

# Desmosomal protein structure and function and the impact of disease-causing mutations

Mohammed, Fiyaz; Chidgey, Martyn

DOI:

[10.1016/j.jsb.2021.107749](https://doi.org/10.1016/j.jsb.2021.107749)

License:

Creative Commons: Attribution (CC BY)

*Document Version*

Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*

Mohammed, F & Chidgey, M 2021, 'Desmosomal protein structure and function and the impact of disease-causing mutations', *Journal of Structural Biology*, vol. 213, no. 3, 107749.  
<https://doi.org/10.1016/j.jsb.2021.107749>

[Link to publication on Research at Birmingham portal](#)

## General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

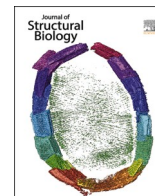
Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.



## Graphical Structural Biology Reviews

## Desmosomal protein structure and function and the impact of disease-causing mutations

Fiyaz Mohammed<sup>a,\*</sup>, Martyn Chidgey<sup>b</sup><sup>a</sup> Institute for Immunology and Immunotherapy, University of Birmingham, Birmingham B15 2TT, UK<sup>b</sup> Institute of Clinical Sciences, University of Birmingham, Birmingham B15 2TT, UK

## ARTICLE INFO

## Keywords:

Desmosome  
Cell–cell adhesion  
Desmosomal proteins  
Disease-related missense mutations  
Arrhythmogenic right ventricular cardiomyopathy

## ABSTRACT

In this graphical review we focus on the structural characteristics of desmosomal proteins, their interactions with each other and with the intermediate filament cytoskeleton. The wealth of structural information that is now available allows predictions to be made about the pathogenic effect of disease-causing mutations. We have selected representative examples of missense mutations that are buried, semi-buried or surface exposed, and demonstrate how such variants could affect the structural fold of desmosomal proteins that are expressed in the heart. We explain how such alterations could compromise desmosomal adhesion, resulting in life threatening diseases including arrhythmogenic right ventricular cardiomyopathy.

## 1. Introduction

Desmosomes are intercellular junctions that anchor cell cytoskeletal intermediate filaments (IFs) and are responsible for maintaining tissue integrity, particularly in tissues that undergo mechanical stress such as the myocardium and epidermis (Fig. 1A–C). They associate with desmin IFs in cardiomyocytes, keratin IFs in epithelial cells and vimentin IFs in follicular dendritic cells. When desmosomal adhesion is compromised, as in some genetically inherited and autoimmune diseases, tissues lose their structural cohesion. For strong intercellular adhesion five core desmosomal proteins, namely a desmocollin (DSC), a desmoglein (DSG), plakoglobin, a plakophilin (PKP) and desmoplakin (Fig. 1D), are required. Mutations in genes encoding these proteins are often associated with a translational frameshift, resulting in unstable truncated proteins and haploinsufficiency. Other mutations induce single residue substitutions in the translated protein, raising the question of whether such alterations invoke neutral or disease-causing effects on protein function. This review examines the structural effects of disease-related missense mutations in desmosomal proteins, and considers their impact on protein structure, stability and intercellular junction assembly. We focus on those mutations that affect desmosomal genes that are expressed in the heart (i.e. *DSC2*, *DSG2*, *JUP*, *PKP2* and *DSP*). For mutations in desmosomal genes (such as *DSG1*, *DSG4*, *DSC3* and *PKP1*) that lead to skin and hair related disorders, we refer readers to a recent review (Lee and McGrath, 2021).

## 2. Desmocollins and desmogleins

DSCs and DSGs are members of the cadherin superfamily and mediate calcium-dependent adhesion at desmosomal junctions. The human genome contains three DSC (*DSC1–DSC3*) and four DSG (*DSG1–DSG4*) genes. The DSCs/DSGs are single-pass membrane spanning proteins (Fig. 2A, B) which exhibit complex tissue- and differentiation-specific expression profiles (Garrod and Chidgey, 2008). Their ectodomains are comprised of five tandem extracellular cadherin (EC) domains with each composed of seven  $\beta$ -strands arranged as two opposed  $\beta$ -sheets (Fig. 2C, D). Both homophilic (i.e. DSC–DSC and DSG–DSG) and heterophilic (i.e. DSC–DSG) interactions have been detected (Spindler et al., 2009; Lowndes et al., 2014; Harrison et al., 2016; Shafraz et al., 2018), and the nature of desmosomal cadherin interactions remains a matter of some dispute (Vielmuth et al., 2018). Regardless the underlying mechanism of dimerization is likely to involve strand exchange between EC1 domains of opposed molecules (Harrison et al., 2016). An atomic model of the desmosome architecture has recently been generated using cryo-electron tomography and molecular dynamics. The model demonstrates that desmosomal cadherins associate at the midline generating a sieve-like pattern with the cadherins engaging in *cis* and *trans* interactions (Sikora et al., 2020).

Desmosomal cadherin cytoplasmic regions can be subdivided into a number of distinct domains based on their sequence (Fig. 2A, B). The DSC cytoplasmic tail consists of an intracellular anchor and catenin

\* Corresponding author.

E-mail addresses: [F.Mohammed@bham.ac.uk](mailto:F.Mohammed@bham.ac.uk) (F. Mohammed), [M.A.Chidgey@bham.ac.uk](mailto:M.A.Chidgey@bham.ac.uk) (M. Chidgey).<https://doi.org/10.1016/j.jsb.2021.107749>

Received 4 January 2021; Received in revised form 19 May 2021; Accepted 20 May 2021

Available online 24 May 2021

1047-8477/© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

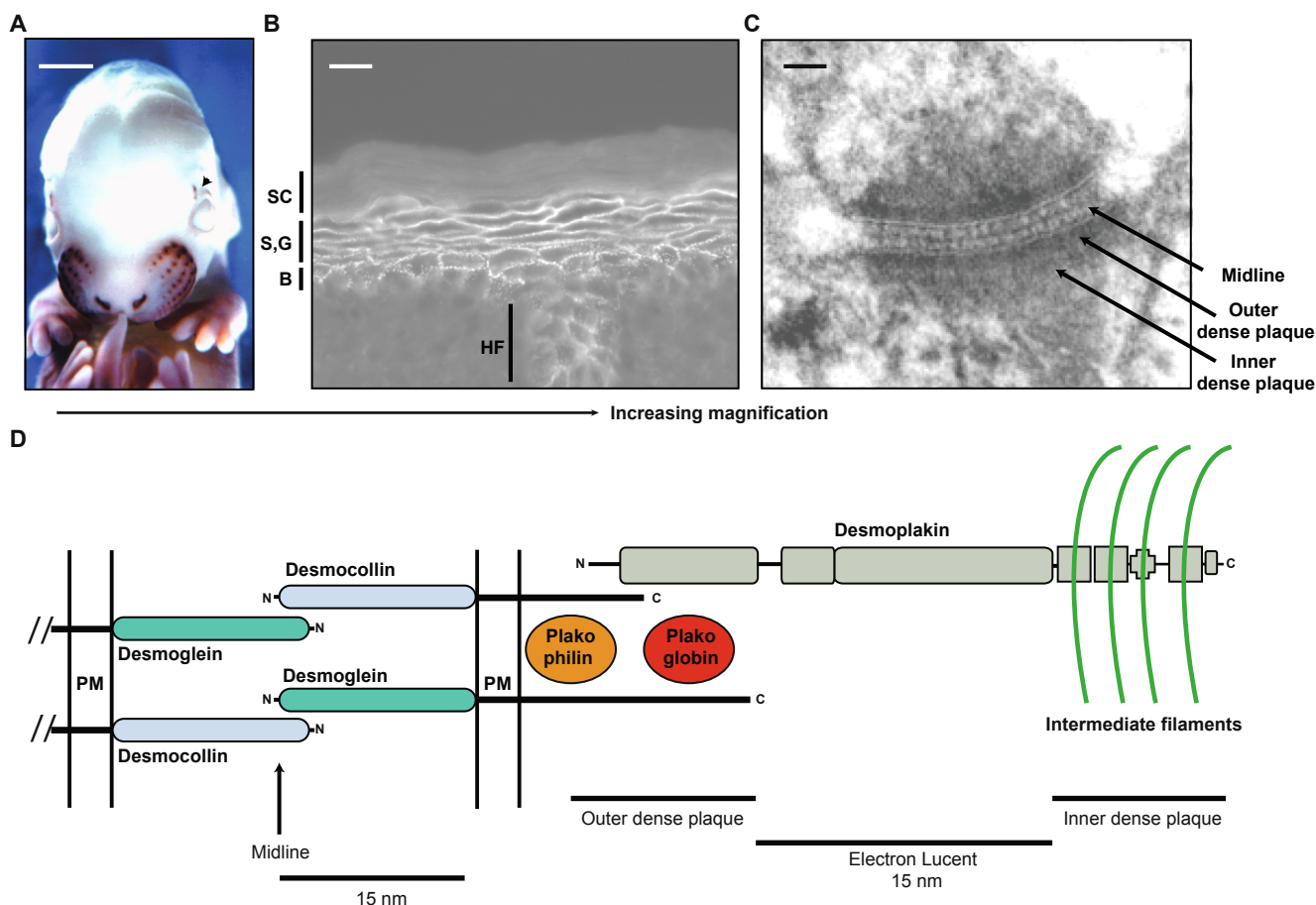
binding domain (CBD). All DSC genes undergo alternative splicing, resulting in longer 'a', and shorter 'b' forms with a truncated CBD that abolishes binding of most ligands. The E-cadherin cytoplasmic domain is intrinsically disordered (Huber et al., 2001) and its CBD interacts with  $\beta$ -catenin and plakoglobin at multiple points, wrapping around their *arm* repeats (Choi et al., 2009). It is envisaged that the DSC/DSG-CBDs engage with plakoglobin in a similar manner given their predicted structural similarity to the E-cadherin-CBD (Fig. 2E, F). DSG cytoplasmic regions are substantially longer than those of DSCs encompassing additional intrinsically disordered segments (Fig. 2B). The function of these extreme C-terminal DSG regions is unclear although they mediate weak binding to other desmosomal proteins (Kami et al., 2009) and may enhance adhesion by inhibiting internalisation (Chen et al., 2012).

The genes encoding desmosomal cadherins expressed in the heart, namely *DSC2* and *DSG2*, are targeted by numerous mutations that result in cardiac-related disorders (Fig. 2G, H). Using structural information available for *DSC2/DSG2-EC* domains we have identified multiple mechanisms by which disease-causing mutations could adversely affect desmosomal cadherin function (Fig. 2I–K). For example, missense mutation I345T in *DSC2-EC2* contributes to arrhythmogenic right ventricular cardiomyopathy (ARVC). The *DSC2-EC2* structure shows that I345 maps to the hydrophobic core region mediating multiple non-polar interactions, supporting the *EC2* fold. The *DSC2-EC2* model encompassing the I345T mutation reveals that these interactions would be abolished,

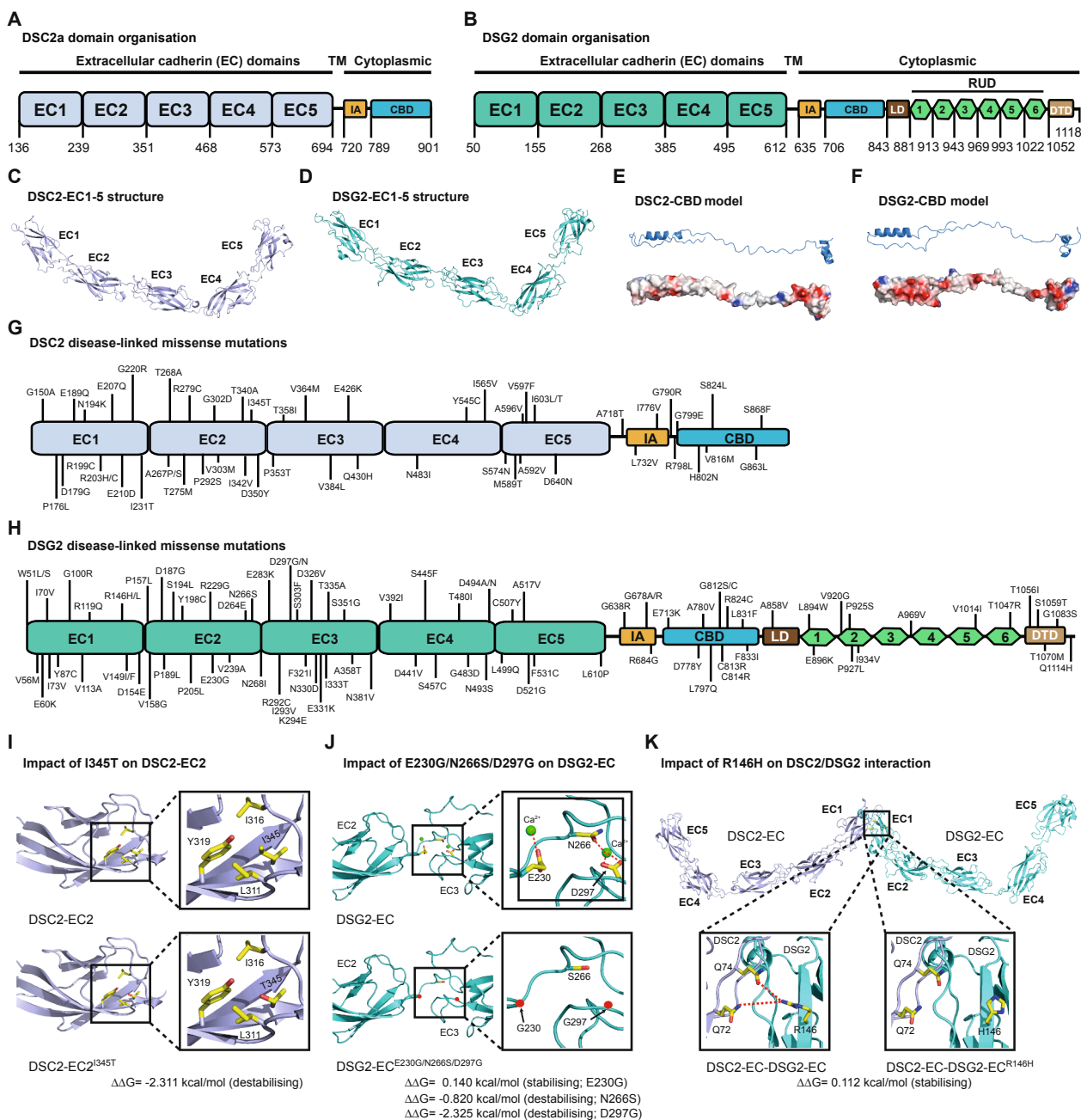
creating structural vulnerabilities within the *EC2* domain. Consistent with this, I345T was predicted to destabilise the *DSC2-EC* fold using DynaMut, a web server that analyses the effects of amino acid substitution on protein stability by calculating the value of  $\Delta\Delta G$  (a metric for predicting how point mutations affect protein stability) (Rodrigues et al., 2018). Other ARVC-causing mutations such as E230G/N266S/D297G map to the *EC2-EC3* interface of *DSG2* and are predicted to compromise calcium binding and interfere with the extended conformations adopted by *EC* domains that are essential for intercellular interactions. Interestingly, only two of these mutations were classified as destabilising using DynaMut highlighting the challenges associated with accurately predicting the effect of point mutations on protein stability. Finally, although the R146H mutation in *DSG2-EC1* was classified as stabilising, such a change may compromise *DSC2/DSG2* heterodimeric interactions leading to impaired desmosome assembly.

### 3. Plakoglobin and plakophilins

The armadillo proteins plakoglobin and PKP serve as linkers in the desmosome-intermediate filament complex (Garrod and Chidgey, 2008) (Fig. 1D). They are characterised by a variable number of imperfect 42-residue *arm* repeats. Plakoglobin is found in both desmosomes and adherens junctions and in the latter engages with the E-cadherin-CBD. Each PKP (PKP1-PKP3) exists in two different splice variants and



**Fig. 1.** Visualisation of desmosome ultrastructure. A) Whole embryo in situ hybridisation showing the location of desmosomes incorporating desmocollin 3 in nasal epidermis, vibrissa hair follicles, whisker pad epithelium and the tactile hair follicles above the eye (arrowhead) at mouse embryonic day of development 13.5. Bar, 1 mm. B) Immunohistochemistry showing desmoplakin containing desmosomes at cell–cell borders in mouse epidermis. The spot-like appearance of the staining in the basal layer keratinocytes shows the location of individual desmosomes. Desmosomes are not present on the basal surface of basal layer keratinocytes. HF, hair follicle; B, basal layer; S,G, spinous, granular layers; SC, stratum corneum. Bar, 25  $\mu$ m. C) Electron microscopy showing desmosome ultrastructure. Bar, 0.1  $\mu$ m. D) Architecture of the desmosome showing the approximate location of desmosomal proteins. Heterophilic interactions are shown between the desmosomal cadherins in the intercellular space, and for simplicity desmoplakin is depicted as a monomer. Structured domains are coloured whereas intrinsically disordered domains are depicted as thin black lines. The intrinsically disordered head and tail domains of plakoglobin and plakophilin are omitted for clarity. PM, plasma membrane.

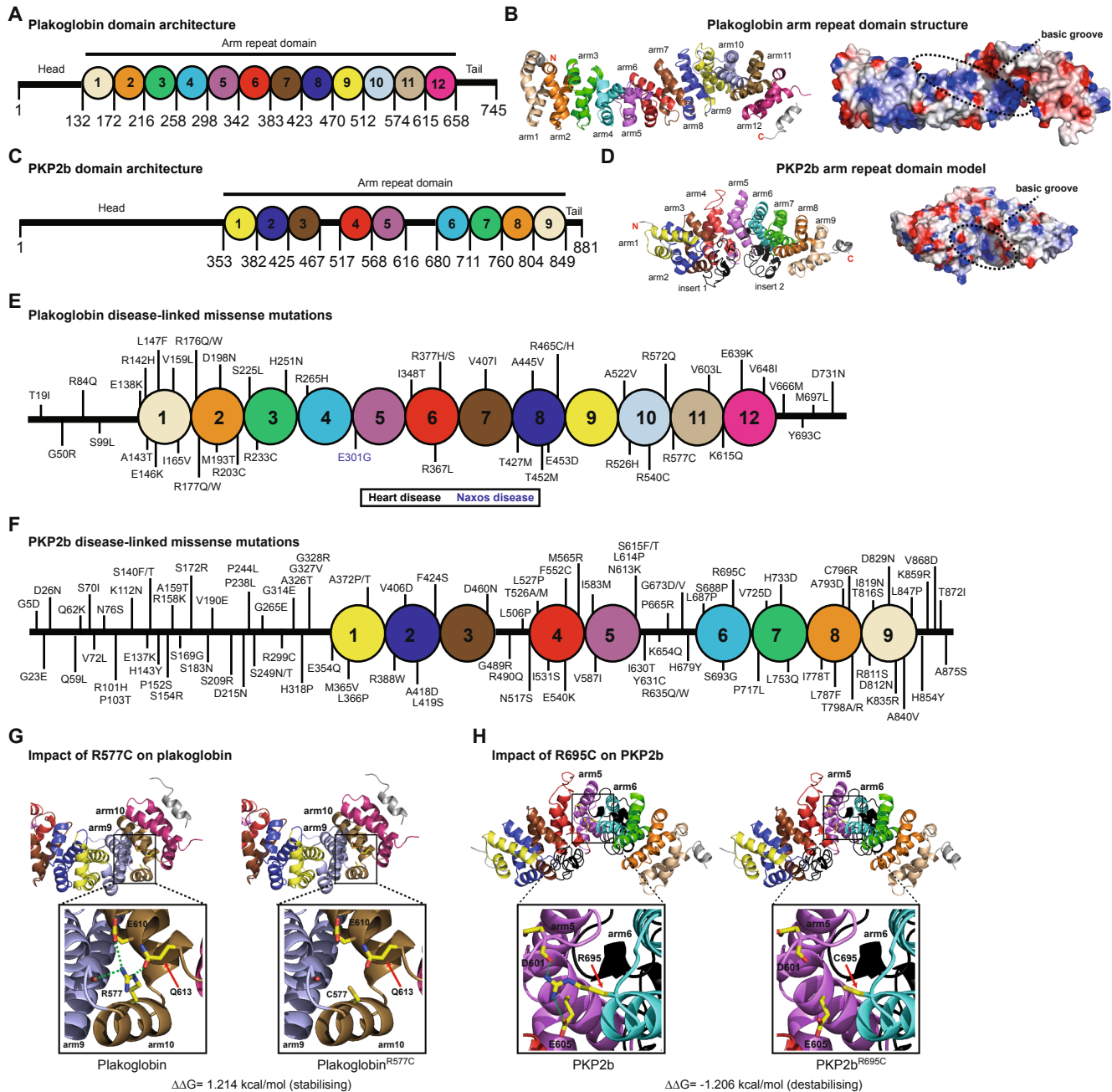


**Fig. 2.** Desmocollin 2a (DSC2a) and desmoglein 2 (DSG2) structures and mapping of disease-related missense mutations. A) Domain architecture of DSC2a showing the five extracellular cadherin (EC) domains, a membrane spanning domain and an intrinsically disordered cytoplasmic domain consisting of an intracellular anchor (IA) and a catenin binding domain (CBD) (also known as the intracellular cadherin-like sequence or ICS). Leader and pro-peptides, which are cleaved during maturation of the protein, are not shown. B) Domain architecture of DSG2 is similar to DSC2 but with the addition of an intrinsically disordered C-terminal region consisting of a linker domain (LD), a repeat unit domain (RUD) and a desmoglein terminal domain (DTD). Note that the linker domain is sometimes referred to as the intracellular proline-rich linker (IPL), a monomer as the domain is not proline rich in DSGs 2, 3 and 4. C) Crystal structure of DSC2 EC domains 1–5 (PDB: 5J5J and 5ERP). D) Crystal structure of the DSG2 EC domains 1–5 (PDB: 5ERG). E) I-TASSER derived model of the DSC2-CBD showing an extended conformation, and electrostatic map revealing its negatively charged surface. The model is derived from the E-cadherin CBD in its bound form with plakoglobin. F) I-TASSER derived model of the DSG2-CBD demonstrating an extended conformation, and electrostatic map revealing its negatively charged surface. The model is derived from the E-cadherin CBD in its bound state with plakoglobin. G) Mapping of DSC2 missense mutations (derived from the HGMD database) that result in heart diseases (such as ARVC and dilated cardiomyopathy) and sudden death. H) Mapping of DSG2 missense mutations that result in heart disease. I) Impact of disease-linked variant I345T on the crystal structure of DSC2-EC. Ribbon diagram of DSC2-EC demonstrating that the side chain of I345 mediates multiple non-polar contacts that stabilise the EC2 core region (top panel). The loss of these interactions with T345 are predicted to destabilise this region (bottom panel). J). Impact of disease-related variants E230G, N266S and D297G on the DSG2-EC crystal structure. Ribbon diagram of DSG2-EC showing that the side chains of E230, N266 and D297 co-ordinate calcium ions (top panel). The disease-associated variants can no longer bind calcium and are predicted to result in a less extended conformation (bottom panel). K) Impact of disease-linked variant R145H on DSC2-DSG2 model. Ribbon diagram of DSC2-DSG2 model demonstrating that the side chain of R145 in DSG2 may mediate multiple polar contacts with DSC2 (top panel). The disease-related variants can no longer mediate such contacts which may impair DSC2-DSG2 interactions (bottom panel). Boxes show close up views of the relevant interactions. The DynaMut server was used to calculate the predicted change in stability (in kcal/mol) for each variant.



exhibits complex tissue-specific expression profiles (Hofmann, 2020). As well as being structural components of desmosomes, plakoglobin and PKPs are involved in diverse signalling pathways and modulate cell behaviour (Aktary et al., 2017, Hofmann, 2020).

Plakoglobin consists of a central *arm* repeat domain flanked by intrinsically disordered N- and C-terminal tails (Fig. 3A). The central region consists of 12 *arm* repeats with all comprising of three  $\alpha$ -helices, H1, H2 and H3, except repeats 1 and 7 which lack H1 (Fig. 3B) (Choi



**Fig. 3.** Plakoglobin and plakophilin 2b (PKP2b) structures and mapping of disease-associated missense mutations. A) Domain architecture of plakoglobin showing intrinsically disordered head and tail domains flanking an *arm* repeat domain consisting of 12 *arm* repeats. B) Crystal structure of the plakoglobin *arm* repeat domain (PDB: 3IFQ). The electrostatic surface profile of the plakoglobin *arm* domain shows a positively charged groove (right panel). C) Domain architecture of PKP2b showing intrinsically disorder head and tail domains flanking an *arm* repeat domain consisting of 9 *arm* repeats. PKP2b is identical to PKP2a with the exception of a 44 amino acid insert between *arm* repeats 3 and 4. D) I-TASSER derived molecular model of the PKP2b *arm* repeat domain. The electrostatic surface profile of the PKP2b *arm* domain shows a positively charged patch in the superhelical groove. E) Mapping of plakoglobin disease-related missense mutations (derived from the HGMD database) that result in heart disease and sudden death, and Naxos disease which is characterised by cardiomyopathy, palmoplantar keratoderma and woolly hair. F) Mapping of PKP2b missense mutations that result in heart disease and sudden death. G) Examining the effect of disease-associated variant R577C on plakoglobin crystal structure. Ribbon diagram of the plakoglobin *arm* domain demonstrating that R577 mediates multiple polar interactions (green dashed lines) with residues emanating from *arm* repeats 9 and 10 (left panel). Introduction of C577 at this position would result in a loss of these intra- and inter-helical stabilising interactions (right panel). H) Probing the effect of disease-associated variant R695C on the PKP2b *arm* repeat domain model. Ribbon diagram of the PKP2b *arm* domain demonstrating that R695 mediates multiple polar interactions (green dashed lines) with residues protruding from *arm* repeat 5 (left panel). Introduction of C695 at this position would result in a loss of these stabilising interactions (right panel). Boxes show close up views of the relevant interactions. The DynaMut server was used to calculate the predicted change in stability (in kcal/mol) for each variant.

et al., 2009). The interface between plakoglobin and the E-cadherin-CBD spans the entire length of the plakoglobin *arm* domain and involves electrostatic interactions between the electropositive groove of plakoglobin and the negatively charged surface of the E-cadherin-CBD. Negatively charged patches are also evident in the DSC/DSG-CBD surfaces suggesting that electrostatic complementarity contributes to stable plakoglobin/desmomesal cadherin attachments. Overall, the negative surface charge on desmomesal cadherin CBDs is less than that on the E-cadherin-CBD which may explain why the latter binds plakoglobin with greater affinity (Choi et al., 2009).

PKPs consist of a central *arm* repeat domain flanked by intrinsically disordered regions (Fig. 3C). The crystal structure of the PKP1a *arm* repeat module consists of 9 *arm* repeats, with a flexible insert between repeats 5 and 6 (Choi and Weis, 2005). The PKP1 *arm* repeats pack together to form a distinct electropositive groove that is comparable to that of plakoglobin. Unsurprisingly, the PKP2b *arm* repeat domain adopts a similar fold to PKP1, albeit with an additional insert between *arm* repeats 3 and 4 (Fig. 3D). PKPs serve as a critical hub, mediating interactions with DSCs, DSGs, plakoglobin and desmoplakin. Although not formally proven the negatively charged DSC-CBD and DSG-CBD may interface with the PKP basic groove. In addition, the intrinsically disordered tails of plakoglobin and the PKPs could facilitate interactions with each other, desmomesal cadherins and desmoplakin.

The genes encoding armadillo proteins expressed in the heart, namely *JUP* and *PKP2*, are targeted by missense mutations that are distributed throughout the head, *arm* and tail domains (Fig. 3E, F). The clinical effects of these mutations can now be interpreted in light of the plakoglobin structure and a molecular model of PKP2b (Fig. 3G, H). For example, the ARVC-causing R577C mutation is located in *arm* repeat 10 of plakoglobin. The guanidinium moiety of R577 forms salt-bridge and hydrogen-bonding interactions with residues protruding from *arm* repeats 9 and 10. In the R577C variant, C577 can no longer contribute to such inter- and intra-helical stabilisation interactions. It is debatable whether the loss of these interactions is solely responsible for disease pathogenesis since the R577C variant is not classified as structurally deleterious based on DynaMut analysis. One possibility is that the thiol group of C577 may be susceptible to oxidation which could lead to the formation of physiologically irrelevant disulphide-linked PKP2b dimers. Similarly, the presence of the sudden death syndrome-related mutation R695C in PKP2b is predicted to introduce structural fragilities within the *arm* domain. In addition, there are numerous mutations that target positively charged *arm* domain residues, which may adversely affect DSC2-CBD and/or DSG2-CBD binding.

#### 4. Desmoplakin

Desmoplakin is a member of the plakin family of cytolinkers that functions as a linker between plakoglobin and the PKPs, and the IF cytoskeleton (Fig. 1D). It consists of N-terminal head, central rod and C-terminal tail domains (Fig. 4A–D). There are two major isoforms of desmoplakin, I and II, with the latter characterised by a shorter rod domain. The head domain comprises of a short intrinsically disordered region followed by a structured plakin domain. The plakin domain consists of six spectrin repeats (SRs) which are organised into long (SR3-6) and short (SR7-8) arms (Al-Jassar et al., 2013). Each SR consists of three helices A, B and C that form an antiparallel triple helical bundle. Helix C and helix A of the subsequent repeat are continuous as SRs are connected by an  $\alpha$ -helical linker. This arrangement yields an elongated rigid structure which is further rigidified by a Src homology-3 domain embedded in SR5 which packs extensively against SR4 (Choi and Weis, 2011). A flexible hinge between the two arms may allow the plakin domain to sweep for ligands during desmosome assembly, or permit extension of the desmoplakin molecule, preventing damage if the molecule is stretched (Al-Jassar et al., 2013). The desmoplakin head domain interacts with plakoglobin and the PKPs in the outer dense plaque. Although these interactions have not been precisely mapped, the

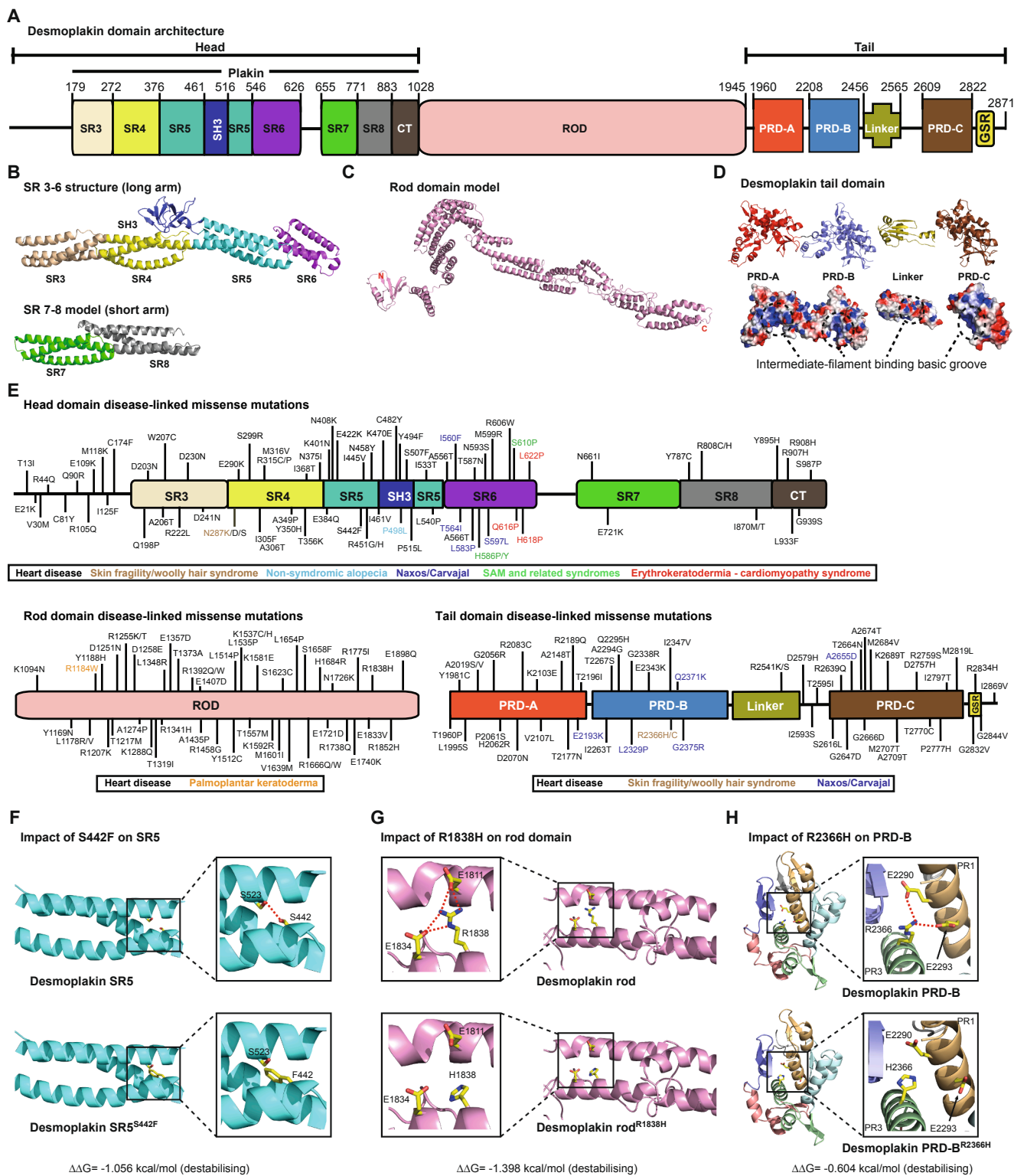
intrinsically disordered N-terminal extremity of desmoplakin is likely to interface with plakoglobin as ARVC mutations V30M and Q90R abolishes the interaction (Yang et al., 2006).

The rod domain is thought to be involved in desmoplakin homodimerization. A molecular model of the rod domain reveals a coiled-coil structure of approximately 20 nm in length (Fig. 4C), allowing it to transverse the electron lucent region between outer and inner dense plaque (Fig. 1D). Electron microscopy of purified desmoplakin I showed the rod domain was on average 130 nm long (O'Keefe et al., 1989), a figure that is likely to represent a fully unfolded version of the domain. The desmoplakin tail domain engages with the IF cytoskeleton in the inner dense plaque, and comprises of three plakin repeat domains (PRDs), a linker module and a glycine-serine-arginine-rich motif (Fig. 4D). The latter undergoes post-translational modifications which modulate desmoplakin-IF interactions (Albrecht et al., 2015). All three PRDs contain 4.5 copies of a 38-residue plakin repeat (PR) motif, which includes an 11-residue hairpin followed by an anti-parallel pair of  $\alpha$ -helices (Choi et al., 2002; Kang et al., 2016). Each PRD encompasses a basic groove that binds poly-acidic motifs on IF rods (Fogl et al., 2016). Given the relative orientations of the IF binding grooves in the tandem PRD-AB structure, it is likely that PRDs A and B engage adjacent IF rods, rather than binding the same rod (Kang et al., 2016; Mohammed et al., 2020). The desmoplakin linker contains two PR-like motifs with an electropositive groove, which also binds IF rods (Kang et al., 2016; Odintsova et al., 2020).

The *DSP* gene encoding desmoplakin is targeted by mutations that result in various heart and skin-related disorders. These disease-causing mutations are scattered throughout the head, rod and tail domains (Fig. 4E). The clinical effects of disease-associated mutations can now be interpreted in light of our increasing structural knowledge of desmoplakin modules (Fig. 4F–H). For example, the ARVC-causing variant S442F maps to SR5 in desmoplakin. In the wild-type protein S442 mediates hydrogen-bonding interaction with a neighbouring polar residue, whereas the bulky aromatic ring of F442 is predicted to clash with the surrounding helix and adversely affect the SR5 fold. Molecular dynamic simulations on the plakin long arm support the idea that the SH3 domain stabilises the module (Daday et al., 2017). Hence ARVC mutations, such as S442F, that map to the SR4-SR5-SH3 interface may compromise desmosome integrity and impair cellular responses to tension. Other disease-associated mutations identified in the rod and PRD domains such as R1838H and R2366H abolish core stabilising polar interactions which could compromise desmoplakin function.

#### 5. Conclusion

Mutations in desmosomal genes are being identified at an unprecedented rate as a result of advances in next-generation sequencing technology. Substantial progress has also been made in our understanding of the structure and function of cardiac-associated desmosomal proteins. This has allowed mapping of pathogen-causing missense mutations and accurate predictions of their impact on structures. Additionally, the rapid growth of computational stability predictors is likely to enhance identification of *bona fide* disease-linked structurally deleterious mutations. These approaches can now be extended to structurally map disease-causing variants in other desmosome-associated proteins (Lee and McGrath, 2021). Despite these advances, the molecular basis of how pathogenic mutations lead to disease remain challenging, particularly for non-structurally deleterious variants. Such non-synonymous mutations could modify the functional properties of the encoded protein by affecting ligand binding or trafficking. Future structural determination of desmosomal protein–ligand complexes will illuminate ligand docking modes and provide further molecular insights into how missense variants drive pathogenicity.



(caption on next page)

**Fig. 4.** Desmoplakin structure and mapping of disease-causing mutations. A) Desmoplakin consists of three major domains, a head, rod and tail domain. The head domain comprises of an intrinsically disordered region at the extreme N-terminus and a plakin domain consisting of long and short arms connected by a flexible linker. The long arm of the plakin domain consists of four spectrin repeats (SR 3–6) whereas the short arm encompasses two (SR 7–8). The SR nomenclature is based on plectin, which incorporates 9 SRs. CT domain sequence diverges from that of the SRs, perhaps reflecting its role in connecting to the rod domain. A non-canonical Src homology 3 (SH3) domain embedded in SR5 mediates interactions with SR4 that may contribute to the rigidity of the long arm. The central rod domain is responsible for dimerization. The C-terminal tail domain consists of three plakin repeat domains (PRDs), termed A, B and C, a linker module interspersed between PRDs B and C and a glycine-serine-arginine (GSR) motif. B) Crystal structure of the plakin long arm (PDB: 3R6N) and I-TASSER derived molecular model of the plakin short arm. C) I-TASSER derived molecular model of the central rod domain of desmoplakin I. D) Ribbon and molecular surface representation of PRDs A and B (PDB: 5DZZ), PRD-C (PDB: 1LM5) and the linker domain (I-TASSER derived model). The intermediate filament binding basic grooves for each domain are highlighted (circular dashed lines). E) Mapping of disease-related mutations (derived from the HGMD database) on the desmoplakin head, rod and tail domains. F) Examining the impact of the disease-associated variant S442F on the crystal structure of desmoplakin SR5. Ribbon diagram of the desmoplakin SR5 crystal structure showing the hydrogen-bonding interaction (red dashed line) between S442 and S523 (top panel). The larger F442 side chain is predicted to clash with S523 resulting in destabilisation of this region (bottom panel). G) Probing the effect of disease-associated variant R1838H on the desmoplakin rod domain model. Ribbon diagram of the desmoplakin rod domain model highlighting the electrostatic interactions (red dashed lines) between R1838 and E1811/E1834 (top panel). Loss of these interactions is predicted to destabilise this region (bottom panel). H) Analysing the impact of the disease-related variant R2366H on desmoplakin PRD-B. Ribbon diagram of the desmoplakin PRD-B structure showing the electrostatic interactions (red dashed lines) between R2366 and E2290/E2293 (top panel). Loss of these interactions is predicted to invoke structural penalties in this region (bottom panel). Boxes show close up views of the relevant interactions. The DynaMut server was used to calculate the predicted change in stability (in kcal/mol) for each variant.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work was supported by the Wellcome Trust (grant number 099266/Z/12/Z).

## References

- Aktary, Z., Alaea, M., Pasdar, M., 2017. Beyond cell-cell adhesion: Plakoglobin and the regulation of tumorigenesis and metastasis. *Oncotarget* 8, 32270–32291.
- Albrecht, L.V., Zhang, L., Shabanowitz, J., Purevjav, E., Towbin, J.A., Hunt, D.F., Green, K.J., 2015. GSK3- and PRMT1-dependent modifications of desmoplakin control desmoplakin-cytoskeleton dynamics. *J. Cell Biol.* 208, 597–612.
- Al-Jassar, C., Bernadomicron, P., Chidgey, M., Overduin, M., 2013. Hinged plakin domains provide specialized degrees of articulation in envoplakin, periplakin and desmoplakin. *PLoS ONE* 8, e69767.
- Chen, J., Nekrasova, O.E., Patel, D.M., Klessner, J.L., Godsel, L.M., Koetsier, J.L., Amargo, E.V., Desai, B.V., Green, K.J., 2012. The C-terminal unique region of desmoglein 2 inhibits its internalization via tail-tail interactions. *J. Cell Biol.* 199, 699–711.
- Choi, H.J., Weis, W.I., 2005. Structure of the armadillo repeat domain of plakophilin 1. *J. Mol. Biol.* 346, 367–376.
- Choi, H.J., Park-Snyder, S., Pascoe, L.T., Green, K.J., Weis, W.I., 2002. Structures of two intermediate filament-binding fragments of desmoplakin reveal a unique repeat motif structure. *Nat. Struct. Biol.* 9, 612–620.
- Choi, H.J., Gross, J.C., Pokutta, S., Weis, W.I., 2009. Interactions of plakoglobin and beta-catenin with desmosomal cadherins: basis of selective exclusion of alpha- and beta-catenin from desmosomes. *J. Biol. Chem.* 284, 31776–31788.
- 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.
- Daday, C., Kolsek, K., Grater, F., 2017. The mechano-sensing role of the unique SH3 insertion in plakin domains revealed by molecular dynamics simulations. *Sci. Rep.* 7, 11669.
- Fogl, C., Mohammed, F., Al-Jassar, C., Jeeves, M., Knowles, T.J., Rodriguez-Zamora, P., White, S.A., Odintsova, E., Overduin, M., Chidgey, M., 2016. Mechanism of intermediate filament recognition by plakin repeat domains revealed by envoplakin targeting of vimentin. *Nat. Commun.* 7, 10827.
- Garrod, D., Chidgey, M., 2008. Desmosome structure, composition and function. *BBA* 1778, 572–587.
- Harrison, O.J., Brasch, J., Lasso, G., Katsamba, P.S., Ahlsen, G., Honig, B., Shapiro, L., 2016. Structural basis of adhesive binding by desmocollins and desmogleins. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7160–7165.
- Hofmann, I., 2020. Plakophilins and their roles in diseased states. *Cell Tissue Res.* 379, 5–12.
- Huber, A.H., Stewart, D.B., Laurents, D.V., Nelson, W.J., Weis, W.I., 2001. The cadherin cytoplasmic domain is unstructured in the absence of beta-catenin. A possible mechanism for regulating cadherin turnover. *J. Biol. Chem.* 276, 12301–12309.
- Kami, K., Chidgey, M., Dafforn, T., Overduin, M., 2009. The desmoglein-specific cytoplasmic region is intrinsically disordered in solution and interacts with multiple desmosomal protein partners. *J. Mol. Biol.* 386, 531–543.
- Kang, H., Weiss, T.M., Bang, I., Weis, W.I., Choi, H.J., 2016. Structure of the intermediate filament-binding region of desmoplakin. *PLoS ONE* 11, e0147641.
- Lee, J.Y.W., McGrath, J.A., 2021. Mutations in genes encoding desmosomal proteins: spectrum of cutaneous and extracutaneous abnormalities. *Br. J. Dermatol.* 184, 596–605.
- Lowndes, M., Rakshit, S., Shafraz, O., Borghi, N., Harmon, R.M., Green, K.J., Sivasankar, S., Nelson, W.J., 2014. Different roles of cadherins in the assembly and structural integrity of the desmosome complex. *J. Cell Sci.* 127, 2339–2350.
- Mohammed, F., Trieber, C., Overduin, M., Chidgey, M., 2020. Molecular mechanism of intermediate filament recognition by plakin proteins. *Biochim. Biophys. Acta, Mol. Cell. Res.* 1867, 118801.
- Odintsova, E., Mohammed, F., Trieber, C., Rodriguez-Zamora, P., Al-Jassar, C., Huang, T. H., Fogl, C., Knowles, T., Sridhar, P., Kumar, J., Jeeves, M., Chidgey, M., Overduin, M., 2020. Binding of the periplakin linker requires vimentin acidic residues D176 and E187. *Commun. Biol.* 3, 83.
- O'Keefe, E.J., Erickson, H.P., Bennett, V., 1989. Desmoplakin I and desmoplakin II. Purification and characterization. *J. Biol. Chem.* 264, 8310–8318.
- Rodrigues, C.H., Pires, D.E., Ascher, D.B., 2018. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res.* 46, W350–W355.
- Shafraz, O., Rubsam, M., Stahley, S.N., Caldara, A.L., Kowalczyk, A.P., Niessen, C.M., Sivasankar, S., 2018. E-cadherin binds to desmoglein to facilitate desmosome assembly. *Elife* 7.
- Sikora, M., Ermel, U.H., Seybold, A., Kunz, M., Calloni, G., Reitz, J., Vabulas, R.M., Hummer, G., Frangakis, A.S., 2020. Desmosome architecture derived from molecular dynamics simulations and cryo-electron tomography. *Proc. Natl. Acad. Sci. U.S.A.* 117, 27132–27140.
- Spindler, V., Heupel, W.M., Efthymiadis, A., Schmidt, E., Eming, R., Rankl, C., Hinterdorfer, P., Muller, T., Drenckhahn, D., Waschke, J., 2009. Desmocollin 3-mediated binding is crucial for keratinocyte cohesion and is impaired in pemphigus. *J. Biol. Chem.* 284, 30556–30564.
- Vielmuth, F., Spindler, V., Waschke, J., 2018. Atomic force microscopy provides new mechanistic insights into the pathogenesis of pemphigus. *Front. Immunol.* 9, 485.
- Yang, Z., Bowles, N.E., Scherer, S.E., Taylor, M.D., Kearney, D.L., Ge, S., Nadvoretzkiy, V. V., Defreitas, G., Carabello, B., Brandon, L.I., Godsel, L.M., Green, K.J., Saffitz, J.E., Li, H., Danieli, G.A., Calkins, H., Marcus, F., Towbin, J.A., 2006. Desmosomal dysfunction due to mutations in desmoplakin causes arrhythmic right ventricular dysplasia/cardiomyopathy. *Circ. Res.* 99, 646–655.