritance in Man, 2023.
Palka, H.L., Green, K.J., 1997. Roles of plakoglobin end domains in desmosome assembly. J. Cell Sci. 110, 2359–2371. https://doi.org/10.1242/jcs.110.19.2359
Pasani, S., Viswanath, S., 2021. A Framework for Stochastic Optimization of Parameters for Integrative Modeling of Macromolecular Assemblies. Life 11, 1183. https://doi.org/10.3390/life11111183
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera: A visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612. https://doi.org/10.1002/jcc.20084
Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., Ferrin, T.E., 2021. UCSF ChimeraX : Structure visualization for researchers, educators, and developers. Protein Sci. 30, 70–82. https://doi.org/10.1002/pro.3943
Pigors, M., Schwieger-Briel, A., Cosgarea, R., Diaconeasa, A., Bruckner-Tuderman, L., Fleck, T., Has, C., 2015. Desmoplakin Mutations with Palmoplantar Keratoderma, Woolly Hair and Cardiomyopathy. Acta Derm. Venereol. 95, 337–340. https://doi.org/10.2340/00015555-1974
Pintilie, G.D., Zhang, J., Goddard, T.D., Chiu, W., Gossard, D.C., 2010. Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. J. Struct. Biol. 170, 427–438. https://doi.org/10.1016/j.jsb.2010.03.007
Robinson, P.J., Trnka, M.J., Pellarin, R., Greenberg, C.H., Bushnell, D.A., Davis, R., Burlingame, A.L., Sali, A., Kornberg, R.D., 2015. Molecular architecture of the yeast Mediator complex. eLife 4, e08719. https://doi.org/10.7554/eLife.08719
Rout, M.P., Sali, A., 2019. Principles for Integrative Structural Biology Studies. Cell 177, 1384–1403. https://doi.org/10.1016/j.cell.2019.05.016
Russel, D., Lasker, K., Webb, B., Velázquez-Muriel, J., Tjioe, E., Schneidman-Duhovny, D., Peterson, B., Sali, A., 2012. Putting the Pieces Together: Integrative Modeling Platform Software for Structure Determination of Macromolecular Assemblies. PLoS Biol. 10, e1001244. https://doi.org/10.1371/journal.pbio.1001244
Šali, A., Blundell, T.L., 1993. Comparative Protein Modelling by Satisfaction of Spatial Restraints. J. Mol. Biol. 234, 779–815. https://doi.org/10.1006/jmbi.1993.1626
Saltzberg, D.J., Viswanath, S., Echeverria, I., Chemmama, I.E., Webb, B., Sali, A., 2021. Using Integrative Modeling Platform to compute, validate, and archive a model of a protein complex structure. Protein Sci. 30, 250–261. https://doi.org/10.1002/pro.3995
Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539. https://doi.org/10.1038/msb.2011.75
Smith, E.A., Fuchs, E., 1998. Defining the Interactions Between Intermediate Filaments and Desmosomes. J. Cell Biol. 141, 1229–1241. https://doi.org/10.1083/jcb.141.5.1229
Sobolik-Delmaire, T., Katafiasz, D., Wahl, J.K., 2006. Carboxyl Terminus of Plakophilin-1 Recruits It to Plasma Membrane, whereas Amino Terminus Recruits Desmoplakin and Promotes Desmosome Assembly. J. Biol. Chem. 281, 16962–16970. https://doi.org/10.1074/jbc.M600570200
Tange, Ole, 2020. GNU Parallel 20200722 ('Privacy Shield'). https://doi.org/10.5281/ZENODO.3956817
Tate, J.G., Bamford, S., Jubb, H.C., Sondka, Z., Beare, D.M., Bindal, N., Boutselakis, H., Cole, C.G., Creatore, C., Dawson, E., Fish, P., Harsha, B., Hathaway, C., Jupe, S.C.,

Kok, C.Y., Noble, K., Ponting, L., Ramshaw, C.C., Rye, C.E., Speedy, H.E., Stefancsik, R., Thompson, S.L., Wang, S., Ward, S., Campbell, P.J., Forbes, S.A., 2019. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res. 47, D941–D947. https://doi.org/10.1093/nar/gky1015

Teraoka, I., 2002. Polymer solutions: an introduction to physical properties. 2 John Wiley & Sons, Inc.

The UniProt Consortium, Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Ahmad, S., Alpi, E., Bowler-Barnett, E.H., Britto, R., Bye-A-Jee, H., Cukura, A., Denny, P., Dogan, T., Ebenezer, T., Fan, J., Garmiri, P., Da Costa Gonzales, L.J., Hatton-Ellis, E., Hussein, A., Ignatchenko, A., Insana, G., Ishtiaq, R., Joshi, V., Jyothi, D., Kandasaamy, S., Lock, A., Luciani, A., Lugaric, M., Luo, J., Lussi, Y., MacDougall, A., Madeira, F., Mahmoudy, M., Mishra, A., Moulang, K., Nightingale, A., Pundir, S., Qi, G., Raj, S., Raposo, P., Rice, D.L., Saidi, R., Santos, R., Speretta, E., Stephenson, J., Totoo, P., Turner, E., Tyagi, N., Vasudev, P., Warner, K., Watkins, X., Zaru, R., Zellner, H., Bridge, A.J., Aimo, L., Argoud-Puy, G., Auchincloss, A.H., Axelsen, K.B., Bansal, P., Baratin, D., Batista Neto, T.M., Blatter, M.-C., Bolleman, J.T., Boutet, E., Breuza, L., Gil, B.C., Casals-Casas, C., Echioukh, K.C., Coudert, E., Cuche, B., De Castro, E., Estreicher, A., Famiglietti, M.L., Feuermann, M., Gasteiger, E., Gaudet, P., Gehant, S., Gerritsen, V., Gos, A., Gruaz, N., Hulo, C., Hyka-Nouspikel, N., Jungo, F., Kerhornou, A., Le Mercier, P., Lieberherr, D., Masson, P., Morgat, A., Muthukrishnan, V., Paesano, S., Pedruzzi, I., Pilbout, S., Pourcel, L., Poux, S., Pozzato, M., Pruess, M., Redaschi, N., Rivoire, C., Sigrist, C.J.A., Sonesson, K., Sundaram, S., Wu, C.H., Arighi, C.N., Arminski, L., Chen, C., Chen, Y., Huang, H., Laiho, K., McGarvey, P., Natale, D.A., Ross, K., Vinayaka, C.R., Wang, Q., Wang, Y., Zhang, J., 2023. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 51, D523–D531. https://doi.org/10.1093/nar/gkac1052

Troyanovsky, S.M., Troyanovsky, R.B., Eshkind, L.G., Krutovskikh, V.A., Leube, R.E., Franke, W.W., 1994a. Identification of the plakoglobin-binding domain in desmoglein and its role in plaque assembly and intermediate filament anchorage. J. Cell Biol. 127, 151–160. https://doi.org/10.1083/jcb.127.1.151

Troyanovsky, S.M., Troyanovsky, R.B., Eshkind, L.G., Leube, R.E., Franke, W.W., 1994b. Identification of amino acid sequence motifs in desmocollin, a desmosomal glycoprotein, that are required for plakoglobin binding and plaque formation. Proc. Natl. Acad. Sci. 91, 10790–10794. https://doi.org/10.1073/pnas.91.23.10790

Tucker, D.K., Stahley, S.N., Kowalczyk, A.P., 2014. Plakophilin-1 Protects Keratinocytes from Pemphigus Vulgaris IgG by Forming Calcium-Independent Desmosomes. J. Invest. Dermatol. 134, 1033–1043. https://doi.org/10.1038/jid.2013.401

Uversky, V.N., 2013. The most important thing is the tail: Multitudinous functionalities of intrinsically disordered protein termini. FEBS Lett. 587, 1891–1901. https://doi.org/10.1016/j.febslet.2013.04.042

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Van Der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., Van Mulbregt, P., SciPy 1.0 Contributors, Vijaykumar, A., Bardelli, A.P., Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C.N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D.A., Hagen, D.R., Pasechnik, D.V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G.A., Ingold, G.-L., Allen, G.E., Lee, G.R., Audren, H., Probst, I., Dietrich, J.P., Silterra, J., Webber, J.T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J.L., De Miranda Cardoso, J.V., Reimer, J., Harrington, J., Rodríguez, J.L.C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N.J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P.A., Lee, P., McGibbon, R.T.,

Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., Oshima, T., Pingel, T.J., Robitaille, T.P., Spura, T., Jones, T.R., Cera, T., Leslie, T., Zito, T., Krauss, T., Upadhyay, U., Halchenko, Y.O., Vázquez-Baeza, Y., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Viswanath, S., Bonomi, M., Kim, S.J., Klenchin, V.A., Taylor, K.C., Yabut, K.C., Umbreit, N.T., Van Epps, H.A., Meehl, J., Jones, M.H., Russel, D., Velazquez-Muriel, J.A., Winey, M., Rayment, I., Davis, T.N., Sali, A., Muller, E.G., 2017a. The molecular architecture of the yeast spindle pole body core determined by Bayesian integrative modeling. Mol. Biol. Cell 28, 3298–3314. https://doi.org/10.1091/mbc.e17-06-0397

Viswanath, S., Chemmama, I.E., Cimermancic, P., Sali, A., 2017b. Assessing Exhaustiveness of Stochastic Sampling for Integrative Modeling of Macromolecular Structures. Biophys. J. 113, 2344–2353. https://doi.org/10.1016/j.bpj.2017.10.005

Wahl, J.K., Sacco, P.A., McGranahan-Sadler, T.M., Sauppe, L.M., Wheelock, M.J., Johnson, K.R., 1996. Plakoglobin domains that define its association with the desmosomal cadherins and the classical cadherins: identification of unique and shared domains. J. Cell Sci. 109, 1143–1154. https://doi.org/10.1242/jcs.109.5.1143

Webb, B., Viswanath, S., Bonomi, M., Pellarin, R., Greenberg, C.H., Saltzberg, D., Sali, A., 2018. Integrative structure modeling with the Integrative Modeling Platform: Integrative Structure Modeling with IMP. Protein Sci. 27, 245–258. https://doi.org/10.1002/pro.3311

Whittock, N.V., Wan, H., Eady, R.A.J., Morley, S.M., Garzon, M.C., Kristal, L., Hyde, P., Irwin McLean, W.H., Pulkkinen, L., Uitto, J., Christiano, A.M., McGrath, J.A., 2002. Compound Heterozygosity for Non-Sense and Mis-Sense Mutations in Desmoplakin Underlies Skin Fragility/Woolly Hair Syndrome. J. Invest. Dermatol. 118, 232–238. https://doi.org/10.1046/j.0022-202x.2001.01664.x