Structural analysis and expression of human desmoglein: a cadherin-like component of the desmosome

LAURA A. NILLES1, DAVID A. D. PARRY2, ERIC E. POWERS1, BRIGITT D. ANGST1, RITA M. WAGNER1 and KATHLEEN J. GREEN1*

1Department of Pathology and the Cancer Center, Northwestern University Medical School, 303 E. Chicago Avenue, Chicago, Illinois 60611, USA
2Department of Physics and Biophysics, Massey University, Palmerston North, New Zealand

* Author for correspondence

Summary

Desmosomes are adhesive cell junctions found in great abundance in tissues that experience mechanical stress. The transmembrane desmosomal glycoproteins have been proposed to play a role in cell adhesion; desmoglein I (DGI) is a major member of this class of desmosomal molecules. However, evidence supporting a role for DGI in cell adhesion or in the plaque is lacking. In order to begin to understand DGI function we have identified human cDNA clones encoding the entire mature polypeptide of 1000 amino acids. Our data suggest that like the bovine DGI molecule human DGI is highly related to the calcium-dependent class of cell adhesion molecules known as cadherins. Four related extracellular domains located in the amino-terminal domain of the molecule contain putative calcium binding sites originally identified in the cadherins. The highest degree of similarity between human N-cadherin and human DGI, and likewise between bovine DGI and human DGI, is greatest in the most amino-terminal extracellular domain. This suggests a conserved functional role for the extracellular domains, perhaps in calcium-mediated cell adhesion. The cytoplasmic portion of the molecule contains a cadherin-like region and, like bovine DGI, a carboxy-terminal tail that is not present in the cadherins, comprising three additional domains. One of these contains a novel repeating motif of 29±1 residues, first identified in bovine DGI. Each of the highly homologous repeating units is likely to consist of two β-strands and two turns with special characteristics. Five amino acids that are identical in bovine and human DGI lie in the second of the two predicted β-strands, and intriguingly contain putative target sites for protein kinase C. On the basis of structural analysis, a model predicting the disposition of human DGI domains in the desmosome is proposed.

Northern analysis suggests that unlike bovine epidermis, which expresses a single mRNA of reported size ~7.6kb, human foreskin and cultured keratinocytes display a complex pattern with bands of ~7.2, 4.0 and 3.0 kb. Each of these cross-hybridizing mRNAs is coordinately expressed in normal human keratinocytes in response to long-term culture and increased calcium.

Key words: desmoglein, desmosome, cadherin, cell adhesion.

Introduction

A variety of mechanisms designed to mediate intercellular and cell-substratum adhesion have developed during vertebrate evolution. These mechanisms employ a number of adhesion molecules and adhesive structures that provide the necessary functional specificity required during embryonic development and adult differentiation (for reviews see Garrod, 1986; Takeichi, 1988; Damásky, 1989; Edelman et al. 1990). It should be noted, however, that these groups are not mutually exclusive, since some CAMs that appear first during development, such as uvomorulin (also known as E-cadherin and L-CAM), are often later localized in intercellular junctions (Takeichi, 1988).

One major type of adhesive cell junction is the desmosome (reviewed by Staehelin, 1974; Bock and Clark, 1987; Kapprell et al. 1990; Schwarz et al. 1990; Garrod et al. 1990; Green and Jones, 1990; Jones and Green, 1991). Desmosomes are found in great abundance in tissues that experience mechanical stress. Although these tissues primarily include epithelia, desmosomes are also found in cardiac muscle, arachnoidal cells of the brain, and dendritic reticulum cells of germinal centers of lymph nodes (Schwarz et al. 1990). A number of desmosomal CJsMs have been identified. These can be divided into two major groups, based on whether they fractionate with the urea-soluble 'plaque' component or the urea-insoluble 'core' (Gorbsky and Steinberg, 1981; Cohen et al. 1983; Jones et al. 1988). The former group comprises proteins...
such as desmoplakins (DP) I and II, which may play a role in linking the intermediate filament cytoskeleton to the cell surface (Schwarz et al. 1990; Green and Jones, 1990; Jones and Green, 1991). Molecules comprising the latter group are thought to be prime candidates for mediating intercellular adhesion (Gorsky and Steinberg, 1981;Cowin et al. 1984). These include desmoglein I (DGI, also known as dgI or band 3), desmoglein II, b (DGII,b), also known as dg2/5, desmocollins I and II, bands 4a,b), and desmoglein III (DGIII also known as dg3) (Bock and Clark, 1987; Miller et al. 1987).

DGI has been described as a triplet, smear or single broad band in the region from ~150 to 165×10^3 M_0 on SDS–PAGE gels (e.g. see Gorsky and Steinberg, 1981; Cohen et al. 1983; Jones et al. 1988; Franke et al. 1981; Cowin and Garrod, 1983; Kapprell et al. 1990; Schwarz et al. 1990). In addition, immunological cross-reactivity among various species and even tissues within the same species has been reported to differ, and indeed DGI has been called a 'family' of polypeptides (Giudice et al. 1984; Samsom is proposed. Immunological evidence also indicates that the expression of some epitopes on DGI may change throughout the levels of a stratified epithelium (Parrish et al. 1986; Jones et al. 1987). Recent studies suggest, however, that DGI probably represents a single glycoprotein species that is an obligate constituent of all desmosomes (Schmelz et al. 1986a,b). Only by an analysis of the primary structure of DGI from various tissues will we be able to answer definitively the question regarding the actual number of DGI isoforms. The transmembranous nature of DGI has been demonstrated by immunolocalization of epitopes in the extracellular space as well as the cytoplasm (Steinberg et al. 1987; Miller et al. 1987; Schmelz et al. 1986a,b). Both DGI and the desmocollins have been shown to bind calcium and exhibit a calcium-sensitive resistance to proteolysis (Steinberg et al. 1987).

In order to gain a more clear understanding of DGI structure and function we have identified human cDNA clones from a λgt11 expression library and have carried out an analysis of the structure and regulation of this constitutively expressed desmosomal molecule. Our data suggest that, like the bovine DGI molecule, as reported recently by Koch et al. (1990), human DGI (HDGI) is highly related to the calcium-dependent class of CAMs known as cadherins (for review see Takeichi, 1988). N-terminal amino acid sequence and, more recently, cDNA sequence analysis have revealed a similar relationship between the desmocollins and cadherins (Holten et al. 1990; Collins et al. 1991), indicating the existence of a new subclass of calcium-dependent desmosomal cadherins. Comparison with the bovine DGI sequence has enabled us to identify regions of the molecule that have been evolutionarily conserved and may have functional significance. On the basis of structural analysis, a model predicting the disposition of HDGI domains in the desmosome is proposed. Finally, evidence is presented that the pattern of human DGI mRNAs is more complex than the pattern reported for the bovine gene. Two additional mRNAs are expressed in human cells and epidermis, and each of these cross-hybridizing mRNAs is coordinately expressed in normal human keratinocytes.

Materials and methods

Antisera

Antisera NW1 and NW3 were generated at Hazelton Research Labs (Denver, PA). Two rabbits were immunized with bovine DGI-polyacrylamide gel slices, boosted after three weeks, and subsequently boosted at two-week intervals. A third DGI antisera used to test phage clones for the production of antibodies was a generous gift from Dr Orest Blaschuk (Royal Victoria Hospital, Montreal).

Library screening and analysis of fusion proteins

The human foreskin Agt11 expression library was generated by Clontech (Palo Alto, CA) using random primers and kindly provided by Dr John Stanley (NIH). A total of approximately 500,000 plaques was screened initially with the NW1 antibody as described by Young and Davis (1983) and modified by Green et al. (1988). After plaque purification the fusion protein produced by each clone was tested for cross-reactivity with three affinity-purified anti-DGI antibodies as previously described (Green et al. 1988). Clones that tested positive were then subjected to epitope selection as described by Weinberger et al. (1985) and modified by Green et al. (1988, 1990). For subsequent screenings randomly primed ^32P-labeled DNA probes were used to isolate additional overlapping clones by standard methods (Sambrook et al. 1989).

DNA sequence analysis

cDNA inserts were subcloned into Bluescript vectors (Stratagene, La Jolla, CA). Both strands were sequenced entirely by Sanger's dideoxy chain termination method as described (Green et al. 1990) using modified bacteriophage T7 DNA polymerase (United States Biochemical Corp., Cleveland, OH). Plasmid DNA was purified from either mini or maxi preps using the alkaline lysis procedure (Sambrook et al. 1989).

Computer analysis

Sequence entry, translation, comparisons and searches for consensus sequences were performed using the PC Gene software package (Intelligenetics, Mountain View, CA). A search for homologous sequences was carried out using the protein identification resource (PIR) database with the UWGCN (University of Wisconsin Genetic Computer Group) software package (Devereux et al. 1984).

RNA and DNA blotting and hybridization

Total RNA was purified from human foreskin and normal human epidermal keratinocytes (NHEK cells) using the method of Chomczynski and Sacchi (1987). Northern and Southern blotting analyses were carried out as previously described (Green et al. 1990). Filters were exposed to Kodak X-Omat AR film in the presence of DuPont Cronex intensifying screens. Resulting autoradiograms were analyzed quantitatively by scanning densitometry using an LKB Ultrascan XL.

Isolation and culture of normal human foreskin keratinocytes

Normal human epidermal keratinocytes (NHEKs) were isolated from foreskins within 24–48 h after they were obtained from the nursery. Each was rinsed in povidone–iodine and 70 % ethanol, followed by 8 rinses in sterile PBS. Following removal of the dermis, tissue was minced into small pieces and incubated with stirring for 45 min at 37°C in 6 ml of 0.25 % trypsin per foreskin. The trypsin solution was then discarded, and 10 ml fresh 0.25 % trypsin solution was added. The tissue was incubated with stirring for another 30–45 min at 37°C. The trypsin solution was removed and the trypsin neutralized by adding 10 ml 10 % FCS–FAD medium (Angst et al. 1990). The cells were then pelleted by centrifugation at 200 g for 3 min and resuspended in 2 ml of KGM (Clonetics, San Diego, CA) containing 0.1 mm calcium. This process was repeated twice, and the cells were pooled. Cells harvested from one or two foreskins were seeded on a 60 mm culture dish and incubated at 37°C/5 % CO_2. The cells were fed every two days with fresh KGM containing 0.1 mm calcium. After 10–14 days in culture or after reaching 80 % confluence, the cells were passaged to 100 mm culture dishes at a seeding density of 4×10^5 per dish. Cells were usually used after five such passages.
Results

Isolation and characterization of cDNA clones encoding human DGI

A human foreskin keratinocyte Agt11 library was screened as previously described with a rabbit polyclonal antibody (NW1) directed against bovine desmoglein I (BDGI). Six positive clones were identified as a result of the primary screen of ~500,000 clones. Fusion proteins produced by plaque-purified phage stocks of each of these clones were tested further using three polyclonal affinity-purified antibodies (described in Materials and methods). Three of the six phage clones reacted strongly with all three of the antibody probes (data not shown). These fusion proteins were further tested by epitope selection for their ability to select DGI-specific antibodies from the complex polyclonal antisera (data not shown). Sequence analysis subsequently provided confirmation that these three clones were overlapping and related. Restriction fragments isolated from the 5′ end of the map were then used to rescreen the cDNA library, resulting in the identification of a series of overlapping clones shown in Fig. 2 (below). (Clones in Fig. 2 are designated by the numbering system used for subcloned plasmids, that is, a ‘p’ followed by a number.)

Northern analysis revealed the presence of multiple mRNAs in human foreskin epidermis and normal human cultured keratinocytes (Fig. 1A). The largest mRNA, 7.2 kb (kilobases), matches approximately that in bovine muzzle described by Koch et al. (1990); however, bands of ~4.0 kb and 3.0 kb were also observed. When six different restriction fragments from along the length of the cDNA were used as probes, all but an 800 bp (base-pairs) fragment within the 3′ untranslated region hybridized to all three bands. This probe did not hybridize to the smallest band of about 3.0 kb (Fig. 1A). In contrast, a single cross-hybridizing band of 5.0 kb was seen in Northern blots of the rat bladder carcinoma cell line 804G (data not shown). Variation in DGI mRNAs among tissues and species is consistent with observations by Koch et al. (1990) and Razsi et al. (1990).

In light of the apparent heterogeneity in DGI mRNAs, we were interested in determining the number of gene(s) encoding DGI. A Southern blot (Fig. 1B) containing human genomic DNA digested with a panel of restriction enzymes was hybridized with 32P-labeled insert from p63 (Fig. 1B). In each case, only one or two bands hybridized. Furthermore, a human DGI genomic clone has been isolated whose restriction patterns correspond to those shown in Fig. 1 (data not shown). These data are consistent with the possibility that there is one gene giving rise to the DGI mRNAs. By comparing the sizes of restriction fragments in the Southern blot with those within the cDNA, it is apparent that processing of the initial transcript is necessary to generate the final message.

Nucleotide and deduced amino acid sequence of human DGI: comparison with the cadherins and bovine desmoglein

The nucleotide sequence of the cDNA map shown in Fig. 2 was determined and then translated in all three reading frames (Fig. 3). Initially, the longest open reading frame in p63 was compared with the PIR database, revealing a significant sequence homology between residues 649 and 713 of HDGI with a restricted region of the cytoplasmic portion of the calcium-dependent cell adhesion molecule chicken N-cadherin (Hatta et al. 1988). Ultimately, the entire mature protein was analyzed using the PC Gene PROSITE program (Intelligenetics, Mountain View, CA), resulting in the identification of cadherin-like segments throughout the molecule. When the bovine DGI sequence recently became available (Koch et al. 1990), it was compared with the entire cDNA and deduced amino acid sequence shown in Fig. 3. The sequences were clearly highly related with an amino acid sequence identity of 82% and similarity of 85%. This finding provided...
Fig. 2. Human DGI cDNAs and sequencing strategy. The human DGI cDNA restriction map is shown schematically. Below the map are the overlapping plasmid subclones used for sequence analysis. Arrows indicate the length and direction of sequence obtained from each primer. The entire map was sequenced completely on both strands. Within the 4.2 kb shown here is encoded the complete mature DGI polypeptide and 24 residues of the processed amino terminus found in the preprotein, represented diagrammatically immediately above the restriction map. p63, p65 and p64 are subclones derived from the inserts of the three original immunopositive phage clones. The cross-hatched rectangle represents the probe within the 3' untranslated region used in Fig. 1a, lane 4.

additional confirmation that the cDNA clones identified in our laboratory indeed encoded human DGI. The comparison also suggested that the entire mature DGI polypeptide is encoded within the 4224 nucleotides shown in Fig. 3. However, like the reported bovine sequence, the HDGI sequence in Fig. 3 lacks a portion of the processed amino terminus. At the 3' end of the mRNA are 1152 nucleotides of untranslated sequence that do not appear to contain a potential polyadenylation site. Since BDGI contains a long 4174-nucleotide untranslated tail, it is expected that polyadenylation site(s) are contained within the 3' portion of the HDGI mRNA for which cDNAs have not yet been identified.

By aligning the HDGI sequence with the BDGI and cadherin sequences, it seems likely that the mature protein begins with E-W-I-K (Koch et al. 1990). These residues are preceded by the final residues of the preprotein, I-R-R-Q-K-R. This motif, particularly the sequence Q-K-R, is also highly conserved in the cadherins. It follows then that the mature, unglycosylated HDG polypeptide is precisely 1000 amino acids with a relative molecular mass of 107703.

The cadherin and desmoglein molecules can all be divided into extracellular, transmembrane and cytoplasmic domains. The extracellular and cytoplasmic domains can each be subdivided further on the basis of sequence comparisons and structural considerations. By identifying regions of highest homology between the desmogleins and cadherins in general, and then between the bovine and human homologs of desmoglein, regions of functional importance might be determined. Fig. 4 compares the domain structure of human N-cadherin (HNCAD; Walsh et al. 1990) with the two desmogleins studied to date.

The extracellular domain can be divided into a total of five regions. The first four have been designated E I–E IV (as in BDGI), each containing approximately 112 residues. Although these regions display limited homology to one another even this is absent in the latter part of E IV. Each of the first three E domains of DGI (and to an extent the fourth domain) consists of five quasi-repeats ~20–22 residues in length. These repeats are manifest in the secondary structure adopted rather than in the sequence per se. Structural analysis suggests that these repeats may comprise alternating pieces of α-helix and β/β-turn structure (e.g. as in residues 1–21 of E I where 1–11 are predicted α-helical and 12–21 are predicted β/β-turn). The degree of similarity in the E I and E II domains of HNCAD and HDGI is ~56%, the highest value of all the domains. Likewise, these two extracellular domains are the regions most highly conserved between the bovine and human homologs of DGI. In spite of this overall similarity in the E I and E II domains, there are significant differences between DGI and the cadherins in regions that have been demonstrated to be important for adhesive specificity (Nose et al. 1990). For instance, although the putative adhesion motif H-A-V (Blaschuk et al. 1990a,b) is replaced by a partially conserved sequence R-A-L (residues 79–81 in HDGI) (Goodwin et al. 1990), the residues immediately surrounding it that have been demonstrated to be important for adhesive specificity are not conserved.

Cadherins (Takeichi, 1988) and DGI (Steinberg et al. 1987) both bind calcium. Specific motifs in the cadherins
have been proposed to be involved in this binding (Ringwald et al. 1997). Six such motifs are located at similar positions in all of the cadherins studied so far, and are thought to lie within hydrophilic loops on the surface of the molecule. Similar sequences have been identified in DGI (underlined in Fig. 3); however, their role in calcium-binding and adhesion has not been demonstrated.

The region immediately amino-terminal of the transmembrane spanning domain (47 residues in HDG, 53 in BDG and 107 in HDNCAD) is designated EA or 'extraellular anchor'. This region is the least highly conserved of all the domains, retaining only 23 % identity/34 % similarity when compared with BDG and 6 % identity/12 % similarity when compared with HNCAD. The difference in length in this domain accounts in large part for the smaller size of the extracellular region of the DGs.

There are four consensus N-glycosylation sites in the mature HDGI as determined by the PC Gene PROSITE program, two of which are located in the extracellular domain (Fig. 3) and are conserved in the bovine homolog. None of these sites is found in the caderhins, all of which have glycosylation sites located at different positions.

The transmembrane spanning domain is 24 amino acids in both the bovine and the human proteins as compared with 32 residues in HNCAD. This domain is very similar in the DG homologs, with only two conservative amino acid changes. As in the cadherins, DG contains a single stretch of hydrophobic amino acids, indicating that the molecule is likely to span the membrane just once.

The cytoplasmic domain of both DGs is substantially larger than that of the cadherins: HNCAD contains 160 residues, whereas BDG contains 471 and HDG contains 480. (It should be noted that BDGI is 35 residues longer than previously reported by Koch et al. (1990), (W. W. Franke, personal communication)). The difference in size between the DGs and the caderhins can be accounted for by the presence of several additional domains at the carboxyl terminus of DGs. Koch et al. (1990) have divided the cytoplasmic domain of BDG into five subdomains based on homology with the cadherins. Using these criteria, a comparison with HDG reveals similar regions, which are designated C I through C IV in Fig. 4. Proximal to the transmembrane spanning domain is a region (residues 521-594) that can be considered an "intracellular anchor" (IA). Although this is equivalent in domain is present in cadherins and DGI, it is larger in DGI. In the cadherins, the first three residues of this domain are basic, while the IA domain of both human and bovine desmogleins is proline- and cysteine-rich. Residues 526-559 of this domain contain two copies of a 17 amino acid repeat in which 7 of 17 residues are identical.

The C I domain (comprising residues 595-717 and termed ICS by Koch et al. (1990)) bears 30 % identity/44 % similarity to HNCAD. Likewise, this highly charged region exhibits 94 % identity to bovine DGI. Thus, the most highly conserved domains within the two DGs so far sequenced are the two most amino-terminal 'E' domains and the cadherin-like region, which corresponds to the cytoplasmic terminus of the cadherins.

However, both the bovine and human DGs contain additional carboxyl-terminal sequences. The bovine molecule has an additional 311 residues, and the human molecule an additional 232 residues compared with HNCAD. This additional sequence can be divided into three domains. The first of these is a proline-rich domain of 59 amino acid residues (718-776) (C II or IPL in BDGI). On the basis of residue content, the boundaries of the three cytoplasmic domains discussed so far, IA, C I and C II, can be defined in an alternative way. Residues 521-569 of the intracellular anchor are quite high in proline content, 18.4 % in HDG. Residues 750-791 on the other hand, contain 27.3 % charged residues and only 2.7 % proline. The remaining proline-rich domain contains 24.6 % prolines. Thus, proline-rich regions each about 50 residues in length flank a strongly acidic block of residues predicted to adopt a p-turn/coil conformation.

The C III region (residues 777-907) is characterized by a novel series of repeat containing 291 ± 1 amino acids that has not been reported in any protein other than the bovine homolog (Fig. 5). The boundaries of C II and C III were defined differently by Koch et al. (1990), who considered the beginning of what they called the RUD region to be the equivalent of residue 790 in HDGI. Therefore, the bovine RUD does not include a partial repeat at the beginning of the domain. Although C III is 89 % identical to the BDG repeat region, the conservation in this domain is not quite as high as that of the cadherin-like region. Each of the complete highly homologous repeating units is likely to consist of two p-strands and two turns as shown diagrammatically in Fig. 5. The first p-strand has one apolar and one variable face. The second strand also has an apolar face although a single charged residue is present. The other face of this strand is also variable in character. Interestingly, there are five amino acids that are identical in all the repeats of both the human and bovine DGIs (Fig. 5). These conserved amino acids lie in the second of the two predicted p-strands. It is intriguing to speculate that they play a role in some conserved function, such as interacting with other desmosomal plaque proteins. The T-E-R motif found in the majority of the repeats (and which contains two of the five perfectly conserved residues just noted) was identified by the PROSITE program as a possible target for phosphorylation by protein kinase C.

The sequence of the final carboxy-terminal domain, C IV (residues 908-1000), diverges significantly in the bovine and human homologs. HDGI is 12 amino acids longer than BDGI and only 52 % identical in this domain. Both molecules are very glycine-rich in this region. Residues 908-964 in HDGI contain ~2 copies of a 28 ± 1 residue repeat, not related to that found in C III. The first of these repeats comprises 8 G-X-Y motifs not found in BDGI; although similar motifs are found also in collagen, X and Y are not typical of residues seen in collagen. The last 35 residues of the C IV domain (designated C IVb in Figs 4 and 7) have a distinctive charged character, with a 5:1 ratio of basic to acidic residues. It is important to note that IA, C I, C II and C III are all very acidic whereas C IV is apolar with a basic tail.

Regulation of human DGI in cultured human keratinocytes

In order to gain a more complete appreciation of desmososome assembly, particularly with regard to the regulated expression of desmosomal constituents, we compared DGI and DFI/II mRNA levels during calcium-induced desmosome assembly in NHEK. In a previous study DFI/II mRNA and protein levels were found to be comparable up to 24 h after human FaDu (pharyngeal) cells were switched to normal calcium levels (Angst et al. 1990). We wondered whether such an increase might
Fig. 3. Nucleotide and predicted amino acid sequence of human DGI. Included within the sequences are the entire mature molecule of 1000 residues, 24 residues of the processed amino terminus and a portion of the untranslated 3' end. The boundaries of the five amino-terminal extracellular (E) domains, the transmembrane spanning domain (M) and the five carboxyl-terminal cytoplasmic (C) domains are designated above the sequence. The six putative calcium binding domains originally described in the cadherins (Ringwald et al. 1987) are underlined. Residues 79-81 (R-A-L), corresponding to the cadherin H-A-V motif, are indicated in bold type. Two potential N-glycosylation sites at residues 61 and 131 are also marked (••).

Fig. 4. Comparison of the domain structure and conservation between human DGI (HDGI) and human N-cadherin (HNCAD), and between HDGI and bovine DGI (BDGI). The orientation of each molecule is shown with respect to the plasma membrane (M). The boundaries of the extracellular domains, E I, E II, E III, E IV and EA (extracellular anchor), and of the cytoplasmic domains, IA (intracellular anchor), C I, C II, C III and C IVa/b, are shown for each molecule. Putative calcium binding domains in the amino terminus are shown as shaded rectangles. Note that both DGI molecules contain an additional three domains at the carboxyl terminus. Percentage identity and similarity are shown. Residues counted as similar were: S/T, R/K, D/E and A/I/L/M/F/Y/V.
have been due to factors arising as a result of long-term culture rather than being strictly calcium dependent.

To investigate this possibility we compared the relative DGI and DPI/II mRNA levels in NHEK cells maintained in low calcium for a 3 day period with levels in cells cultured in normal calcium. NHEK were grown to 80% confluence and then switched to normal (1.3 mM) calcium KGM or maintained in low calcium (0.1 mM) for 1, 2 and 3 days. Total RNA was isolated and Northern blot analysis carried out using GAPDH (glyceraldehyde-3-phosphate dehydrogenase) as an internal loading standard. Both DPI/II and DGI mRNAs were found to exhibit a significant increase in accumulated levels in low and normal calcium over a 3 day period in culture (Fig. 6). In low calcium, DPI/II levels increased ~2- to 4-fold between 1 and 3 days, whereas in normal calcium DPI/II increased ~6- to 7-fold. Interestingly, all three DGI bands exhibited a greater increase than the DPs in both low and normal calcium, with calcium producing an ~2-fold higher relative increase. The three DGI mRNAs increased approximately 13- (7.2 kb mRNA), 28- (4.0 kb mRNA) and 15-fold (3.0 kb mRNA) between day 1 and 3 in normal calcium.

**Discussion**

An adhesive role for desmosomes has been implicated in the formation of the early blastocyst stage embryo where they first appear adjoining the cells of the developing trophectoderm (Ducibella et al. 1975). Desmosomes are also thought to play an important role in later development during the movement of cell sheets (Lentz and Trinkaus, 1971). In the adult, the relatively high abundance of desmosomes in skin and other tissues experiencing mechanical stress has suggested to investigators their importance in maintaining the integrity of epithelial cell sheets (Arnn and Staehelin, 1981). Evidence that specific desmosomal cell adhesion molecules are involved in these processes is limited. The desmogleins (including DGI and desmocollins I and II) have long been thought to be involved in the adhesive aspect of desmosome function (e.g. see Gorbsky and Steinberg, 1981; Garrod, 1986). These molecules are all calcium-binding transmembrane glycoproteins specifically localized in desmosomes. Experiments by Cowin et al. (1984) suggested that desmosome-dependent adhesion could be prevented by incubating cells with Fab' fragments of antibodies directed against the desmocollins. The structural analysis presented in this paper, as well as those by Holton et al. (1990), Koch et al. (1990), Cowin et al. (1990) and Collins et al. (1991), now provide strong support for an evolutionary and perhaps functional relationship between the desmogleins and the calcium-dependent class of adhesion molecules known as cadherins. In addition, the availability of cDNA clones

![Diagram of β-turn structure](image-url)
RNAs were isolated from cultured keratinocytes on days 1, 2 and 3 following a switch to 1.3 mM calcium (NCa), or from keratinocytes maintained in 0.1 mM calcium (LCa). Three Northern blots were hybridized with DP, DGI and GAPDH cDNA probes, and mRNA levels were quantitated by scanning densitometry. mRNA sizes in Mr\texttimes10^3 are shown at the left.

![Fig. 6. Expression of DPI/II and DGI in NHEK cultures. Total RNAs were isolated from cultured keratinocytes on days 1, 2 and 3 following a switch to 1.3 mM calcium (NCa), or from keratinocytes maintained in 0.1 mM calcium (LCa) at the same time points. Northern blots were hybridized with DP, DGI and GAPDH cDNA probes, and mRNA levels were quantitated by scanning densitometry. mRNA sizes in Mr\texttimes10^3 are shown at the left.](image)

Now provides an opportunity for testing such proposed functions in living cells and tissues.

Unlike the bovine homolog, human DGI probes reveal the presence of three cross-hybridizing species in Northern blots of human keratinocytes and of human foreskin. mRNAs of different molecular weights and numbers are seen in other cell lines, such as the rat bladder 804G cell line. Similar differences in expression have been reported using bovine cDNAs as probes (Koch et al. 1990; Razai et al. 1990). Southern blot analysis indicates that it is likely that DGI is encoded by a single gene. Therefore, the multiple bands seen in Figs 1 and 6 may represent alternatively spliced variants and/or products of multiple transcriptional start sites or polyadenylation sites. Northern mapping analysis (Fig. 1) is consistent with the possibility that the entire coding region is contained within the 3.0 kb transcript and that the 7.2 and 4.0 kb mRNAs represent alternate forms that contain additional 3' and possibly 5' untranslated sequences. This can be confirmed when clones encoding the end portions become available.

Although tissue-specific differences in the level of DPI/II mRNAs have been reported previously (Angst et al. 1990; Schwarz et al. 1989; Green et al. 1991), to date no evidence has been presented that the expression of desmosomal molecules may be modulated. Here we show that both DPI/II and DGI mRNAs increase in accumulated mRNA levels over a 3 day period in culture whether maintained in low calcium or switched to normal calcium levels. Whether these changes are due to transcriptional or post-transcriptional regulation has not been determined. Calcium is not the only factor inducing changes in DPI/II and DGI mRNA levels. Other factors related to long-term culture such as an increase in cell density, detachment from the substratum during stratification, or the release of autocrine regulatory factors (Watt, 1989), may contribute even more significantly to changes in accumulated mRNA levels of these molecules. It is important to note here that unlike murine keratinocytes, NHEK cells continue to proliferate in KGM medium containing 1.3 mM calcium (Clonetics Corporation, 1989), thus leading to increases in cell density and release of cells from the substratum. These results are in line with observed increases in desmosomes seen in stratified foci of some tumor cells maintained in normal levels of calcium that do not form desmosomes in monolayer cultures (Green et al. 1991). More importantly, the conditions inducing changes in mRNA levels in culture are similar in many respects to those occurring during normal differentiation of a stratified epithelium and could account for the increases in desmosome number observed in upper layers of such an epithelium (White and Gohari, 1984).

DGI exhibits a particularly striking increase in accumulated mRNA levels, especially in the presence of calcium, with levels increasing up to ~28-fold for the 4.0 kb DGI message. In a previous report (Angst et al. 1990), an increase in accumulated mRNA levels for DPI/II over a 5 day period after the switch to normal calcium was not ruled out. However, the effect of calcium was not appreciated at that time, possibly because the increase in accumulated DP mRNA levels was observed at only one, namely the 5 day, timepoint and may have been an effect solely of long-term culture. mRNA levels clearly do not appear to change over the short term following the calcium switch in any of the cells we have studied (Angst et al. 1990, and unpublished observations). Because normal keratinocytes are likely to be more sensitive to increased extracellular calcium than the tumor cell line (Kulesza et al. 1983), the calcium effect on expression of desmosomal components may be more pronounced in NHEK than FaDu cells.

Our analysis of HDGI demonstrates a high degree of conservation with BDGI, with an overall amino acid sequence identity of 82%. The major differences between the two molecules lie in the regions flanking the transmembrane-spanning domain and the domain at the carboxyl terminus. It is likely that the former regions act primarily to anchor DGI in the membrane, as suggested previously by Koch et al. (1990). Likewise, the basic tail of the carboxyl terminus (C IVb in Fig. 4) may act as an anchor, preventing this region from being pulled through its neighboring non-charged segment (C IVa). A similar feature has been seen in M-protein, the surface spike transmembrane protein of Streptococcus, and other related proteins (Fischetti et al. 1988). As with BDG, the human molecule can be divided into a series of extra- and intracellular domains. The most highly conserved of these include the first two extracellu-

Human desmoglein structure and expression 817
lar domains, E I and E II, and the cytoplasmic cadherin-like region. The E domains have been shown to be of critical importance for mediating homophilic adhesion between cadherin molecules, as well as for determining the specificity of adhesion among the different classes of cadherins (E-, N-, P) (Nose et al. 1990; Blaschuk et al. 1990a,b). Furthermore, a peptide containing the second proposed calcium binding domain of uvomorulin (found in domain E II) has recently been shown to bind calcium, providing the first direct evidence that these motifs actually function as proposed (Ozawa et al. 1990a).

Expression of uvomorulin molecules in which aspartic acid residue 134, found within this domain, was altered by site-specific mutagenesis resulted in the abrogation of cell adhesion in L cells. Interestingly, this residue is identical in both BDGI and HDGI. On the other hand, residues demonstrated to be involved in determining the specificity of adhesion are not conserved (Nose et al. 1990). This supports the possibility, suggested by Goodwin et al. (1990a), that different residues in the desmosomal adhesion molecules are responsible for the segregation of DGs into specific domains on the cell surface.

The first intracellular domain, C I, is also highly conserved among the cadherins and desmogleins. This domain has been proposed to play a role in the interaction of cadherins with the microfilament-containing cytoskeleton (Ozawa et al. 1990b). Uvomorulin/E-cadherin (Boiler et al. 1985) and A-CAM (N-cadherin) (Volk and Geiger, 1986) have been localized to intercellular adherens junctions. It has been proposed that three polypeptides called catenins may be involved in linking cell adhesion molecules to the actin-containing cytoskeleton (Ozawa et al. 1989, 1990b). By constructing chimeric molecules containing the final 72 residues of uvomorulin, Ozawa et al. (1990b) have recently provided convincing evidence that this domain is responsible for catenin binding. The similarity between DG and cadherins in this region poses a problem in thinking about the assembly of adherens junctions and desmosomes in the same cell. If this domain is involved in linking cadherins to the microfilament cytoskeleton, then why aren't microfilaments also assembled onto desmosomal plaques? There are a number of possible reasons for the specific attachment of cytoskeletal elements to these junction types, one being the existence of the extra carboxy-terminal portion in the DG molecule. This region may act as a specific binding site for desmosomal plaque molecules and in addition might modify the ability of the cadherin-like domain to bind to catenins. Likewise, catenin binding might be prevented by conformational differences in the structure of the cadherin-like domain resulting from post-translational modifications such as phosphorylation or cross-linking of cysteine residues.

The question remains, however, as to why the C I domain is so highly conserved in cadherins and desmogleins. One possibility is that this region also provides a binding site for plakoglobin, a component of both adherens junctions and desmosomes. Using immunoprecipitation, this molecule has been reported to form complexes with DGI (Korman et al. 1989). It should be noted that another possible complicating factor is the report by Jones (1988) that desmosomes also contain a $125 \times 10^3$ M$^+$ glycoprotein that reacts with antibodies directed against E-cadherin (uvomorulin).

Unlike adherens junctions, the cytoplasmic plaque of desmosomes acts as a specific binding site for intermediate filaments and contains the unique plaque proteins, DPI and II. The presence of the additional carboxyl tail containing a novel repeat domain may play a role in the specific attachment of IF to desmosomes via plaque proteins such as the DPs. Five residues in the repeat regions are identical in all of the repeats in both the bovine and human forms. These are clear candidates for involvement in binding or regulation of binding to other plaque components. Interestingly, among these highly conserved residues is a threonine-containing motif that might act as a phosphorylation substrate for specific kinases such as PKC. (It should be noted that these putative sites may not be in the optimal context (Kemp and Pearson, 1990)).

Phosphorylation has been suggested by several investigators as a possible mechanism for regulation of desmosome assembly (e.g. see Garrod et al. 1990; Schwartz et al. 1990).

As noted by Koch et al. (1990), the Gly-Ser-rich portion of the carboxyl terminus is similar to the end domains of the epidermal keratin molecules. In fact, when we performed a comparison of the HDGI C IV domain from residues 908–994 with an equivalent length piece from the carboxyl terminus of mouse epidermal keratin K2, a 32% identity was revealed, mainly in the Gly and Ser residues. It could be argued that these regions of high predicted flexibility might then interact with one another, providing a mechanism for keratin intermediate filament (IF)–desmosome association. However, since vimentin-containing IF interact with desmosomes in the arachnoidal cells of the brain and desmin-containing IF interact with desmosomes in cardiac myocytes, other types of interactions must mediate an IF–desmosome association in these cells. In a previous report (Green et al. 1990) we suggested that the carboxyl terminus of DPI/II and the 1B rod domain of intermediate filaments might interact on the basis of a similar periodicity of 9.5 in their charged residues. Perhaps the reported preferential association of keratin IF with desmosomes in cells containing both types of filament (Kartenbeck et al. 1984) is due to stabilizing influences resulting from multiple types of IF interactions with desmosomal molecules.

On the basis of structural analysis, we have generated a model predicting the disposition of DGI domains in the desmosome (Fig. 7). Because there is some disagreement in the literature on the values for the width of each portion of the plaque, particularly in the so-called 'satellite' region (Garrod et al. 1990), we carried out our own measurements of desmosomes in conventionally fixed bovine tongue (Fig. 7). The values we generated are in general agreement with those of others (Miller et al. 1987; Steinberg et al. 1987; Garrod et al. 1990). The portion of the satellite region designated 'inner dense plaque' by Steinberg et al. (1987) is perhaps the most difficult portion of the plaque to define. The value of 15–19 nm shown here falls somewhere between those suggested by Garrod et al. (1990) and Steinberg et al. (1987). The value of 3.5 nm assigned to the plasma membrane corresponds to the width determined by the commonly observed 23 or 24 α-helical residues that span it, i.e. 23 nm×0.1485 nm rather than the value measured in electron micrographs, which is somewhat larger.

The general orientation of BDGI has been defined by Koch et al. (1990); that is, the carboxyl terminus of the molecule is located within the cytoplasm, and the amino terminus comprises the extracellular domain. This prediction is consistent with the reactivity of specific antibodies to cytoplasmic versus extracellular domains. Our model predicts that the first three cytoplasmic domains, IA, C I
Fig. 7. Model of the disposition of two DGI molecules within a schematic representation of a desmosome. Approximate values for the width of each layer are given in nanometers. On one half of the desmosome, layers are designated as: CDS (central dense stratum), ICS (intercellular space), M (plasma membrane), ODP (outer dense plaque), IDP (inner dense plaque). Intermediate filaments (IF) are shown looping through the IDP. The approximate locations of plakoglobin (PG) and desmoplakins (DP) are indicated by brackets. On the other half of the desmosome, the DGI domains are designated using terminology from Figs 3 and 4. The E I and E II domains are shown as overlapping here, implying a possible role in homophilic cell adhesion. However, it should be noted that there is no experimental evidence to support this idea.

and C II, lie within the ~15 nm outer dense plaque (ODP). These domains have a high degree of predicted symmetry, with proline-rich segments flanking a negatively charged block of residues. The next segment, C III, is very different in character, and it also contains a high density of putative phosphorylation sites, which might best be situated out of regions of dense protein packing. Thus, it is predicted that this domain will span the ~8 nm region between the outer and inner dense plaques (Steinberg et al. 1987), that is the portion of the 'satellite' region nearest the plasma membrane (Miller et al. 1987; Garrod et al. 1990). It seems most likely that the apolar domain and the basic tail comprising the C IV domain extend outward through the IDP. Since IF impinge upon this region, this localization for C IV might be consistent with the hypothesis suggested above, that the Gly-Ser-rich regions of C IV might be involved in IF binding. The final 35 residues of HDGI (C IVb) might then anchor the molecule in place at the junction of this space and the IDP.

Immunolocalization data presented by Miller et al. (1987) suggest that most of the DGI reactivity in the cytoplasm is restricted to the ODP, although Steinberg and colleagues (1987) suggested that some reactivity extends into the satellite region. Interestingly, plakoglobin also seems to be restricted to the ODP, the region where C I, the cadherin-like domain, is predicted to lie. DPI/II, on the other hand, lies predominantly within the satellite region.

Clearly this model is highly speculative and awaits testing. However, the availability of human cDNAs now provides an opportunity of addressing questions regarding not only the disposition of various domains within the desmosome, but (1) the mechanism of segregation into specific intercellular junction domains during assembly; (2) the mechanism of specific cell adhesion to other desmosomal CAMs; (3) the role of calcium in desmosomal adhesion; (4) the identity of desmoglein-associated proteins in the plaque; and (5) the mechanisms by which these associated proteins come together to form functional complexes involved in the binding and organization of intermediate filaments.

The authors thank Dr John Stanley for the human keratinocyte agt11 library and Dr Orest Blaschuk for one of the DGI antibodies. Thanks go also to Dr Rex Chisholm for his generous help with data base searches and comparisons, and for many useful discussions.

This work was supported by grants to K.J.G. from the Illinois (#90-27) and National (#BE-56) American Cancer Society and grant #2432 from The Council for Tobacco Research USA, Inc. K.J.G. is also supported by NIH HD24430, an American Cancer Society Junior Faculty Research Award, and a March of Dimes Basil O'Connor Starter Scholar Research Award.

References


Human desmoglein structure and expression


(Received 8 March 1991 – Accepted 2 May 1991)