



# The intracellular domain of BP180/collagen XVII is intrinsically disordered and partially folds in an anionic membrane lipid-mimicking environment

Jussi Tuusa<sup>1</sup> · M. Kristian Koski<sup>2</sup> · Salla Ruskamo<sup>2</sup> · Kaisa Tasanen<sup>1</sup>

Received: 10 January 2020 / Accepted: 16 March 2020 / Published online: 27 March 2020  
© The Author(s) 2020

## Abstract

The trimeric transmembrane collagen BP180, also known as collagen XVII, is an essential component of hemidesmosomes at the dermal–epidermal junction and connects the cytoplasmic keratin network to the extracellular basement membrane. Dysfunction of BP180 caused by mutations in patients with junctional epidermolysis bullosa or autoantibodies in those with bullous pemphigoid leads to severe skin blistering. The extracellular collagenous domain of BP180 participates in the protein's triple-helical folding, but the structure and functional importance of the intracellular domain (ICD) of BP180 are largely unknown. In the present study, we purified and characterized human BP180 ICD. When expressed in *Escherichia coli* as glutathione-S-transferase or 6×histidine tagged fusion protein, the BP180 ICD was found to exist as a monomer. Analysis of the secondary structure content by circular dichroism spectroscopy revealed that the domain is intrinsically disordered. This finding aligned with that of a bioinformatic analysis, which predicted a disordered structure. Interestingly, both anionic detergent micelles and lipid vesicles induced partial folding of the BP180 ICD, suggesting that in its natural environment, the domain's folding and unfolding may be regulated by interaction with the cell membrane or accompanying proteins. We hypothesize that the intrinsically disordered structure of the ICD of BP180 contributes to the mechanism that allows the remodeling of hemidesmosome assembly.

**Keywords** Bullous pemphigoid · Circular dichroism · Hemidesmosome · Induced folding

## Abbreviations

Aa	Amino acids	His	Histidine
BP	Bullous pemphigoid	ICD	Intracellular domain
CD	Circular dichroism	MALS	Multi-angle light scattering
GST	Glutathione-S-transferase	NC	Non-collagenous
		SEC	Size-exclusion chromatography

Handling Editor: S. Dai.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00726-020-02840-5>) contains supplementary material, which is available to authorized users.

✉ Jussi Tuusa  
jussi.tuusa@oulu.fi  
Kaisa Tasanen  
kaisa.tasanen@oulu.fi

<sup>1</sup> PEDEGO Research Unit, Department of Dermatology, Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland

<sup>2</sup> Biocenter Oulu and Faculty of Biochemistry and Molecular Medicine, University of Oulu, Oulu, Finland

## Introduction

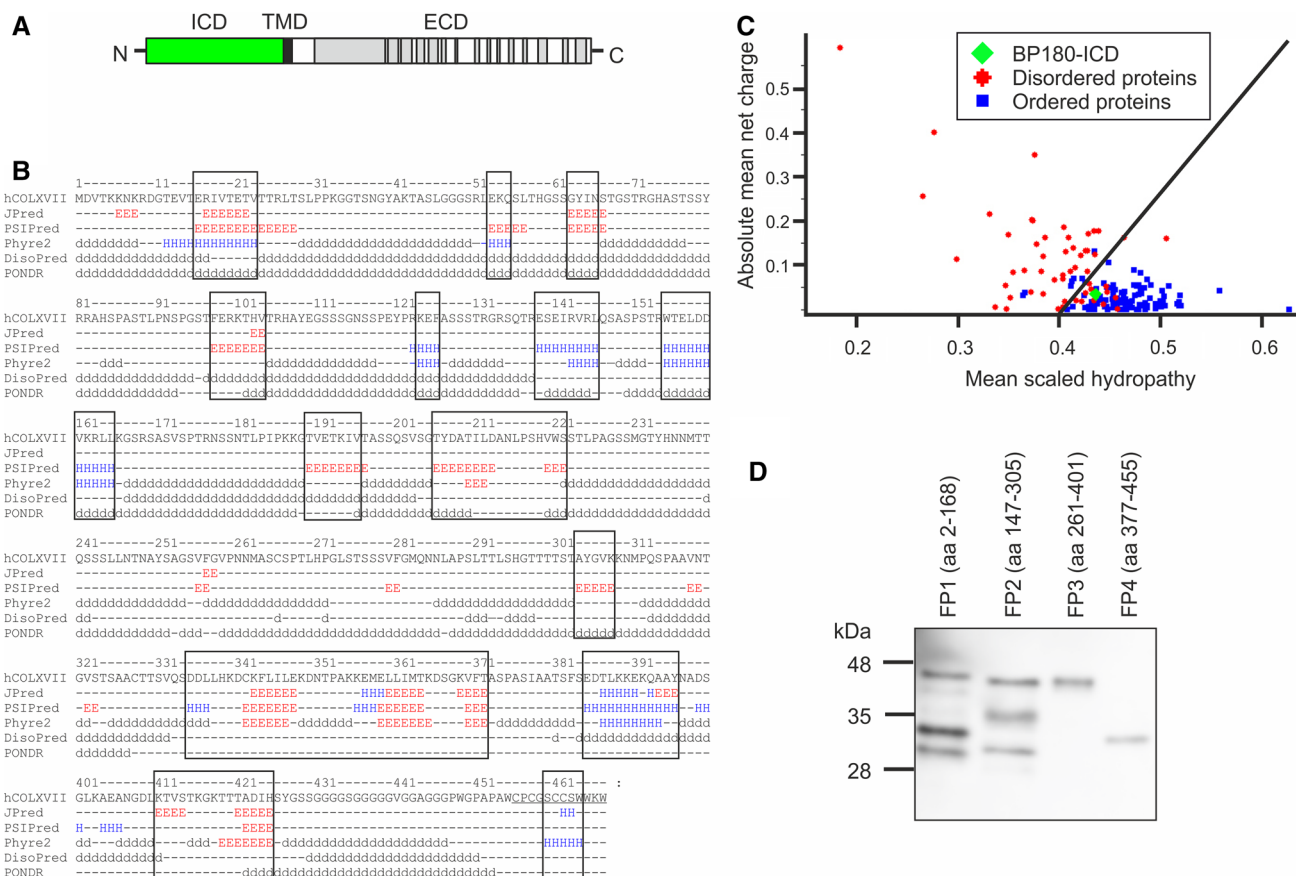
BP180, also known as collagen XVII and BPAG2, is a trimeric transmembrane collagen and a structural component of hemidesmosomes that anchors epidermal keratinocytes to the underlying basement membrane (Franzke et al. 2005; Walko et al. 2015). The importance of BP180 for epidermal adhesion is demonstrated by hereditary and acquired blistering skin diseases: junctional epidermolysis bullosa in which COL17A1 mutations causing either absence or dysfunction of BP180 lead to severe, lifelong blistering of skin and mucosa, and bullous pemphigoid in which autoantibodies against BP180 cause severe itch and widespread blistering in elderly people (Franzke et al. 2005).

BP180 belongs to the family of type II transmembrane proteins with collagenous structure (Franzke et al. 2005). This group consists of collagens XIII, XVII, XXIII, and XXV, which are mainly involved in cell–matrix interactions, and the more distantly related gliomedin, ectodysplasin A, macrophage scavenger receptors I and II, and macrophage receptor MARCO proteins, which have functions in the maintenance of the nodes of Ranvier in myelinated axons, the development of ectodermal organs, ligand binding, and internalization and microbial defence, respectively (Franzke et al. 2005). In this protein family, all four collagens, gliomedin, and ectodysplasin all undergo a shedding of their extracellular (ecto) domain.

BP180 has a large 120 kDa ectodomain with 16 non-collagenous (NC) domains interrupted by 15 collagenous sequences (Fig. 1a). The ectodomain, which is shed by

ADAM (a disintegrin and metalloprotease) proteases both constitutively and following inducible stimuli, is involved in wound healing (Franzke et al. 2002, 2004; Jackow et al. 2016). In contrast to the other transmembrane collagens that harbor only a short cytoplasmic stub, the intracellular domain of BP180 (BP180 ICD) consists of 467 amino acids (aa) and has a molecular weight of approximately 50 kDa (Fig. 1a) (Franzke et al. 2005). While the function of BP180's ectodomain has been the focus of intensive research, the structure and functions of its intracellular domain are yet to be fully elucidated.

BP180 ICD is known to interact with other components of the hemidesmosome, such as integrin  $\beta$ 4, the plakin family proteins BP230 and plectin and acidic14-3-3 $\sigma$ , and thereby connects hemidesmosomes with keratin intermediate filaments (Koster et al. 2003; Fontao et al. 2004;



**Fig. 1** Sequence modeling and immunoblotting suggest that the BP180 ICD has both disordered and folded regions. **a** The primary structure of BP180. The intracellular domain (ICD) is marked in green, the transmembrane domain in black, collagenous and non-collagenous regions in the extracellular domain (ECD) with gray and white, respectively. **b** Alignment of secondary structure/disorder status of the human BP180 ICD. Sequences predicted to have extended,  $\beta$ -strand secondary structure are marked with 'E', helices with 'H' and disordered regions with 'd'. Boxes indicate subdomains that may be folded. aa 456–467, not present in the truncated construct, are

underlined. **c** Plot of hydropathy vs. net charge by PONDR modeling. Red and blue dots represent known disordered and folded proteins, respectively. The position of BP180 ICD is marked with a green diamond. **d** GST-BP180 fusion proteins FP1 (aa 2–168), FP2 (147–305), FP3 (261–401) and FP4 (377–455) were expressed in *Escherichia coli* and glutathione sepharose pull downs (Tuusa et al. 2019) were analyzed by immunoblotting with anti-GST antibody. In the case of FP1 and FP2, the polypeptide of highest molecular weight represents the full-length fusion protein. Experiment was replicated three times

Li et al. 2007; Walko et al. 2015). It has also been implicated in the cytoplasmic signaling cascade that follows the shedding of the ectodomain and allows squamous carcinoma cell survival and proliferation (Galiger et al. 2018). Furthermore, phorbol ester-induced protein kinase C (PKC)-mediated phosphorylation of BP180 ICD leads to the dissociation of BP180 from hemidesmosomes (Kitajima et al. 1992, 1995), and anti-BP180 IgGs induce the PKC-dependent macropinocytosis-mediated internalization of BP180 (Iwata et al. 2016). To date, the structure of human BP180 ICD has been described only at the level of its primary architecture. However, electron microscopy of Triton X-100 solubilized purified bovine BP180 suggests a coiled-coil trimeric protein with a flexible extracellular tail and a globular intracellular head domain (Hirako et al. 1996). The amino acid sequence of BP180 ICD is known to harbor a few special features including four short repeats, a glycine tract, and a cysteine cluster in its carboxyterminal end, however, without known function (Giudice et al. 1992).

Here we have expressed the human BP180 ICD in *Escherichia coli* as a soluble polypeptide, purified, and characterized its physicochemical properties. Bacterially expressed BP180 ICD is a monomeric and intrinsically disordered in solutions, but undergoes partial folding in the presence of anionic detergent micelles or lipid vesicles.

## Materials and methods

### Bioinformatics

JPred4 (<https://www.compbio.dundee.ac.uk/jpred/>) (Drozdetskiy et al. 2015), PSIPred 4.0 (<https://bioinf.cs.ucl.ac.uk/psipred/>) (Buchan and Jones 2019) including DisoPred3 (Jones and Cozzetto 2015), PONDR (<https://www.pondr.com/>) (Romero and Dunker 1997; Li et al. 1999; Uversky et al. 2000) and Phyre2 (<https://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) (Kelley et al. 2015) servers were used to predict secondary structures and disordered regions.

### Expression and purification

BP180 ICD (encoding aa 2–467 and 2–455) was PCR amplified from human BP180 cDNA (Franzke et al. 2004), cloned into pGEX-2T (Tuusa et al. 2019) or modified pET-21a vector. It was then expressed in *Escherichia coli* BL21(DE3) in ZYM-5052 auto-induction medium (Studier 2014) at 18 °C as Glutathione-S-transferase (GST) or N- or C-terminal 6xhistidine (His)-tagged fusion proteins. To prepare the N-terminal His-tag, oligonucleotides 1 and 2 (Table 1) were hybridized to get a tag linker which was ligated into *NheI* and *BamHI* digested pET21a to obtain pET21a(2xHis). BP180 ICD (2–455) cDNA was amplified with oligonucleotide 3 and 4 (Table 1) primers, digested with *BamHI* and *MfeI* and ligated into *BamHI* and *EcoRI* digested pET21a(2xHis). A construct with a C-terminal His-tag was cloned by ligating oligonucleotide 5 and 6 (Table 1)—primer amplified cDNA directly into the *NheI* and *HindIII* restriction sites of pET21a. The DNA constructs encoding BP180 ICD fragments and affinity purification of the GST-fusion proteins were as described previously (Tuusa et al. 2019). To purify the His-tagged BP180 ICD, cell pellets were lysed by sonication in a buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM MgCl<sub>2</sub>, 1 mM dithiothreitol, 10 mM imidazole, 100 µg ml<sup>-1</sup> lysozyme, 10 µg ml<sup>-1</sup> DNaseI, and cOmplete EDTA-free protease inhibitor cocktail (Merck, KGaA, Darmstadt, Germany). Cytoplasmic extract was prepared by adjusting NaCl concentration to 300 mM and pelleting the debris by 20,000×g centrifugation for 40 min. The residual nucleic acids were removed by 0.1% polyethylenimine precipitation and centrifugation. BP180 ICD was purified with a Ni<sup>2+</sup>-NTA affinity column (Macherey–Nagel GmbH & co. KG, Düren, Germany) with a step-wise imidazole gradient. Peak fractions were combined, supplemented with 1 mM EDTA, dialyzed against 20 mM HEPES pH 7.8, 150 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, 5% glycerol and further purified using the ÄKTA Explorer 10 system (Pharmacia Biotech, Uppsala, Sweden), 1 ml HiTrap SP-FF cation exchange column (GE Healthcare Life Sciences, Chicago, USA) with a linear 150–600 mM NaCl gradient. Finally, the ÄKTA Pure system with HiLoad 16/60 Superdex 200 Increase size exclusion chromatography (SEC) column (GE

**Table 1** Oligonucleotides used in the cloning of DNA constructs

Name	Sequence	Polypeptide/tool/restriction site
1	5'-CTAGCCACCACCACCACCACCACG-3'	His-tag/linker
2	5'-GATCCGTGGTGGTGGTGGTGGTGG-3'	His-tag/linker
3	5'-CGATGGATCCGATGTAACCAAGAAAACA-3'	BP180 ICD/PCR primer/ <i>BamHI</i>
4	5'-AGCCCAATTGTCACCAGGCTGGC-3'	BP180 ICD/PCR primer/ <i>MfeI</i>
5	5'-CGATGCTAGCGATGTAACCAAGAAAACA-3'	BP180 ICD/PCR primer/ <i>NheI</i>
6	5'-AGCCAAGCTTCCAGGCTGGCGC-3'	BP180 ICD/PCR primer/ <i>HindIII</i>

Healthcare Life Sciences) was used to purify the protein to near homogeneity.

### Static light scattering

Molecular mass and oligomerization state were analyzed with a MiniDawn multi-angle light scattering (MALS) device (Wyatt Technology Corporation, Santa Barbara, USA) connected to a Shimadzu high-performance liquid chromatography (HPLC) unit (Shimadzu Corporation, Kyoto, Japan) with a Superdex 200 Increase SEC column (GE Healthcare Life Sciences). The RID-10A refractive index detector (Shimadzu Corporation) connected to the HPLC system was used as a concentration source for the calculations. ASTRA software (Wyatt Technology Corporation) was used to calculate molecular weight.

### Circular dichroism spectroscopy

Samples were dialyzed against a 10 mM potassium phosphate pH 7.5, 150 mM NaF buffer. Then a Chirascan™ circular dichroism (CD) spectrometer (Applied Photophysics Ltd, Leatherhead, UK) with 1 mm pathlength quartz cuvettes was used to measure CD spectra at 190–280 nm wavelengths under the following conditions: sample in buffer alone; immediately after the addition of sodium dodecyl sulfate (SDS), *n*-dodecylphosphocholine (DPC) (Merck), unilamellar lipid vesicles containing dimyristoylphosphatidylcholine (DMPC)(Sigma) and dimyristoylphosphatidylglycerol (DMPG)(Sigma) in 1:1 molar ratio, or calcium acetate; immediately after the addition of both lipid vesicles and calcium acetate. The unilamellar lipid vesicles were prepared as described before (Tuusa et al. 2017). Every spectrum is an average of three scans. The Bestsel (Micsonai et al. 2015) web server was used for spectrum deconvolution and to provide estimations of the contents of the protein's secondary structure.

### SDS-PAGE, co-sedimentation analysis and immunoblotting

Proteins were analyzed by separating them in 4–20% polyacrylamide gradient TGX gels (Bio-Rad Laboratories, Inc. Hercules, USA) then subjecting them to PageBlue (Thermo Fisher Scientific, Inc. Waltham, USA) staining or immunoblotting. Proteins were identified by the Proteomics and Protein Analysis core service at the Biocenter of the University of Oulu, Finland, using in-gel digestion and MALDI-TOF mass spectrometry. For co-sedimentation analysis of lipid vesicles, 21 μM BP180 ICD was incubated with DMPC:DMPG vesicles using a protein:lipid molar ratio of 1:20 for 1 h at 22 °C followed by 1 h centrifugation at 20,000×g. Samples from supernatants and dissolved pellets

were analyzed by immunoblotting using the Endo-2 anti-BP180 ICD antibody (Franzke et al. 2002). GST fusion proteins were detected by an anti-GST monoclonal antibody (clone 8–326; Thermo Fisher Scientific, Inc).

## Results

### Bioinformatic analysis suggests that the BP180 ICD has a disordered structure

Sequence analysis did not provide strong homology-based hints regarding the structure of the BP180 ICD. PSI-BLAST searches showed only a few non-collagen XVII protein sequences that aligned locally with the BP180 ICD, and those had low (<30%) sequence identity. These included the heavily O-glycosylated extracellular domain of CD45 tyrosine phosphatase, which likely has a disordered structure (data not shown). Phyre2 modeling, which was done both for full-length BP180 and for its ICD, indicated some folded regions dispersed among disordered sequences (Fig. 1b and data not shown). The modeling of the ICD as a part of the full-length protein provided a few modest hits for bacterial glycosyltransferases. It is estimated that 70% of the ICD is composed of disordered regions.

We then used the JPred4 and PSIPred 4.0 software packages to determine secondary structures and the DisoPred3 and PONDR packages to map disordered regions. Taking these findings together, it appears that the BP180 ICD might have a few secondary structures, in its N-terminal half and a larger folded region between aa 335 and 425 (Fig. 1b). In the net charge vs. hydropathy plot, the BP180 ICD was mapped on the side of the folded proteins, but close to the border between the ordered and disordered regions (Fig. 1c). These predictions are in line with our observations that the GST-BP180 fusion proteins FP1 (containing aa 2–168) and FP2 (aa 147–305) were accompanied by degradation and/or truncated translation products typical for disordered proteins, when expressed in *Escherichia coli*, whereas the fusion proteins FP3 (aa 261–401) and FP4 (aa 377–455) were expressed as full-length soluble proteins without degradation or premature termination (Fig. 1d). Taken together, while the bioinformatic modeling does not exclude the possibility of a stable folded structure, it suggests that the BP180 ICD is intrinsically partly disordered.

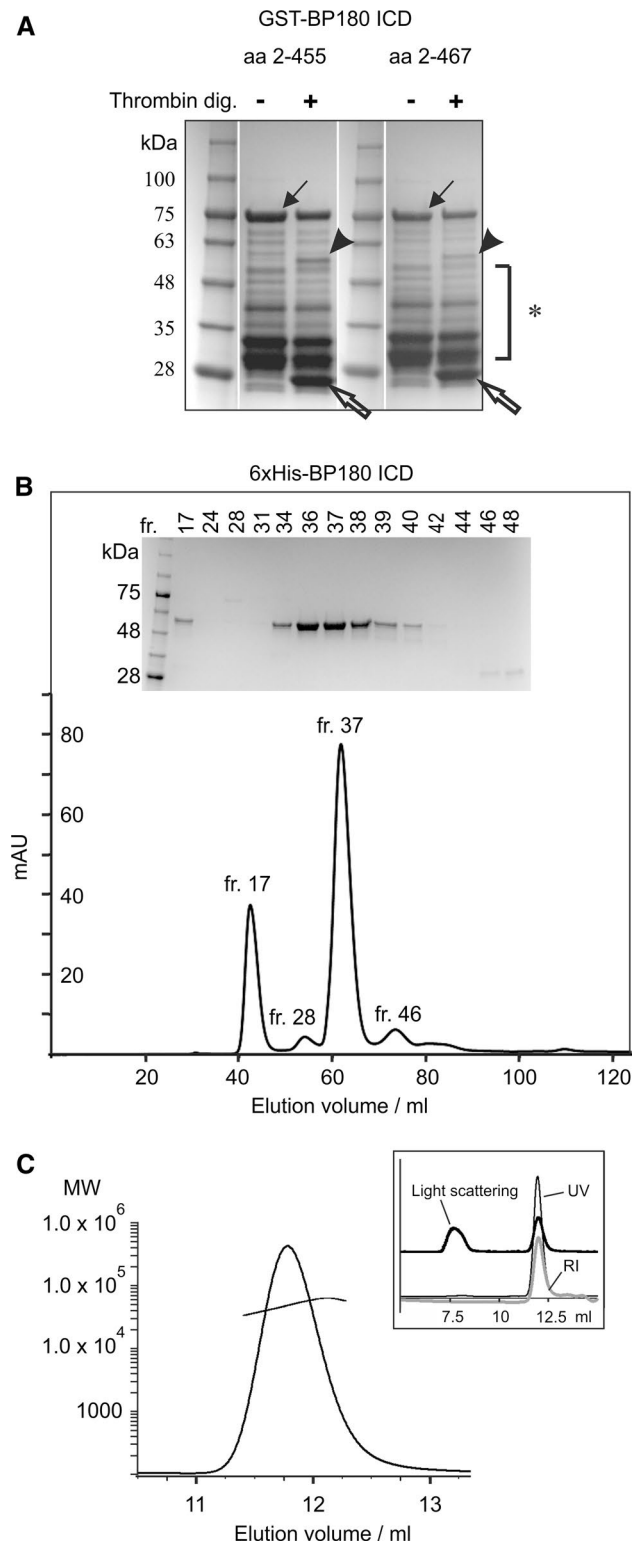
### Bacterially expressed and purified BP180 ICD is prone to aggregation in an environment with low ionic strength

To study the structural properties of the BP180 ICD, we first expressed GST-BP180 ICD using constructs encoding aa residues 2–467 and 2–455 in the *Escherichia coli* strain

**Fig. 2** In *Escherichia coli*, the BP180 ICD is expressed as a soluble monomer. **a** SDS-PAGE analysis of the BP180 ICD (aa2–455 and aa2–467). The solid arrows indicate the 75 kDa GST-BP180 ICD with overlapping *Escherichia coli* DnaK contamination. The thrombin cleaved BP180 ICD, GST and truncated translation/degradation products are marked with an arrowhead, open arrow and an asterisk respectively. **b** The UV-curve of HiLoad 16/60 SEC and corresponding protein gel of the BP180 ICD with a N-terminal 6xHis-tag. **c** A sample was analyzed with Superdex 200 Increase SEC-coupled MALS-instrument. Light scattering analysis of the elution peak, mass distribution, and molecular weight of purified His-tagged BP180 ICD. The UV, refractive index, and light scattering profile are shown in the box. MALS was performed twice from both of the two protein purifications

BL21 (DE3), assuming that the GST tag would promote the folding of a fusion partner. Both GST fusion proteins were expressed as soluble proteins with the accompanying degradation and/or truncated translation products (Fig. 2a). The GST fusion BP180 ICD aa 2–455 (hereafter GST-BP180 ICD), which lacks 12 C-terminal amino acids including four cysteine residues, had a slightly higher yield (Fig. 2a) and was selected for further analysis. The GST-BP180 ICD was purified by glutathione sepharose affinity chromatography followed by dialysis, thrombin digestion to remove the cleaved GST tag and cation exchange chromatography. Despite the comprehensive optimization of its expression and purification conditions, the BP180 ICD was found to be prone to aggregation in low-salt conditions during thrombin digestion. Furthermore, the MALDI mass spectrometry analysis of tryptic fragments of the 75 kDa protein band visible in SDS-PAGE confirmed the identity of GST-BP180, but also revealed the co-purification of the bacterial chaperone protein DnaK, which appears to have the same molecular weight (data not shown), a finding that may indicate the presence of unburied hydrophobic residues in the BP180 ICD.

Since the relatively large GST (28 kDa) might also sterically hinder the folding of the BP180 ICD, we changed to a short 6xHis affinity tag. Both N- or C-terminal 6xHis-tagged fusion proteins (His-BP180 ICD, BP180 ICD-His) were expressed as soluble proteins. The His-BP180 ICD was purified in four steps: The removal of the DNase resistant nucleic acid remnants by polyethylenimine precipitation was followed by Ni affinity, cation exchange, and size exclusion chromatographies. This process yielded near-homogenous His-BP180 ICD (Fig. 2b). At low concentration ( $\leq 1$  mg  $\text{ml}^{-1}$ ) in a high-salt buffer ( $\geq 300$  mM), the His-BP180 ICD was stable at +4 °C for several weeks. However, the His-BP180 ICD was prone to aggregation at higher protein concentrations, in dialysis against a low-salt buffer or after a freeze–thaw cycle. This suggests that the BP180 ICD can be expressed in a bacterial expression system as a soluble protein and purified to homogeneity, but may have a natively unfolded structure under these conditions.



### Bacterially expressed BP180 ICD is a monomer

In mammalian cells, the full-length BP180 is a trimer in which the extracellular NC16A domain forms a coiled coil, triggering the collagenous domains to form a triple helical

assembly (Kroeger et al. 2017). However, it is unknown whether the cytoplasmic ICD domains are able to bind to each other independently. This was examined by static light scattering.

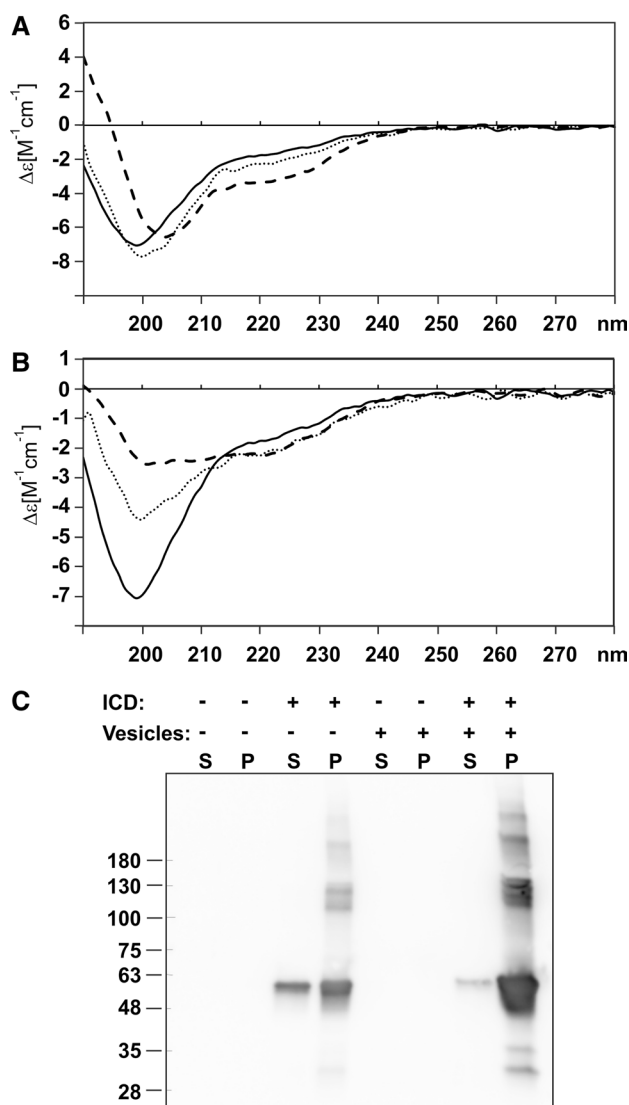
Purified His-BP180 ICD was run through SEC column and analyzed by refractive index detector and MALS detector unit. Two scattering peaks were detected (Fig. 2c). The first peak eluted in the void volume at a low protein concentration (below  $UV_{280nm}$  detection) and the main light scattering peak with a high UV absorption eluted at 24 min (Fig. 2c). The determined molecular weight, 50 kDa (Fig. 2c), corresponds with both the theoretical value of 48 kDa calculated from the amino acid sequence and the apparent size of 55 kDa in SDS-PAGE. Importantly, the light scattering data indicated that His-BP180 ICD, without the transmembrane and extracellular domains of full-length BP180, exists as a monomer in solution.

### BP180 ICD is intrinsically disordered but partially folds in the presence of an anionic membrane mimicking detergent micelles and lipid vesicles

We used circular dichroism (CD) spectroscopy to investigate whether the BP180 ICD has a folded structure and to measure its secondary structure elements. The CD spectrum of BP180 ICD resembles one typical of a random coil structure (Fig. 3a), suggesting that bacterially expressed BP180 ICD is a disordered protein.

In mammalian cells, the positively charged ICD ( $pI=9.61$ ) of BP180 is localized next to the cell membrane, which generally has a negative net charge on the cytoplasmic leaflet (McLaughlin 1989). Therefore, we tested whether a low concentration of negatively charged sodium dodecyl sulfate (SDS) affects the structure of the BP180 ICD. SDS micelles induced an increase of CD signal between 190 and 205 nm, a clear shift of the minimum wavelength from 200 to 205 nm and a change in the shape of the spectrum at 220 nm (Fig. 3a). The Bestsel analysis indicated an increase from 12.0 to 23.1% in the relative proportion of  $\alpha$ -helices upon SDS addition (Fig. 3a). In contrast, the zwitterionic detergent *n*-dodecylphosphocholine (DPC) did not have any effect (Fig. 3a).

Next, we measured the CD spectrum for BP180 ICD in the presence of negatively charged unilamellar DMPC:DMPG lipid vesicles with a 1:100 protein:lipid molar ratio. A clear change in the CD spectrum was seen (Fig. 3b), with a more pronounced negative CD signal at the 222 nm wavelength as well as a decrease in the negative peak at 198 nm. This indicates a partially  $\alpha$ -helical conformation with some regions remaining disordered. Interestingly, the inclusion of 1 mM calcium with the lipid vesicles partially reversed the effect of the lipids (Fig. 3b). Finally, we expressed and purified C-terminally tagged BP180 ICD-His to ensure that the



**Fig. 3** The BP180 ICD is an intrinsically disordered protein that undergoes partial folding in an anionic lipid environment. **a** CD spectra of BP180 ICD (solid curve,  $n=3$ ), BP180 ICD+neutral lipid *n*-dodecyl phosphocholine (dotted curve,  $n=3$ ) and BP180 ICD+anionic lipid sodium dodecyl sulphate (dashed curve,  $n=3$ ). **b** CD spectra of BP180 ICD (solid curve,  $n=3$ ), BP180 ICD+DMPC:DMPG lipid vesicles (dashed curve,  $n=3$ ) and BP180 ICD+DMPC:DMPG lipid vesicles with 1 mM calcium acetate (dotted curve,  $n=2$ ). CD spectra were averaged of three scans in each of three (two in case of calcium) independent measurements. Representative series are shown. **c** Samples with or without BP180 ICD and lipid vesicles were incubated at room temperature and centrifuged at  $20,000\times g$  for 60 min. Pellets (P) and supernatants (S) were analyzed by immunoblotting against BP180 ICD. Sedimentation analysis was performed twice with identical results

lack of secondary structures was not due to the non-native N-terminus. The CD spectrum measured after two purification steps showed the BP180 ICD-His to have a disordered structure (Online Resource—Fig. S1). Taken together, these CD spectroscopy findings suggest that bacterially expressed

BP180 ICD is intrinsically disordered in solution but can acquire partial folding upon interaction with negatively charged lipid vesicles.

### BP180 ICD binds to lipid vesicles

Finally, we further analyzed the interaction of BP180 ICD with lipid vesicles. It is known that many peripheral membrane-binding proteins, such as myelin protein P2, can aggregate membrane vesicles in vitro (Tuusa et al. 2017). Therefore, we incubated His-BP180 ICD with and without DMPC-DMPG vesicles (1:100 protein:lipid molar ratio) and centrifuged sedimentable material. The majority of BP180 ICD was co-sedimented with DMPC-DMPG vesicles, while a small amount of protein stayed in the supernatant (Fig. 3c). Little aggregation and sedimentation of BP180 ICD was visible in the absence of lipid vesicles. These results confirm the binding of the BP180 ICD to negatively charged lipid vesicles.

### Discussion

BP180 is unique among the transmembrane collagens, having a relatively large cytoplasmic domain. The structure of its intracellular N-terminus is poorly characterized and the current knowledge of its function is limited to its interactions with other hemidesmosomal proteins. In the present study, we expressed the human BP180 ICD in *Escherichia coli*, purified it to homogeneity and characterized it as an intrinsically disordered polypeptide.

The BP180 ICD was expressed as a soluble protein with or without its cysteine-rich C-terminus, or its native N-terminus and independently of tag size (GST or His). Previously, electron microscopy of full-length bovine BP180 has suggested that the ICD has a globular shape (Hirako et al. 1996) and protein–protein-interaction motifs within the first 400 amino acids have been mapped using a yeast two-hybrid system (Fontao et al. 2004). Our findings suggest that bacterially expressed BP180 ICD lacks a stable global folding pattern as suggested by bioinformatics and demonstrated by the presence of degradation and/or prematurely terminated translation products with the GST fusions of the N-terminal ICD fragments, co-purification of bacterial chaperone DnaK, the instability of the protein in high concentration, in low ionic strength or during freezing–melting and most importantly by the CD spectrum typical for disordered proteins.

We cannot exclude the possibility that the folding of BP180 ICD requires chaperoning or post-translational modifications that are not present in the bacterial expression system. Also, although the GST-BP180 ICD (aa 2–467) with the cysteine rich terminus had rather lower yield than the truncated BP180 ICD (aa 2–455) and cysteines should be

reduced in the physiological cytosolic environment in basal keratinocytes (Wolf et al. 2014), we cannot totally exclude the role of rare cytosolic disulfides to the folding of the BP180 ICD in mammalian cell. Another possible explanation is that the folding of BP180 ICD may depend upon the transmembrane and extracellular domains of BP180 and/or the trimerization mediated by the extracellular domain (Balding et al. 1997). However, our results raise the possibility that even in hemidesmosomes, the BP180 ICD may be natively unfolded, especially in the N-terminal regions, unless they interact with membrane lipids and/or acidic protein partners. Both the bioinformatics analysis and observations that membrane proximal part of the BP 180 ICD may have higher tendency to fold compared to N-terminus (Fig. 1b, d) support the hypothesis that lipid membrane may affect the structure of BP 180 ICD. Interestingly, the isoelectric point (pI) of the region comprising the fibronectin domains III and IV of beta-4 integrin is 5.15 and these domains are critical for the interaction of beta-4 integrin with the BP180 ICD (Schaapveld et al. 1998). Similarly, the acidic region of BP230 (aa 1–555, pI = 5.41) has been shown to interact with the BP180 ICD. In contrast, the BP180 ICD is basic (pI = 9.61), particularly so in the region of aa 145–230 (pI = 9.99), which is critical for the interaction between the BP180 ICD and BP230 (Koster et al. 2003). Finally, the acidic regulatory protein 14-3-3 $\sigma$  binds to the BP180 ICD (Li et al. 2007).

Interestingly, we found that the interactions of BP180 ICD with anionic detergent micelles or glycerophospholipids induce partial folding, which hints that the folding and unfolding of the BP180 ICD are connected to its interaction with anionic membrane head groups and/or other negatively charged macromolecules. We observed that a divalent cation (calcium) partially blocked the folding induced by the negatively charged lipid vesicle. This finding may be related to the physiological phenomenon, whereby an increase in cytosolic calcium and the associated activation of protein kinase C leads to the phosphorylation of the BP180 ICD and dissociation of BP180 from hemidesmosomes (Kitajima et al. 1992, 1995). Taken together, these findings suggest that coulombic interactions can regulate the structure and function of BP180 on the cytoplasmic side of membrane. Importantly, it has been estimated that approximately 50% of transmembrane proteins have at least one disordered region at least 30 amino acids in length, and that these are concentrated on the cytoplasmic side (Burgi et al. 2016). Of note, macrophage scavenger receptor I, a transmembrane collagen family protein with type II topology, has a disordered cytoplasmic domain 50 amino acids long, which tends to bind to molecular chaperones (Nakamura et al. 2002; Franzke et al. 2005). Disordered regions may allow interaction-induced folding that provides a means by which to regulate

cellular functions. This has previously been demonstrated in proteins such as cyclin-dependent kinase inhibitor p27kip1 (Lacy et al. 2004),  $\alpha$ -synuclein (Davidson et al. 1998) and myelin basic protein (Raasakka et al. 2017). It is an interesting question whether the intrinsically disordered character of the BP180 ICD is involved in the regulation of hemidesmosome assembly and disassembly and/or to other protein–protein or protein–lipid interactions.

BP180 is the major antigen in bullous pemphigoid. The immunodominant epitope is the extracellular juxtamembraneous NC16A domain, but autoantibodies against the carboxyterminal extracellular epitopes are also frequently found. The phenomenon of epitope spreading has been found to take place in bullous pemphigoid and includes autoantibodies directed against epitopes located in the intracellular domain (Di Zenzo et al. 2011). Whether the latter are pathogenic is currently unknown, but our recent study demonstrated that some autoantibodies against the BP180 ICD are common among patients with Alzheimer's disease and those with multiple sclerosis and not rare among healthy subjects. These autoantibodies do not bind native protein in the skin and are likely not pathogenic. Our current finding indicates that the BP180 ICD is intrinsically disordered making it prone to degradation and altered processing, which may explain neoepitope presentation by BP180 and its relative instability compared with the other hemidesmosomal components (Liu et al. 2019).

In conclusion, our results suggest that the cytoplasmic domain of BP180 is intrinsically disordered and has potential for partial folding when exposed to negatively charged lipids. Further research is required to characterize the structure and dynamics of the BP180 ICD in its natural context.

**Acknowledgements** Open access funding provided by University of Oulu including Oulu University Hospital. We address thanks to Drs Ulrich Bergmann and Hongmin Tu, (Biocenter Oulu, Proteomics and protein analysis core facility) for their expertise in mass spectrometry and CD spectrometry, respectively, Ville Ratas (Biocenter Oulu, Protein crystallography core facility) and Ms Sirpa Rannikko (Department of Medical Microbiology and Immunology, Research Unit of Biomedicine and Medical Research Center) for help with light scattering and size-exclusion chromatography, and with ion-exchange chromatography, respectively, all in the University of Oulu, Finland.

**Author contributions** All authors contributed to the study design and suggested improvements to manuscript draft. Experiments were performed and the data were analyzed by Jussi Tuusa, M. Kristian Koski and Salla Ruskamo. Jussi Tuusa wrote the first draft of the manuscript. Kaisa Tasanen supervised the research. All authors read and approved the final manuscript.

**Funding** This research was funded by grants from the Sigrid Juselius Foundation to Jussi Tuusa and Kaisa Tasanen, and from the Academy of Finland to Kaisa Tasanen (#294738) and Salla Ruskamo (#275225).

## Compliance with ethical standards

**Conflict of interest** The authors declare no conflict of interest.

**Informed consent** All authors are aware of the details of their research work that are presented in the current manuscript and gave their consent to publication.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Balding SD, Diaz LA, Giudice GJ (1997) A recombinant form of the human BP180 ectodomain forms a collagen-like homotrimeric complex. *Biochemistry* 36:8821–8830. <https://doi.org/10.1021/bi970675n>
- Buchan DWA, Jones DT (2019) The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Res* 47:W402–W407. <https://doi.org/10.1093/nar/gkz297>
- Burgi J, Xue B, Uversky VN, van der Goot FG (2016) Intrinsic disorder in transmembrane proteins: roles in signaling and topology prediction. *PLoS ONE* 11:e0158594. <https://doi.org/10.1371/journal.pone.0158594>
- Davidson WS, Jonas A, Clayton DF, George JM (1998) Stabilization of alpha-synuclein secondary structure upon binding to synthetic membranes. *J Biol Chem* 273:9443–9449. <https://doi.org/10.1074/jbc.273.16.9443>
- Di Zenzo G, Thoma-Uszynski S, Calabresi V, Fontao L, Hofmann SC, Lacour JP, Sera F, Bruckner-Tuderman L, Zambruno G, Borradori L et al (2011) Demonstration of epitope-spreading phenomena in bullous pemphigoid: results of a prospective multicenter study. *J Invest Dermatol* 131:2271–2280. <https://doi.org/10.1038/jid.2011.180>
- Drozdetskiy A, Cole C, Procter J, Barton GJ (2015) JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* 43:W389–W394. <https://doi.org/10.1093/nar/gkv332>
- Fontao L, Tasanen K, Huber M, Hohl D, Koster J, Bruckner-Tuderman L, Sonnenberg A, Borradori L (2004) Molecular consequences of deletion of the cytoplasmic domain of bullous pemphigoid 180 in a patient with predominant features of epidermolysis bullosa simplex. *J Invest Dermatol* 122:65–72. <https://doi.org/10.1046/j.0022-202X.2003.22125.x>
- Franzke CW, Tasanen K, Schacke H, Zhou Z, Tryggvason K, Mauch C, Zigrino P, Sunnarborg S, Lee DC, Fahrenholz F et al (2002) Transmembrane collagen XVII, an epithelial adhesion protein, is shed from the cell surface by ADAMs. *EMBO J* 21:5026–5035. <https://doi.org/10.1093/emboj/cdf532>
- Franzke CW, Tasanen K, Borradori L, Huotari V, Bruckner-Tuderman L (2004) Shedding of collagen XVII/BP180:



- structural motifs influence cleavage from cell surface. *J Biol Chem* 279:24521–24529
- Franzke CW, Bruckner P, Bruckner-Tuderman L (2005) Collagenous transmembrane proteins: recent insights into biology and pathology. *J Biol Chem* 280:4005–4008
- Galiger C, Loffek S, Stemmler MP, Kroeger JK, Mittapalli VR, Fauth L, Esser PR, Kern JS, Meiss F, Lassmann S et al (2018) Targeting of cell surface proteolysis of collagen XVII impedes squamous cell carcinoma progression. *Mol Ther* 26:17–30
- Giudice GJ, Emery DJ, Diaz LA (1992) Cloning and primary structural analysis of the bullous pemphigoid autoantigen BP180. *J Invest Dermatol* 99:243–250. <https://doi.org/10.1111/1523-1747.ep12616580>
- Hirako Y, Usukura J, Nishizawa Y, Owaribe K (1996) Demonstration of the molecular shape of BP180, a 180-kDa bullous pemphigoid antigen and its potential for trimer formation. *J Biol Chem* 271:13739–13745
- Iwata H, Kamaguchi M, Ujiie H, Nishimura M, Izumi K, Natsuga K, Shinkuma S, Nishie W, Shimizu H (2016) Macropinocytosis of type XVII collagen induced by bullous pemphigoid IgG is regulated via protein kinase C. *Lab Invest* 96:1301–1310. <https://doi.org/10.1038/labinvest.2016.108>
- Jackow J, Schlosser A, Sormunen R, Nystrom A, Sitaru C, Tasanen K, Bruckner-Tuderman L, Franzke CW (2016) Generation of a functional non-shedding collagen XVII mouse model: relevance of collagen XVII shedding in wound healing. *J Invest Dermatol* 136:516–525. <https://doi.org/10.1016/j.jid.2015.10.060>
- Jones DT, Cozzetto D (2015) DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* 31:857–863. <https://doi.org/10.1093/bioinformatics/btu744>
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10:845–858. <https://doi.org/10.1038/nprot.2015.053>
- Kitajima Y, Owaribe K, Nishizawa Y, Jokura Y, Yaoita H (1992) Phorbol ester- and calcium-induced reorganization of 180-kDa bullous pemphigoid antigen on the ventral surface of cultured human keratinocytes as studied by immunofluorescence and immunoelectron microscopy. *Exp Cell Res* 203:17–24
- Kitajima Y, Owada MK, Fujisawa Y, Seishima M, Yaoita H, Hirako Y, Owaribe K (1995) A hemidesmosomal transmembrane collagenous molecule, the 180-kDa bullous pemphigoid antigen (BPA II), is phosphorylated with 12-O-tetradecanoylphorbol-13-acetate in a human squamous cell carcinoma cell line (DJM-1). *Epithelial Cell Biol* 4:70–75
- Koster J, Geerts D, Favre B, Borradori L, Sonnenberg A (2003) Analysis of the interactions between BP180, BP230, plectin and the integrin alpha6beta4 important for hemidesmosome assembly. *J Cell Sci* 116:387–399. <https://doi.org/10.1242/jcs.00241>
- Kroeger JK, Hofmann SC, Leppert J, Has C, Franzke CW (2017) Amino acid duplication in the coiled-coil structure of collagen XVII alters its maturation and trimerization causing mild junctional epidermolysis bullosa. *Hum Mol Genet* 26:479–488. <https://doi.org/10.1093/hmg/ddw404>
- Lacy ER, Filippov I, Lewis WS, Otieno S, Xiao L, Weiss S, Hengst L, Kriwacki RW (2004) p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nat Struct Mol Biol* 11:358–364. <https://doi.org/10.1038/nsmb746>
- Li X, Romero P, Rani M, Dunker AK, Obradovic Z (1999) Predicting protein disorder for N-, C-, and internal regions. *Genome Inform Ser Workshop Genome Inform* 10:30–40
- Li Y, Lin X, Kilani RT, Jones JC, Ghahary A (2007) 14–3–3 sigma isoform interacts with the cytoplasmic domain of the transmembrane BP180 in keratinocytes. *J Cell Physiol* 212:675–681. <https://doi.org/10.1002/jcp.21064>
- Liu N, Matsumura H, Kato T, Ichinose S, Takada A, Namiki T, Asakawa K, Morinaga H, Mohri Y, De Arcangelis A et al (2019) Stem cell competition orchestrates skin homeostasis and ageing. *Nature* 568:344–350. <https://doi.org/10.1038/s41586-019-1085-7>
- McLaughlin S (1989) The electrostatic properties of membranes. *Annu Rev Biophys Chem* 18:113–136. <https://doi.org/10.1146/annurev.bb.18.060189.000553>
- Micsonai A, Wien F, Kernya L, Lee YH, Goto Y, Refregiers M, Kardos J (2015) Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc Natl Acad Sci USA* 112:E3095–E3103. <https://doi.org/10.1073/pnas.1500851112>
- Nakamura T, Hinagata J, Tanaka T, Imanishi T, Wada Y, Kodama T, Doi T (2002) HSP90, HSP70, and GAPDH directly interact with the cytoplasmic domain of macrophage scavenger receptors. *Biochem Biophys Res Commun* 290:858–864. <https://doi.org/10.1006/bbrc.2001.6271>
- Raasakka A, Ruskamo S, Kowal J, Barker R, Baumann A, Martel A, Tuusa J, Myllykoski M, Burck J, Ulrich AS et al (2017) Membrane association landscape of myelin basic protein portrays formation of the myelin major dense line. *Sci Rep* 7:4974017–5053643. <https://doi.org/10.1038/s41598-017-05364-3>
- Romero O, Dunker K (1997) Sequence data analysis for long disordered regions prediction in the calcineurin family. *Genome Inform Ser Workshop Genome Inform* 8:110–124
- Schaapveld RQ, Borradori L, Geerts D, van Leusden MR, Kuikman I, Nievers MG, Niessen CM, Steenbergen RD, Snijders PJ, Sonnenberg A (1998) Hemidesmosome formation is initiated by the beta4 integrin subunit, requires complex formation of beta4 and HD1/plectin, and involves a direct interaction between beta4 and the bullous pemphigoid antigen 180. *J Cell Biol* 142:271–284. <https://doi.org/10.1083/jcb.142.1.271>
- Studier FW (2014) Stable expression clones and auto-induction for protein production in *E. coli*. *Methods Mol Biol* 1091:17–32. [https://doi.org/10.1007/978-1-62703-691-7\\_2](https://doi.org/10.1007/978-1-62703-691-7_2)
- Tuusa J, Raasakka A, Ruskamo S, Kursula P (2017) Myelin-derived and putative molecular mimic peptides share structural properties in aqueous and membrane-like environments. *Mult Scler Demyelinating Disord* 2:4
- Tuusa J, Lindgren O, Tertunen HM, Nishie W, Kokkonen N, Huilaja L, Izumi K, Herukka SK, Miettunen J, Shimizu H et al (2019) BP180 autoantibodies target different epitopes in multiple sclerosis or Alzheimer's disease than in bullous pemphigoid. *J Invest Dermatol* 139:293–299. <https://doi.org/10.1016/j.jid.2018.09.010>
- Uversky VN, Gillespie JR, Fink AL (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* 41:415–427. [https://doi.org/10.1002/1097-0134\(2000115\)41:3%3c415:AID-PROT130%3e3.0.CO;2-7](https://doi.org/10.1002/1097-0134(2000115)41:3%3c415:AID-PROT130%3e3.0.CO;2-7)
- Walko G, Castanon MJ, Wiche G (2015) Molecular architecture and function of the hemidesmosome. *Cell Tissue Res* 360:529–544. <https://doi.org/10.1007/s00441-015-2216-6>
- Wolf AM, Nishimaki K, Kamimura N, Ohta S (2014) Real-time monitoring of oxidative stress in live mouse skin. *J Invest Dermatol* 134:1701–1709. <https://doi.org/10.1038/jid.2013.428>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.