Absence of cell-surface EpCAM in congenital tufting enteropathy

Ulrike Schnell¹,†, Jeroen Kuipers¹,†, James L. Mueller², Anneke Veenstra-Algra¹, Mamata Sivagnanam³ and Ben N.G. Giepmans¹,*

¹Department of Cell Biology, University Medical Center Groningen, University of Groningen, PO Box 196, 9700 AD, Groningen, The Netherlands, ²Division of Allergy, Immunology, and Rheumatology, Department of Pediatrics, University of California San Diego, Ludwig Institute of Cancer Research, San Diego Branch, La Jolla, CA, USA and ³Division of Gastroenterology, Hepatology, and Nutrition, Department of Pediatrics, University of California San Diego, La Jolla; Rady Children’s Hospital, San Diego, CA, USA

Received December 3, 2012; Revised and Accepted February 27, 2013

Mutations in the epithelial cell adhesion molecule (EpCAM; CD326) gene are causal for congenital tufting enteropathy (CTE), a disease characterized by intestinal abnormalities resulting in lethal diarrhea in newborns. Why the different mutations all lead to the same disease is not clear. Here, we report that most mutations, including a novel intronic variant, will result in lack of EpCAM’s transmembrane domain, whereas two mutations allow transmembrane localization. We find that these mutants are not routed to the plasma membrane, and that truncated mutants are secreted or degraded. Thus, all epcam mutations lead to loss of cell-surface EpCAM, resulting in CTE.

INTRODUCTION

Congenital tufting enteropathy (CTE) is a rare form of intractable diarrhea that develops within days after birth (1). The disease is characterized by morphological abnormalities in the intestinal epithelium, including subtotal villous atrophy with crypt hyperplasia and focal crowding of surface enterocytes, resembling tufts (2). Most patients are dependent on parenteral nutrition to allow normal growth and development. The disease persists throughout life and imparts significant morbidity and mortality (3). CTE was first described in 1994 (1) with an annual incidence of approximately 1 out of 100 000 live births in western Europe. As a result of its severity and rarity, the pathogenesis of CTE remains poorly understood.

Mutations in the gene encoding the epithelial cell adhesion molecule (EpCAM; CD326) have been discovered as the genetic basis of CTE (Fig. 1A) (3–7). To date, eight distinct epcam mutations have been identified, resulting in a single amino acid exchange, premature truncation or partial deletion of the EpCAM protein (3–7). The transmembrane cell-surface protein EpCAM (Fig. 1B) has originally been discovered as one of the first carcinoma markers (8). EpCAM has been shown to promote tumor growth, but also to play a morphoregulatory role during development (9–11). In the intestinal epithelium, EpCAM is localized at the basolateral membrane of cells, and a gradient in the EpCAM expression level can be observed from villi (low) to crypts (high) (reviewed in 9). EpCAM has proposed functions in cell–cell adhesion (reviewed in 9), as well as in receptor-like mediated control of cell proliferation (12). Although it remains to be explored how these EpCAM functions are separated or connected, EpCAM’s function at the cell surface is not debated (9,12).

Interestingly, in a recent report, the absence of EpCAM in a knockout mouse was found to result in a phenotype that is similar to CTE (13). Here, we set out to uncover why the different epcam mutations lead to the same disease. We found that all CTE-related EpCAM variants prevent the protein from being located at the plasma membrane, revealing why the different mutations lead to the same disease. In addition, a previously unidentified intronic variant in the only supposed heterozygous CTE patient confirms the recessive...
### Table A

<table>
<thead>
<tr>
<th>No.</th>
<th>Coding DNA</th>
<th>Protein</th>
<th>Zygosity</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>c.491+1G&gt;A</td>
<td>Exon 4i (a)</td>
<td>homozygous</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>c.426-1G&gt;A*</td>
<td>Exon 4i (b)</td>
<td>homozygous</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>c.556-14A&gt;G</td>
<td>Intron 5</td>
<td>homozygous</td>
<td>new</td>
</tr>
<tr>
<td>4</td>
<td>c.197G&gt;A*</td>
<td>Exon 5</td>
<td>compound heterozygous</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>c.412C&gt;T</td>
<td>Exon 5</td>
<td>homozygous</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>c.498insC</td>
<td>Exon 5</td>
<td>homozygous</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>c.491+1G&gt;A</td>
<td>see #1</td>
<td>compound heterozygous</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>c.316A&gt;T</td>
<td>Exon 5</td>
<td>p.K106X</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>c.492-2A&gt;G</td>
<td>Exon 5i (b)</td>
<td>homozygous</td>
<td>7</td>
</tr>
<tr>
<td>10</td>
<td>c.498insC</td>
<td>see #6</td>
<td>compound heterozygous</td>
<td>7</td>
</tr>
</tbody>
</table>

### Diagram B
- **N-terminus**
- **Extracellular**
- **Intracellular C-terminus**

### Diagram C

#### Wt
- **human EpCAM**
- **Structural changes**
  - W143_T164del
  - C66Y
- **Premature truncations**
  - K106X
  - R138X
  - Q167Psfx21
  - A165Mfsx24
  - Y186Fsfx6
RESULTS

All identified CTE mutations have been found to be homozygous or compound heterozygous (Fig. 1A), consistent with autosomal recessive inheritance (3,6). Notably, C66Y had been found as the only heterozygous mutation (3). Here, we report that further genetic screening of this patient identified an additional intronic variation (c.556-14A>G), which predictably leads to a frameshift and premature truncation within EpCAM’s ectodomain (p.Y186FfsX6; Fig. 1; Supplementary Material, Fig. S1). Although most of the EpCAM mutations found in CTE lead to premature STOP codons, resulting in the loss of the transmembrane domain, two mutations (W143_T164del and C66Y) result in other structural changes, retaining the transmembrane domain (Fig. 1C). To identify the molecular basis of CTE, all mutants were mimicked at the cDNA level and expressed in HEK293T cells followed by the analysis of subcellular localization (Fig. 2A–C). Whereas wtEpCAM is predominantly localized at the plasma membrane, both the deletion mutant (W143_T164del) and the point mutant (C66Y) are retained in the endoplasmic reticulum (ER). Experiments using Golgi markers versus ER markers in living cells further showed the preferential accumulation in the ER (Supplementary Material, Fig. S2A and B), with similar localization of the deletion mutant and the C66Y point mutant (Supplementary Material, Fig. S2C). The ER localization was shown even more strikingly when the mutant proteins were expressed with an ER marker protein and immunostained for giantin, a Golgi-resident enzyme (Supplementary Material, Fig. S3). In these two mutant proteins, the level of EpCAM multimers is increased (Fig. 2D). All EpCAM mutants with premature STOP codons are minimally detectable (Fig. 2A). Only in a subset of cells, the mutant proteins are present in the ER (Fig. 2A). FACS analysis shows that none of the mutants is present at the cell surface (Fig. 2B, Supplementary Material, Table S1). Western blot analysis of cell lysates and concentrated cell-free medium reveals that all mutant proteins are expressed (Fig. 2C), and that mutants R138X and K106X are secreted. Note that a weak band shows that also Y186FfsX6 is secreted to a lesser extent. Inhibition of protein transport with nocodazole increases the levels of truncated mutants in cell lysates (Fig. 2C, middle panel). Notably, both wtEpCAM and its cleaved ectodomain are present in cell-free medium due to the presence of the protein on microvesicles and ectodomain shedding (Fig. 2C, bottom).

DISCUSSION

Although the CTE mutations in epcam lead to different gene products, all mutations cause loss of cell-surface EpCAM, thus explaining the similar phenotype. ER retention of the deletion mutant and C66Y might be caused by misfolding, resulting in increased multimerization and aggregation. Loss of the cys66, which is normally engaged in a disulfide bond (14), leads to free cys99, possibly resulting in intermolecular dimerization. A similar model of intermolecular aggregation has been suggested to occur in gelatinous drop-like corneal dystrophy, involving a C119Y mutation in EpCAM’s homolog TROP-2 (15).

EpCAM mutants that lack the transmembrane domain are either secreted or seem to be degraded. Notably, only the mutants with unchanged sequence prior to the STOP codon are secreted (K106X and R138X), whereas mutants with frame shifts, resulting in regions of novel amino acids, are not detectable in the medium and hardly detectable in cell lysates. These mutants may be recognized as targets for ER-associated protein degradation (ERAD), leading to ubiquitination and fast degradation by the proteasome (16). Mutant Y186FfsX6, which only has six novel amino acids added, is secreted at very low levels in comparison with the other secreted mutants. Thus, the relatively long regions of novel amino acids in CTE mutants Q167PfsX21 and A165MfsX24 might play a role in the substrate recognition for ERAD. Considering EpCAM’s morphoregulatory role and its high expression in the intestinal crypts (reviewed in 9) where the stem cells are located (17), the absence of EpCAM might disturb the balance between cell proliferation/stem cell renewal and differentiation and thereby the normal development of the intestinal epithelium (3). EpCAM’s absence from the cell surface may not only preclude its adhesive function, but also EpCAM-mediated signaling, which involves proteolytic cleavage (12,18) and association with tetraspansins and other transmembrane proteins in the plasma membrane (19,20). Loss of EpCAM in a knockout mouse model indeed affects other cell adhesion molecules, notably E-cadherin and β-catenin, as recently reported by Guerra et al. (13), which is in line with our findings on the absence of EpCAM from the cell surface due to mutations in patients.

In conclusion, we show that CTE is a truly autosomal recessive disease, characterized by mutations in epcam. All mutations identified lead to the absence of EpCAM from the cell surface, impairing normal development of the intestinal epithelium. Revealing the effects of epcam mutations might help in the development of CTE therapies, i.e. the prevention of protein aggregation to reroute EpCAM to the cell surface. In addition, these mutations may be instrumental in understanding EpCAM’s normal role in regulating cell signaling and/or adhesion.

Figure 1. (A) epcam mutations linked to CTE. Nucleotide and amino acid coordinates are given (The “A” of the start codon is 1). a: donor splice site; b: acceptor splice site; single asterisks denote coordinates deviating from reference; double asterisks denote predictions using the Alamut software. (B) Schematic representation of EpCAM. Gray: signal peptide; green: EpCAM motif-1; purple: thyroglobulin type 1A-like repeat; black: transmembrane domain. (C) EpCAM mutant proteins in CTE. Colors as in (B), red: CTE mutations. The single asterisk denotes Swiss-Prot: P16422.2.
Figure 2. (A) Mutants are absent from the plasma membrane. 293T cells expressing an EGFP-ER marker and wt or mutant EpCAM, immunostained with an antibody recognizing EpCAM motif-1 followed by an Alexa647-conjugated secondary antibody. Bar: 20 μm. (B) Living 293T cells transfected and immunostained as in (A) analyzed by FACS. For percentages of gated cells per quadrant, see Supplementary Material, Table S1. ER-EGFP was used to identify transfected cells. (C) Truncated mutants are secreted or degraded. Western blot analysis of lysates and concentrated medium of 293T cells, expressing wt or mutant EpCAM as indicated. Middle blot: cells were treated with 10 μM nocodazole for 24 h to inhibit vesicle transport. MW indication of proteins includes N-glycosylation. (D) EpCAM mutants multimerize. Western blot analysis of 293T cells, expressing wtEpCAM (wt), C66Y, W143_T164del (Del) or empty vector (dashes). The single asterisk denotes monomers; the double asterisk, dimers; the triple asterisk, oligomers.
MATERIALS AND METHODS

Genotyping
Informed consent of the subjects/parents was obtained according to the local institutional review board guidelines. Available unaffected subjects (parents) were also recruited, and, after informed consent, blood samples were collected, and genomic DNA was extracted and sequenced as previously described (3).

Cell culture and transfection
Human embryonic kidney cells (293T; ATCC, USA) were cultured in DMEM (1 g/l glucose) supplemented with 5% FCS and penicillin (100 U/ml)/streptomycin (1 mg/ml). Cells were cultured at 37°C in a humidified incubator in the presence of 5% CO₂. Media and supplementation were obtained from PAA (Austria). 293T cells were transfected with polyethylenimine (linear, MW 25000; Polysciences, Inc., USA) (21). Briefly, 0.5–1 μg of DNA and polyethylenimine (ratio 1:5) were incubated in serum-free medium for 30 to 60 min at room temperature and added to the cells during plating (six-well plate). Cells were used for experiments at least 24 h after transfection.

Plasmids
For this study, human EpCAM cDNA (22) (a kind gift of V. Cirulli) was subcloned into pcDNA3.1. Using this as template, the C66Y mutation was introduced (GACT→GGCT substitution as in Fig. 1A, no. 4) by QuikChange Mutagenesis (Stratagene, The Netherlands). Also mutants K106X (A→T substitution; Fig. 1A, no. 8), R138X (C→T substitution; Fig. 1A, no. 5) and Q167PfsX21 (insertion C; Fig. 1A, no. 6) were generated using QuikChange Mutagenesis. The W143_T164del mutation (Fig. 1A, nos. 1 and 2) was introduced by combining N-terminal and C-terminal PCR fragments, ligated together following a BsiWI cut, resulting in the TY-AL amino acid sequence of the predicted mutant (23). Also, the W143_T164del mutation was introduced to the deletion mutant W143_T164del, ligating N- and C-terminal PCR fragments via an introduced BsiWI restriction site. The predicted mutation Y186FfsX6 was created by using standard cloning techniques. Plasmids were verified by sequencing. The ER marker ER-superluc-EGFP (calnexin pre-sequence) (23) and pcDNA3.1-EGFP (24) and pcDNA3.1-MannII-EGFP-4C (25) have been described.

Flow cytometry
Transfected 293T cells were trypsinized and kept in ice-cold PBS containing 1% BSA and 0.1% Na₂HPO₄ (also used for antibody dilutions; all incubations on ice). For immunostaining, the living, non-permeabilized cells were incubated for 15 min with monoclonal antibody MOC31 (26), which is directed against the first motif in EpCAM’s extracellular domain (wtEpCAM as well as mutants are recognized). After washing, cells were incubated for 15 min with an Alexa Fluor 647-conjugated secondary antibody (Invitrogen, The Netherlands) and used for FACS analysis (FACSCalibur flow cytometer, BD Biosciences, USA).

Confocal microscopy and immunofluorescent staining
Fluorescent images were acquired using a Zeiss LSM 780 confocal microscope (Zeiss, Germany). For immunostaining, transfected 293T cells were fixed by 10% formalin for 30 min (Sigma-Aldrich, The Netherlands), washed (PBS) and permeabilized with 0.1% Triton X-100 (Merck, The Netherlands) in PBS (15 min). After blocking with 1% BSA in PBS for 15 min, cells were incubated with first and secondary antibodies (60 min each, diluted in 1% BSA/PBS) as described for FACS. Nuclei were counterstained using Hoechst 33258 (Sigma-Aldrich, The Netherlands). All immunostaining steps were executed at room temperature.

Western blotting
SDS–polyacrylamide gel electrophoresis and western blotting procedures were conducted according to standard protocols (25) have been described.

SUPPLEMENTARY MATERIAL
Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS
We thank Drs van IJzendoorn, Kampinga and de Leij (UMCG) and Yaron Avitzur for material and discussions, Dr Cirruli (University of Washington) for EpCAM cDNA, M. den Besten (UMCG) for technical assistance and Dr Pasmooij (UMCG) for assistance with mutation prediction.

Conflict of Interest statement. None declared.
REFERENCES


