CONGENITAL CHLORIDE DIARRHEA (CLD) was recognized as a disease entity in 1945 by Gamble et al. (8) and Darrow (5), but it was not until 1971 that it was shown to be caused by an autosomal recessive gene (Ref. 24; also see Ref. 25 [MIM #214700]). The main clinical feature of CLD is prenatal onset of watery diarrhea that in utero leads to polyhydramnios and often premature birth. Ultrasound diagnosis is possible, based on dilated bowel loops (21). Babies have distended abdomens and absence of meconium, and they fail to thrive. They develop metabolic alkalosis, hyperbilirubinemia, dehydration, and severe electrolyte disturbances with hypochloremia, hypokalemia, and hyponatremia. If suspected, diagnosis can be verified by measuring fecal Cl− concentration, which is always >90 mmol/l in a patient with corrected electrolyte balance. Without treatment, the disease is often fatal; occasionally, patients survive in a state of chronic contraction, causing early kidney failure (16). CLD is successfully treated in newborns by intravenous replacement and in older children and adults by oral replacement of fluid and electrolyte losses. Detailed guidelines for therapy have been given by Holmberg (16).

In normal intestine, absorption of Cl− and other electrolytes occurs either passively along the concentration and/or electrostatic gradient or actively against the gradient. Passive transport predominates in the permeable mucosa of upper small intestine, but, in the ileum and especially in the colon, where the intestinal mucosa is tight, active transport is essential for electrolyte trafficking. In the ileum and colon, intestinal NaCl absorption is mediated by the operation of two separate exchangers: Cl−/HCO3− and Na+/H+, which are indirectly coupled by H+ economy. In CLD, the Cl−/HCO3− exchange is absent or defective, causing a severe intestinal Cl− absorption defect (17). Massive amounts of Cl− are lost in the stools, and the patients develop hypochloremia. The respective defect in HCO3− secretion leads to metabolic alkalosis and the acidification of intestinal content, which further inhibit the absorption of Na+ through the Na+/H+ exchanger. In the intestine, the high luminal electrolyte content leads to diarrhea by osmotic mechanisms. Na+ and water losses cause secondary hyperaldosteronism and K+ wastage, leading to both hyponatremia and hypokalemia (16).

IDENTIFICATION OF THE CLD GENE

The first recognized patients were of European descent; altogether, at least 100 patients have been described in many different populations, including the United States, Canada, almost all European countries, the Middle East, and Asian countries such as Japan, Hong Kong, and Vietnam (see Ref. 18 for references). There are three recognized geographic areas where the disease is more common: Finland, Poland, and Saudi Arabia and Kuwait (1, 15, 19, 22, 24, 30). In Finland, the disease occurs in the eastern provinces, with an approximate incidence of 1 in 20,000. In Poland, the incidence has been estimated at 1 in 200,000 across the country, and, in Saudi Arabia and Kuwait, figures as high as 1 in 5,000 have been proposed for some areas. However, consanguineous marriages in these latter countries may cause unusually high local incidences (12).

Although the basic defect in CLD has been determined by physiological studies that involve deficient Cl−/HCO3− exchange in the distal ileum and colon, the biochemical mechanism remained totally unknown until the gene mutated in CLD was identified. Taking advantage of the suggested genetic founder effect in the Finnish population, Kere et al. (20) studied three candidate genes and serendipitously mapped the CLD
gene close to the cystic fibrosis transmembrane conductance regulator (CFTR) gene. More detailed genetic and physical mapping studies (13, 15) suggested a candidate gene known as DRA (for downregulated in adenoma; Ref. 26). DRA was initially cloned as a candidate tumor-suppressor gene, based on its abundant expression in the colon and its strongly reduced expression in colon adenomas and carcinomas. Later, it was observed to have high protein homology to a newly cloned sulfate transport protein encoded by the diastrophic dysplasia gene DTDST (10). Thus the genetic map position of DRA, its site of expression, and its homology to a sulfate transporter made it a passable candidate for the CLD gene. A mutation search in Finnish CLD patients revealed homozygosity for a single mutation, deletion of valine at position 317 (V317del) in all 32 Finnish patients studied, in accordance with the expected founder effect. A nonsense mutation, 344delT, was found in two Polish patients, and a missense mutation, H124L, was found in another two. These results implicated DRA as the CLD gene (14).

**WORLDWIDE MUTATION SPECTRUM**

The identification of the gene mutated in CLD and determination of its genomic structure (9) allowed molecular epidemiologic studies worldwide. The studies of CLD patients from diverse populations have revealed in all cases mutations in the DRA gene; thus it appears that CLD as a clinical entity is caused by mutations in a single gene (Table 1). So far, at least 20 different mutations have been identified in CLD patients (6, 11, 12, 14). Each of the three populations with an observed high incidence has its own founder mutation: in Finland, the V317del mutation is present in all but one chromosome studied to date (98%); in Saudi Arabia and Kuwait, the G187X mutation is present in 17 of 18 chromosomes (94%); and, in Poland, the most common mutation I675–676ins accounts for 16 of 34 disease chromosomes (47%). Thus the Polish founder mutation gives rise to homozygotes but also compound heterozygotes together with a set of other mutations that would only rarely result in the disease in the absence of the founder mutation. In contrast, the rare patients from other populations have been found to be compound heterozygotes carrying different mutations in each chromosome. The fact that CLD has been observed in many populations worldwide and the wide spectrum of mutations suggest that new mutations in the CLD gene are not uncommon, but none of the CLD mutations has undergone such an enrichment as the F508del mutation of its chromosomal neighbor, the CFTR gene.

**BIOCHEMISTRY, FUNCTION, AND REGULATION OF THE CLD PROTEIN**

The CLD (or DRA) protein is expressed in the epithelial cells, particularly in the brush-border cells of normal ileum and colon, but is also expressed in a few

---

**Table 1. Mutations of the CLD gene**

<table>
<thead>
<tr>
<th>cDNA Change</th>
<th>Codon</th>
<th>Predicted Coding Change</th>
<th>Population</th>
<th>No. of Patients</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>177–1788insC</td>
<td>60</td>
<td>Frameshift leading to nonsense change at codon 60 followed by STOP at codon 70</td>
<td>North American</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>268–269insAA</td>
<td>90</td>
<td>Frameshift leading to nonsense change at codon 91 followed by STOP at codon 93</td>
<td>Hong Kong</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>344delT</td>
<td>115</td>
<td>Frameshift leading to nonsense change at codon 115 followed by STOP at codon 133</td>
<td>Polish</td>
<td>3</td>
<td>11, 14</td>
</tr>
<tr>
<td>358G &gt; A</td>
<td>120</td>
<td>Glycine &gt; serine, G120S</td>
<td>Polish, Swedish</td>
<td>2</td>
<td>11, 12</td>
</tr>
<tr>
<td>371A &gt; T</td>
<td>124</td>
<td>Histidine &gt; leucine, H124L</td>
<td>Polish</td>
<td>3</td>
<td>11, 14</td>
</tr>
<tr>
<td>392G &gt; G</td>
<td>131</td>
<td>Proline &gt; arginine, P131R</td>
<td>North American</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>559G &gt; T</td>
<td>187</td>
<td>Glycine &gt; STOP, G187X</td>
<td>Kuwaiti, Saudi Arabian</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>IVSS-1G &gt; T</td>
<td></td>
<td>Change in the intron acceptor site AG at nt −1 of intron 5 leading to the loss of a splice site</td>
<td>North American</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>IVSS-2A &gt; G</td>
<td></td>
<td>Change in the intron acceptor site AG at nt −2 of intron 5 leading to the loss of a splice site</td>
<td>North American</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>915C &gt; A</td>
<td>305</td>
<td>Tyrosine &gt; STOP, Y305X</td>
<td>Polish</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>951–953delGGT</td>
<td>317</td>
<td>In-frame loss of a valine, V317del</td>
<td>Finnish</td>
<td>32</td>
<td>14</td>
</tr>
<tr>
<td>IVSS-1G &gt; A</td>
<td></td>
<td>Change in the intron acceptor site AG at nt −1 of intron 11 leading to the loss of a splice site</td>
<td>Polish</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>1487T &gt; G</td>
<td>496</td>
<td>Leucine &gt; arginine, L496R</td>
<td>Hong Kong</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>1516delC</td>
<td>505</td>
<td>Frameshift leading to nonsense change at codon 505 followed by STOP at codon 534</td>
<td>Polish</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>1526–1527delTT</td>
<td>509</td>
<td>Frameshift leading to nonsense change at codon 509 followed by STOP at codon 517</td>
<td>Japanese</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>1548–1551delAACC</td>
<td>516</td>
<td>Frameshift leading to nonsense change at codon 518 followed by STOP at codon 534</td>
<td>Polish</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>1578–1580delTTA</td>
<td>527</td>
<td>In-frame loss of a tyrosine, Y527del</td>
<td>Polish</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>1609delA</td>
<td>537</td>
<td>Frameshift leading to nonsense change at codon 537, STOP at codon 575</td>
<td>North American</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>2025–2026insATC</td>
<td>676</td>
<td>In-frame addition of an isoleucine, 1675–676ins</td>
<td>Polish</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>2116delA</td>
<td>706</td>
<td>Frameshift leading to nonsense change at codon 706, STOP at codon 711</td>
<td>Finnish</td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>

CLD, congenital chloride diarrhea.
other tissues (4, 14) (Fig. 1). It is predicted to be a 10-,
12-, or 14-transmembrane segment integral membrane
protein, and it is variably N-glycosylated (4) (Fig. 2).
There are no recognizable ATP-binding domains. With
the use of two different expression systems (Xenopus
laevis oocytes and Sf9 insect cells), the protein has
been shown to constitute an Na\(^+\)-independent sulfate
transporter that can be inhibited by the anion trans-
porter inhibitor DIDS (3, 27). Its anion specificity also
includes at least oxalate and, as shown recently in a
Xenopus oocyte model, Cl\(^-\) (23). In this cell model, Cl\(^-\)
transport mediated by the CLD protein was also shown
to have features compatible with a Cl\(^-\)/OH\(^-\) ex-
changer (23). The exact biochemical and physiological
role of the CLD protein and its relationship to Cl\(^-\)/
HCO\(_3\)\(^-\) exchange in ileum and colon are still subject
to study. However, the known functions and the role of
the mutated protein in the pathogenesis of CLD
provide several hypotheses that can now be directly
tested.

The simplest and perhaps most intriguing model,
is that the CLD protein alone constitutes in brush-
border cells the apical Cl\(^-\)/HCO\(_3\)\(^-\) exchanger whose
physiological role is to absorb up to 900 mmol of Cl\(^-\)
daily and expel an equal amount of HCO\(_3\). This hypo-
thesis, if true, would add another structurally different
protein family to the known Cl\(^-\)/HCO\(_3\) exchangers. In
humans, there are three anion exchanger (AE) genes
designated as AE1, AE2, and AE3, which encode for the
erthrocyte band 3 protein and two homologous anion
transporters, respectively. Mutations in the AE1 gene
cause red blood cell abnormalities (29), and knockout
mice show spherocytosis and hemolytic anemia due to
reduced membrane stability (28). The human AE gene
family members show \(\sim 20\%\) homology over \(\sim 100\) amino
acids to the CLD protein, whereas the two other members
of the human sulfate transporter gene family (see below)
are 32–48\% homologous to the CLD protein over more
than 700 amino acids. Thus the AE family and sulfate
transporter gene family are structurally distinct.

Fig. 1. Immunohistochemical staining of the congenital chloride diarrhea (CLD) protein in normal colon (left). Preimmune serum control is also shown (right). Note the positively staining epithelial cells along the surface but the scarcity of positive cells in crypts.
Other models are more complex. The CLD protein may be part of a multiprotein structure with a Cl\(^{-}/\)HCO\(_3\)\(^{-}\) exchanger function. In this hypothesis, the CLD protein must have a critical regulatory role, to accommodate for the fact that mutations in it cause CLD in vivo and a Cl\(^{-}\) transport defect in a Xenopus oocyte model.

In yet another model, the Cl\(^{-}/\)HCO\(_3\)\(^{-}\) exchange may require two (or more) transporter proteins having functions that are tightly coupled by a common substrate. In this model, mutation in just one of the proteins, CLD, would block the functions of both transporters, causing the observed overall defect in Cl\(^{-}/\)HCO\(_3\)\(^{-}\) exchange.

The availability of several naturally occurring mutations associated with CLD disease will facilitate the testing of hypotheses about the protein’s functional

### Table 2. Three structurally related human genes with an anion transport function and mutations in distinct diseases

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosomal Location</th>
<th>Structural Information</th>
<th>Site of Expression</th>
<th>Transported Anions</th>
<th>Disease Caused by Mutations</th>
<th>Number of Known Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLD (DRA)</td>
<td>7q22-q31.1</td>
<td>21 exons, 764 aa</td>
<td>Differentiated cells in ileal and colonic mucosa, prostate, thyroid, other?</td>
<td>Sulfate, oxalate, chloride, others?</td>
<td>CLD</td>
<td>20</td>
</tr>
<tr>
<td>PDS</td>
<td>7q22-q31.1</td>
<td>21 exons, 780 aa</td>
<td>Thyroid, other?</td>
<td>Not known</td>
<td>Pendred syndrome</td>
<td>15</td>
</tr>
<tr>
<td>DTDST</td>
<td>5q32-q33.1</td>
<td>4 exons, 739 aa</td>
<td>Many tissues, more in cartilage and ileum</td>
<td>Sulfate, chloride, others?</td>
<td>Diastrophic dysplasia, congenital achondrogenesis type II, atelosteogenesis type II</td>
<td>14</td>
</tr>
</tbody>
</table>

aa, Amino acids; DRA, downregulated in adenoma gene; PDS, putative sulfate transporter gene; DTDST, diastrophic dysplasia gene.
Mutations seem to occur in all parts of the protein, but at least two potentially interesting features emerge. First, there seem to be three clusters of several tightly spaced mutations (12). They may represent mutation-prone segments of DNA or, alternatively, protein domains that are critical for intact function. Second, there are several mutations that localize in the long COOH-terminal tail in the 10- and 12-transmembrane segment models and in a large intracellular loop in the 14-transmembrane segment model (Fig. 2). It is less likely that these mutations would affect the synthesis or integration of the protein in the cell membrane; rather, this COOH-terminal, mutation-prone domain may have a distinct regulatory function. This inference is supported also by the presence of several protein kinase C and casein kinase II target sites in the COOH-terminal segment.

So far, the biochemical effects of mutations on CLD mRNA or protein processing and function remain largely unknown. Most data are available on the V317del mutation common in Finland. Both CLD mRNA and protein have been detected in colon epithelia of patients, whereas expression of the protein in Xenopus oocytes showed deficient sulfate and Cl\(^-\) transport in comparison to the wild-type protein (23). Further biochemical studies are needed to understand the effect of different mutations on CLD protein processing and function.

It is well established that the downregulation of the CLD gene correlates with colon tumor progression (2, 26). The wide spectrum of different mutations associated with CLD suggests that the primary pathogenetic effect of mutations is indeed in anion transport, and downregulation in cancer cells may well be a secondary effect. The specific expression patterns of the CLD gene in different tissues and different sites in the gut epithelium and in normal and malignant cells make its regulation a highly interesting subject for study.
A FAMILY OF TISSUE-SPECIFIC ANION TRANSPORTERS

In addition to the CLD and DTDST genes mutated in a disorder of gut anion exchange and in growth retardation syndromes, respectively, a third closely related gene, PDS, was recently found to be defective in Pendred syndrome, characterized by congenital sensorineural deafness and goiter (7) (Table 2). Structurally, all three human genes belong to the family of sulfate transporters, which, as we gather more information about its members’ physiological roles, may turn out to be a misnomer. This protein family and the relationships of structurally related proteins across all organisms are illustrated in a phylogenetic tree (Fig. 3). We wish to emphasize that the tree was not constructed to illustrate relationships between the different organisms or between the functional roles of the different proteins but rather to give a homology-based grouping of the protein members of the family. Structurally related proteins with conserved domains are present among bacteria, yeast, fungi, plants, and animals. It is, however, striking that the animal branch has the smallest number of known members. Furthermore, 7 of 13 animal entries represent proteins from Caenorhabditis elegans whose genes are almost exhaustively listed at present. In all other animal species, only six members have been identified, three of them from humans; each of the human genes has a distinct, tissue-specific function that has been associated with a specific disorder by positional cloning (Table 2).

These observations suggest that numerous other members of this gene and protein family may remain unidentified. Speculatively, it is possible that additional family members would exist in humans, with perhaps distinct, tissue-specific functions, as suggested by the known CLD, DTDST, and PDS genes, proteins, and diseases. Their transport specificity may not be restricted to mostly sulfate in vivo, but it is indeed possible that the transport of other anions is physiologically more important. The functional study of the known proteins and the identification of additional members of this gene family provide rich sources for a student of the physiology of anion transport and will promote the understanding of specific transport mechanisms in different parts of the body.

We thank Albert de la Chapelle and Christer Holmberg for numerous discussions, collaboration, and support; Siru Haila and Merja Nissinen for excellent laboratory work; and our many clinical collaborators for making mutation studies possible.

Our research on CLD is supported by the Finnish Foundation for Pediatric Research, Ulla Hiett Fund, Helsinki University Central Hospital research funds, Sigrid Jusélius Foundation, Academy of Finland, and Research and Science Foundation of Farmos.

Address for reprint requests: J. Kere, Finnish Genome Center, PO Box 21 (Tukholmankatu 2), Univ. of Helsinki, 00014 Helsinki, Finland.

REFERENCES


Downloaded from http://ajpgi.physiology.org/ by 10.220.33.2 on October 29, 2017


