

Transmembrane topology of a CLC chloride channel

THOMAS SCHMIDT-ROSE AND THOMAS J. JENTSCH†

Center for Molecular Neurobiology Hamburg, ZMNH, Hamburg University, Martinistrasse 52, D-20246 Hamburg, Germany

Communicated by William A. Catterall, University of Washington School of Medicine, Seattle, WA, May 12, 1997 (received for review February 13, 1997)

ABSTRACT CLC chloride channels form a large and conserved gene family unrelated to other channel proteins. Knowledge of the transmembrane topology of these channels is important for understanding the effects of mutations found in human myotonia and inherited hypercalciuric kidney stone diseases and for the interpretation of structure–function studies. We now systematically study the topology of human CIC-1, a prototype CLC channel that is defective in human myotonia. Using a combination of *in vitro* glycosylation scanning and protease protection assays, we show that both N and C termini face the cytoplasm and demonstrate the presence of 10 (or less likely 12) transmembrane spans. Difficult regions were additionally tested by inserting cysteines and probing the effect of cysteine-modifying reagents on CIC-1 currents. The results show that D3 crosses the membrane and D4 does not, and that L549 between D11 and D12 is accessible from the outside. Further, since the modification of cysteines introduced between D11 and D12 and at the extracellular end of D3 strongly affect CIC-1 currents, these regions are suggested to be important for ion permeation.

Voltage-gated chloride channels of the CLC family are highly conserved during evolution and are expressed in organisms ranging from bacteria (1) and yeast (2) to plants (3) and animals (4). Their physiological functions in higher organisms include the regulation of cell volume, control of electrical excitability, and transepithelial transport.

The CLC proteins were identified by expression cloning of the *Torpedo* chloride channel CIC-0 (5). At present, nine mammalian members are known (for review, see ref. 6). CIC-1 is nearly specific for skeletal muscle (4). It ensures its electrical stability, which is evident from CIC-1 mutations leading to myotonia congenita, a disease characterized by defective muscle relaxation. Many different CIC-1 mutations were found in human myotonia and in animal models (7–11). The importance of CLC channels is underscored by another, recently identified disease. Inactivating mutations of CIC-5 cause a syndrome that is characterized by low molecular weight proteinuria, hypercalciuria, and kidney stones (12). CIC-2 may play a role in cell volume control (13) and in setting the intracellular chloride concentration, which in turn is important for synaptic transmission in certain neurons (14). However, the function of most CLC proteins is still unclear.

CLC proteins are structurally unrelated to other ion channel classes. Topological models are still mainly based on hydrophathy analysis, suggesting the presence of about 12–13 transmembrane domains, originally termed D1 to D13 (5). This model has already had to be revised to accommodate new experimental findings (15, 16).

CLC channels function as multimers of identical (11) or homologous subunits (17). Dominant negative mutations suggested that CIC-1 channels have more than three subunits (11).

However, CIC-0 is a homodimeric channel with one pore per subunit (16, 18, 19); this may also apply for CIC-1 (20).

In structure–function studies, mutations have been introduced into CIC-0 (18, 21), CIC-1 (9, 11), CIC-2 (13, 22), and CIC-5 (12). These mutations showed that several protein regions are important for gating and permeation. Unfortunately, the interpretation of these data is limited by the lack of structural information.

We put the topology of CLC channels on a firm experimental basis. We primarily used the glycosylation scanning procedure (23–25), which we complemented with protease protection assays (26, 27) and by introducing cysteines that were then probed with extracellular cysteine-modifying reagents for effects on currents (28–30). In addition to clarifying the transmembrane topology, the latter technique identified novel CIC-1 residues, which are important for channel function.

EXPERIMENTAL PROCEDURES

Construction of Glycosylation Mutants 1T2 to 12T13. We started with pTLB·H1NQ, a mutant CIC-1 cDNA in the expression vector pTLN (17). The CIC-1 glycosylation site (N430) was mutated to glutamine and an epitope (LYPYD-VPDYVSG) was inserted behind G972; the latter is of no importance for the present work. For C-terminally truncated pTLB·H1NQΔC, the sequence TGATAAGAATTC containing two stop codons and an *EcoRI* site was introduced after L629. Glycosylation transplants, encompassing a part (F413 to G444, referred to as “T,” or G409 to V456, indicated by “T*”) of the D8/D9 loop of CIC-1 including the native glycosylation site, were inserted into hydrophilic loops by recombinant PCR (31) as detailed in Fig. 1. All PCR-derived fragments were sequenced. Full-length glycosylation constructs for functional tests were gained from the truncated forms by swapping the respective stretches from pTLB·H1NQΔC into pTLB·H1NQ.

Construction of CIC-1/Prolactin Fusion Proteins. Full-length prolactin cDNA was obtained by reverse transcription-PCR from bovine pituitary gland poly(A)⁺ RNA using Pfu DNA polymerase (Stratagene) and was cloned into pTLN to yield pTLN·Prl.

Constructions of NPrl, 1Prl, . . . 8Prl, 10Prl, 12Prl, 13Prl: Serial truncations of human CIC-1 cDNA, starting at the initiator ATG and extending to positions before or behind potential transmembrane spans, were fused to the prolactin cDNA C-terminal of Y74 (referred to as “prolactin epitope”). Fusions were generated by PCR-amplification of CIC-1 fragments and prolactin by using one *Bam*HI-site containing primer, respectively, so that ligation introduced two additional residues (G and S) at the fusion site. In constructs ND3Prl, ND4Prl, and ND6Prl domains D3 (M193 to F235), D4 (K231 to Y262), or D6 (T293 to D340), respectively, are fused between the N terminus of CIC-1 (ending at G115) and the prolactin epitope on the C-terminal side. In SPD4Prl, the N terminus of CIC-1 of construct ND4Prl was replaced by the signal peptide of pre-prolactin, starting with the initiator

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1997 by The National Academy of Sciences 0027-8424/97/947633-6\$2.00/0
PNAS is available online at <http://www.pnas.org>.

Abbreviations: MTSES, sodium (2-sulfonatoethyl)methanethiosulfonate; WT, wild type.

†To whom reprint requests should be addressed.

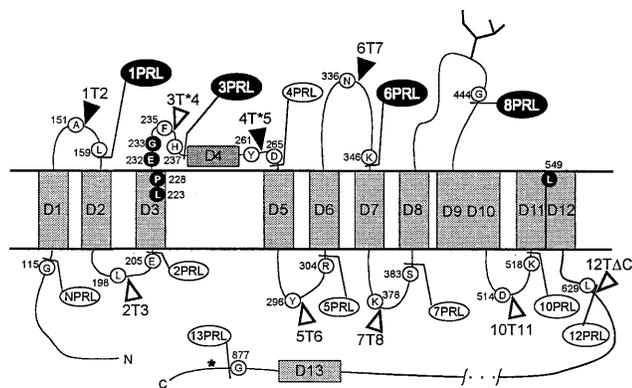


FIG. 1. Topological model of human CIC-1 illustrating the constructs used in this work. Hydrophobic domains are labeled D1 to D13 according to the original nomenclature for CIC-0 (5) based on Kyte–Doolittle hydropathy analysis. For amino acid sequence, see ref. 4. WT-glycosylation is represented by a branched line in loop D8/D9. An unused glycosylation site after D13 is indicated by an asterisk. Insertion sites of glycosylation transplants are indicated by triangles (solid = glycosylation positive), fusion sites with prolactin by ball-and-stick (solid = proteinase protection positive) after the amino acids shown in open circles. Amino acids in black have been replaced by cysteines and are sensitive to cysteine specific reagents.

methionine and ending with D61. N34T*5PrI consists of the CIC-1 N terminus fused to the sequence encoding D3 through D5 (M193 to R304), including the glycosylation transplant between D4 and D5 as in construct 4T*5. The C terminus was then fused to the prolactin epitope.

In Vitro Transcription. Capped cRNA was transcribed by SP6 RNA polymerase from 0.5 μ g of plasmid after linearization with *Mlu*I using the mMessageMachine cRNA synthesis kit (Ambion).

In Vitro Translation. Glycosylation constructs were translated using rabbit reticulocyte lysate (Promega) in the presence or absence of canine pancreatic microsomes (Promega) and in the presence of [³⁵S]methionine. For SDS/PAGE, 2–5 μ l of translation mixture was diluted in 5–10 μ l of sample buffer and separated on 8% gels. After drying, gels were examined with a bioimage-analyzer (Fuji BAS1500). Translation for protease protection experiments was done with microsomes but without radioactivity.

Protease Protection Assay and Western Blot Analysis. Translation mixture was brought to 10 mM CaCl₂ and chilled on ice. Aliquots of 10 μ l were incubated with proteinase K (30 μ g/ml, Boehringer Mannheim) in the presence or absence of 1% Triton X-100. Controls remained without proteinase and detergent. Proteolysis proceeded on ice for 60 min, and was stopped by adding 5 mM phenylmethylsulfonyl fluoride. After 10 min on ice, 50 μ l of preheated (95°C) sample buffer was added and the sample was boiled for 15 min to inactivate the protease. Proteins were separated on a 5–20% SDS/PAGE gradient, blotted onto nitrocellulose (Schleicher & Schuell), and probed with an anti-prolactin antiserum (ICN) using protein A-coupled horseradish peroxidase (Bio-Rad) and the Renaissance chemiluminescence kit (DuPont/NEN).

Electrophysiology. cRNA (10–15 ng) in 50 nl was injected into *Xenopus* oocytes, which were kept in modified Barth's solution (88 mM NaCl/2.4 mM NaHCO₃/1.0 mM KCl/0.41 mM CaCl₂/0.33 mM CaNO₃/0.82 mM MgSO₄/10 mM HEPES, pH 7.6) for 2 to 3 days and analyzed in ND96 saline (96 mM NaCl/2 mM KCl/1.8 mM CaCl₂/1 mM MgCl₂/5 mM HEPES, pH 7.4). Standard two electrode voltage clamp measurements were performed at room temperature using a Turbotec amplifier (Npi Instruments) and pCLAMP 5.5 software (Axon Instruments, Foster City, CA).

Probing Cysteine Residues with SH Reactive Compounds.

We substituted L223, P228 in D3, E232 and G233 between D3 and D4, and L549 in the D11/D12 block with cysteine. PCR mutagenesis, cRNA synthesis, and injection into *Xenopus* oocytes were performed as described. Currents were measured in ND96 with 20 mM HEPES (pH 7.4) (NDHep) and monitored for reproducibility. The bath was then changed for 3 min to 10 mM sodium (2-sulfonatoethyl)methanethiosulfonate (MTSES) in NDHep, during which currents reached steady state. Oocytes were then washed extensively with NDHep and recorded again. By comparing currents before and after MTSES application, effects due to noncovalent interactions were excluded.

RESULTS

Probing CIC-1 Topology by Glycosylation. The highly conserved consensus site for N-linked glycosylation between D8 and D9 is used in CIC-0, CIC-1, and the CIC-K channels (15, 16). Thus, in contrast to the initial topology model (5), this region must face the extracellular space. To discover which other regions are extracellular, we inserted glycosylation consensus sites between various hydrophobic domains of CIC-1 (Fig. 1). To detect exclusively the effects of newly introduced sites, we based these constructs on a mutant (CIC-1N430Q) lacking the wild-type (WT) glycosylation site. This mutant is functional (15). Unfortunately, most hydrophilic loops of CIC-1 are shorter than 33 residues, which is a lower limit for reliable glycosylation (32, 33). Therefore, we inserted a copy of the D8–D9 loop that is glycosylated in native CLC proteins (15). A similar strategy was used successfully, e.g., in the analysis of Glut-1 glucose transporter topology (34).

Core glycosylation increases the molecular weight by about 2 kDa. With CIC-1 (\approx 110 kDa) the relative increase in mass is just \approx 2% if a single site is glycosylated. The resulting shift is difficult to detect by SDS/PAGE. To increase the relative mass difference, we used a truncated construct which ends about 50 amino acids downstream of the last transmembrane domain. As membrane integration is a sequential process starting at the N terminus, this should not influence upstream topology.

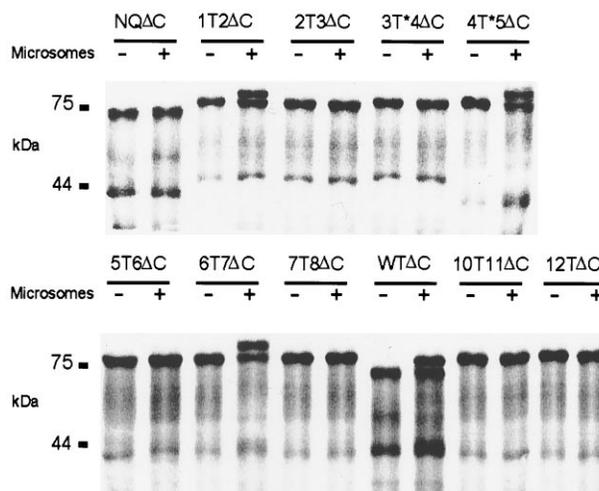


FIG. 2. Autoradiograph of *in vitro* translation products with glycosylation sites engineered between hydrophobic domains. RNA was translated in reticulocyte lysate in absence (–) or presence (+) of microsomal membranes. All constructs are shortened at the C terminus as indicated by Δ C. WT Δ C contains the native glycosylation site. In NQ Δ C the glycosylation site is removed by mutating N430 to Q. “T” in other constructs indicates the standard glycosylation transplant, “T*” the longer variant, and the flanking numbers the respective hydrophobic domains. Numbers at left indicate the molecular mass in kDa.

Fig. 2 shows the size analysis of WT and mutant proteins synthesized *in vitro* in the presence and absence of microsomes. The shortened WT protein, WT Δ C, and the equivalent glycosylation knockout, NQ Δ C, yielded a \approx 74-kDa band when translated without microsomes. With microsomes, an additional band of \approx 77 kDa reflecting the glycosylated protein was seen with WT Δ C but not with NQ Δ C. The shift to higher masses was incomplete, indicating that part of the protein is still synthesized at free ribosomes in the cytosol also with microsomes. This could not be improved by changing the cRNA/microsome ratio. The \approx 42-kDa band in Fig. 2 is due to a fraction of truncated cRNA, which results from a premature transcription stop of the SP6 RNA polymerase (C. Lorenz and T.J.J., unpublished observations). This band was larger in constructs 1T2 Δ C, 2T3 Δ C, and 3T*4 Δ C because it contained the glycosylation transplant.

Glycosylation occurred between D1/D2, D6/D7, and D8/D9 (equivalent to WT glycosylation). In contrast, the standard glycosylation transplant could not be glycosylated when inserted in the short hydrophilic stretches flanking D4 (data not shown). Since these are the shortest interdomain segments, we introduced a larger transplant. The segment behind D4 could now be glycosylated (construct 4T*5 Δ C), while the one before D4 (construct 3T*4 Δ C) remained unglycosylated.

Chloride Channel Function of Glycosylation Constructs. To exclude that these transplants had perturbed channel structure, we expressed full-length cRNAs containing these inserts in *Xenopus* oocytes and measured whole cell currents. It is highly unlikely that mutants with a changed transmembrane topology still yield functional chloride channels. In contrast, a loss of currents does not prove an altered topology, as the insertions may have abolished channel function by other means.

In general, currents from glycosylation constructs were smaller than those expressed from identical amounts of WT cRNA (Fig. 3). Western blot analysis of *Xenopus* oocyte crude membrane preparations showed that the reduction in currents was paralleled by lower protein levels (data not shown). Thus, insertion constructs are translated less efficiently or degraded more efficiently, or both. Currents of 3T*4, 5T6, and 12T13 were not significantly different from water-injected controls. By contrast, the remaining six constructs elicited currents of 5–30% as compared with WT and NQ mutants. Thus, these insertions did not alter the transmembrane topology. Importantly, this includes all mutants that were glycosylated. Currents of constructs 1T2 and 6T7 showed qualitative changes, indicating that insertions at these positions changed channel properties, but did not abolish its function. With 1T2 the voltage dependence was shifted to more positive values, and

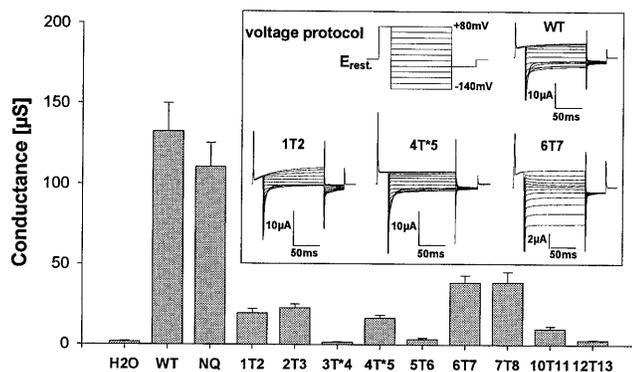


FIG. 3. Two-electrode voltage clamp experiments of *Xenopus* oocytes expressing full-length glycosylation constructs. Mean conductance at -40 mV of WT CIC-1 and its mutants in comparison to water injected control oocytes (mean \pm SEM, $n = 4$ to 9). (Inset) Pulse protocol and typical current traces for WT and constructs 1T2, 4T*5, and 6T7.

outward currents slowly activated. Mutant 6T7 has almost completely lost the inactivation behavior at hyperpolarizing potentials and its rectification is strongly changed.

Proteinase Protection Assay. Glycosylation of a given site proves that it is extracellular. However, absence of glycosylation is no proof that it is intracellular, because the access of the glycosyltransferase could be impaired by sterical hindrance. Proteinase protection, however, can probe for an intracellular localization. An epitope is fused to the protein at different positions. If the epitope ends up in the lumen of the endoplasmic reticulum, it can be degraded by proteinase K only if the integrity of the membrane is destroyed by detergents (35). Conversely, complete degradation in the absence of detergent indicates a cytosolic location.

We constructed a series of truncated CIC-1 proteins in which a prolactin epitope was fused behind putative transmembrane domains (Fig. 1). Portions of prolactin lacking its endogenous signal sequence are topogenically neutral and respond faithfully to upstream signal-anchor and stop-transfer sequences (23, 36, 37). These constructs were translated *in vitro* in the presence of microsomes and treated with proteinase K in the absence or presence of Triton X-100. Fragments containing the prolactin epitope were detected by Western blot analysis.

As expected, the hydrophilic N terminus of CIC-1 confers no protection on the epitope in construct NPrl (Fig. 4), whereas D1 (construct 1PrI) acts as a signal-anchor and translocates the epitope into the endoplasmic reticulum. The protected fragment is \approx 20 kDa in size, which results from 17.7 kDa of the prolactin domain and 2.3 kDa of the transmembrane domain D1. Protection of prolactin-containing fragments of similar sizes are found with constructs 6PrI and 8PrI, consistent with the extracellular localization indicated by our glycosylation analysis. 3PrI only shows a weak signal at \approx 20 kDa, suggesting that the translocation of the epitope occurred less efficiently. Addition of detergent always led to complete digestion. No protection could be observed with 2PrI, 4PrI, 5PrI, 7PrI, 10PrI, and 13PrI.

The lack of protease protection with 4PrI apparently contradicts the extracellular localization of the D4–D5 linker inferred from glycosylation of construct 4T*5 Δ C (Fig. 2). The glycosylation assay is more trustworthy because no transmembrane domains were deleted in that construct and because the full-length protein 4T*5 yielded typical chloride currents. This

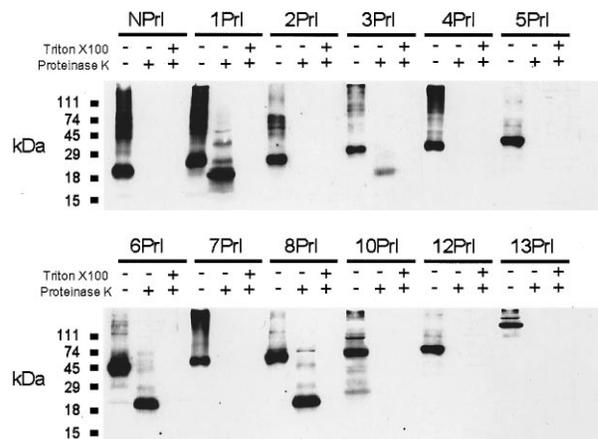


FIG. 4. Proteinase protection assay. Increasing portions of CIC-1 starting at the N terminus (N) were fused C-terminally to a prolactin reporter-epitope (PrI). Numbers indicate the last hydrophobic domain of CIC-1 that is included in the fusion protein. *In vitro* translation was performed in the presence of pancreatic microsomes. Products were treated with proteinase K with (+) or without (-) Triton X-100. Controls were treated equivalently, but proteinase K and detergent were omitted (-/-).

suggests that the transfer of D4 across the membrane may be more complicated than that of other domains.

We therefore tested the capability of D3 or D4 to act individually as signal-anchor sequences (Fig. 5). We fused the CIC-1 N terminus to D3 and D4 and then to the prolactin epitope (constructs ND3PrI and ND4PrI). Prolactin was protected against digestion in control constructs ND1PrI and ND6PrI, but neither D3 nor D4 alone were able to translocate the epitope. We also tested whether D4 might act as a stop-transfer sequence by replacing the CIC-1 N terminus in ND4PrI by the prolactin signal sequence to initiate translocation (construct SPD4PrI). When translated with microsomes, the product was shortened by the signal peptidase. The truncated peptide was completely resistant to proteinase K. This indicates that D4 cannot redirect the C terminus into the cytoplasm.

Thus, D3 and D4 cannot individually translocate the prolactin epitope. However, construct 3PrI (Fig. 4) indicates that D3 can translocate the epitope (albeit inefficiently) if preceded by D1 and D2, and glycosylation shows that the region before D5 is extracellular. In addition, constructs 2PrI and 5PrI suggest that the stretches preceding D3 and following D5, respectively, are cytoplasmic. To test whether domains D3, D4, and D5 may insert into the membrane only if translated together, we designed construct N34T*5PrI. Following the CIC-1 N terminus, the construct contains the complete D3–D5 region (including the extended glycosylation transplant tested in construct 4T*5ΔC) fused to the reporter epitope. Translating this construct in the presence of microsomes indeed resulted in glycosylation (as indicated by its larger size), whereas the prolactin epitope was digestible by proteinase K (Fig. 5). Thus, the stretch after D3/D4 is only translocated if it is followed by D5. Glycosylation provides strong evidence that the stretch after D4 is extracellular, while the partial protease protection of construct 3PrI indicates that this also applies for the segment preceding D4. The latter evidence, however, is somewhat weaker because it was obtained with a severely truncated protein.

Cysteine Substitution Mutants. We reprobated the D3/D4 region in the context of the full-length channel. As enzymatic modification of residues may be hindered by steric constraints, we replaced single residues by cysteines and monitored their acces-

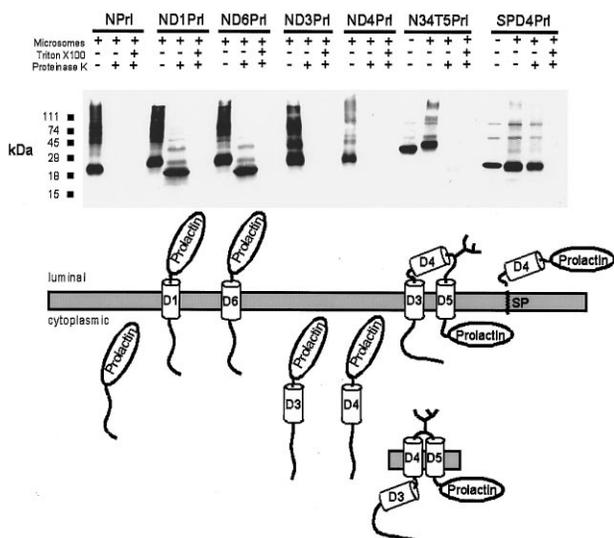


FIG. 5. Testing of individual domains for their topogenic activity. Modular fusions of individual domains with the CIC-1 N terminus and the prolactin reporter epitope. In construct SPD4PrI the N terminus is formed by the signal sequence of pre-prolactin (SP). (Upper) Immunoblot of *in vitro* translation products treated as described above. (Lower) Schematic drawings of topologies compatible with the experimental results.

sibility to comparatively small cysteine-specific reagents using electrophysiology. Methanethiosulfonate derivatives have been used successfully to probe the pore-forming residues of several channel classes (29, 38–40), including the cystic fibrosis transmembrane conductance regulator chloride channel (41, 42). These reagents covalently add protruding charged moieties ($-\text{SCH}_2\text{CH}_2\text{SO}_3^-$ in the case of MTSES) to cysteines, which may interfere with ion permeation if they are next to the pore. If these membrane impermeable reagents affect currents, this proves that the respective residues are accessible from the exterior. A negative result, however, does not yield any information, as the residue may have been modified without any functional effect. In contrast to methods detecting modified cysteines biochemically (43), our functional assay does not require a cysteine-less CIC-1 mutant (that may be nonfunctional) as a starting construct, provided that WT CIC-1 currents are unaffected by MTSES. This is indeed the case (Fig. 6).

We replaced L223 and P228 in D3 as well as E232 and G233 between D3 and D4 (see Fig. 1 for positions), since we knew that mutations at similar positions in CIC-0 affect pore properties (M. Pusch and T.J.J., unpublished data). These substitutions did not abolish currents when expressed in *Xenopus* oocytes. While currents from mutant P228C were like WT, L223C and E232C showed a weaker inactivation at hyperpolarizing potentials. The currents for mutant G233C were smaller and slightly outwardly rectifying. Currents from mutants L223C, P228C, E232C, and G233C, but not from WT CIC-1, were drastically reduced by incubating the oocytes for 3 min in 10 mM of MTSES (Fig. 6). This shows that the end of D3 and amino acids before D4 are accessible to small charged molecules from the exterior and confirm the protease protection experiment with construct 3PrI. In mutant E232C a negatively charged amino acid is replaced by a neutral one.

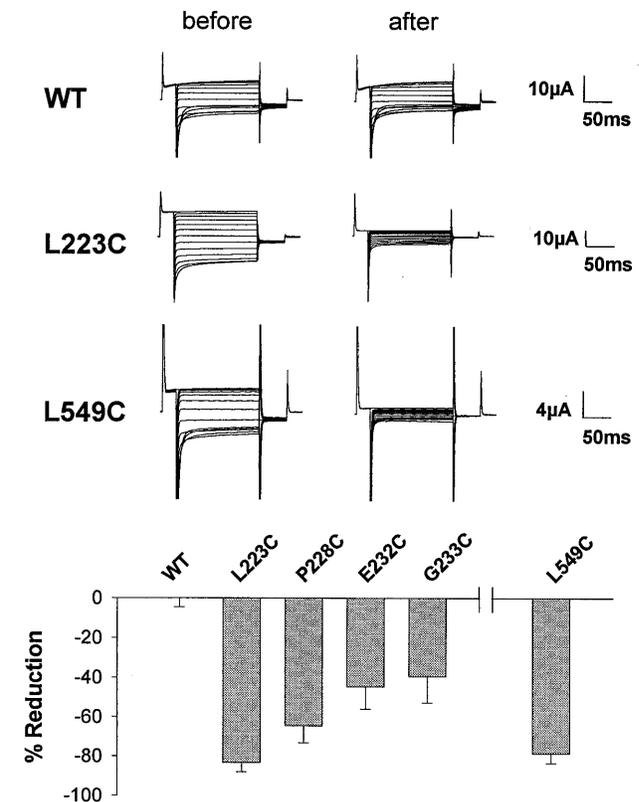


FIG. 6. Effects of MTSES on *Xenopus* oocytes expressing CIC-1 WT or cysteine mutants indicated in Fig. 1. (Upper) Examples of current traces. (Lower) Relative changes in conductance at -40 mV 3 min after application of 10 mM MTSES (mean \pm SEM, $n = 5$ to 9) and subsequent washout.

Although a modification with MTSES reintroduces the negative charge, WT inactivation behavior was not restored.

We used the same technique to test whether residues in the middle of the broad hydrophobic region D11/D12, which enters the cytoplasm at both ends (constructs 10PrI and 12PrI; Fig. 4), may be accessible from the exterior. When we mutated leucine 549, which is located next to a highly conserved glutamate residue, to cysteine, channel activity was retained and like WT. Extracellular application of MTSES drastically inhibited these currents (Fig. 6). Thus, D11/D12 crosses the membrane twice.

Several other introduced cysteines in the vicinity of those described above did not confer MTSES-sensitivity (data not shown).

DISCUSSION

Several techniques are available to probe the transmembrane topology of multispan transmembrane proteins, including probing with epitope-specific antibodies, gene fusions with reporter enzymes, glycosylation scanning, protease protection assays, and cysteine scanning (for review, see ref. 44). All of these techniques have their specific problems, and sometimes different techniques have suggested different topologies. Here we chose three complementary techniques to increase the reliability of the results.

In the first set of experiments we used an *in vitro* glycosylation scanning procedure. Since N-glycosylation exclusively takes place in the endoplasmic reticulum, carbohydrate transfer to an artificially introduced acceptor site indicates its extracellular position. Absence of glycosylation, however, does not prove an intracellular localization. We therefore reprobated CIC-1 topology with the protease protection assay. It indicates a cytoplasmic localization of a reporter epitope by its lability to added protease, while those in the lumen are protected. Additionally, we checked two particular areas by modifying artificially introduced cysteines with extracellularly applied membrane-impermeable reagents. Importantly, all crucial constructs from the first and third set of experiments yielded functional chloride channels, demonstrating that their topology was unchanged.

The first CLC topology models were exclusively based on a hydropathy analysis according to Kyte and Doolittle (45) that indicated 13 hydrophobic stretches long enough to potentially span the lipid bilayer. These were termed D1 to D13 (ref. 5; see Fig. 1). The broad hydrophobic region preceding the hydrophilic C terminus was arbitrarily divided into domains D9 through D12. Hydropathy analysis also identified a C-terminal region (D13) of weak hydrophobicity. In contrast to the D12–D13 interdomain stretch, D13 is highly conserved.

CLC proteins have no cleavable signal peptide. This is compatible with (but does not prove) a cytoplasmic N terminus. The orientation of signal-anchor sequences is strongly influenced by the charge distribution in flanking sequences. Positive charges are placed preferentially in the cytosol ("positive-inside rule"; refs. 46 and 47). With CIC-1, the first transmembrane span D1 was directly preceded by five positive and two negative charges, but there were only two positive residues directly behind D1. This suggests an intracellular N terminus. This was confirmed by our experiments, which demonstrate that D1 is a signal-anchor sequence that assumes type II orientation (46). The second hydrophobic domain, D2, was a stop-transfer sequence whose orientation was again compatible with the charge distribution at its ends.

The D3/D5 stretch was the only region in which a protease protection experiment apparently contradicted the glycosylation assay. The latter results suggest that the D4–D5 linker is extracellular—and this is strong evidence because the full-length glycosylation construct is functional—while the prolactin fused at the end of D4 was not translocated into the endoplasmic reticulum. We could resolve this problem by

showing that D3–D5 correctly insert into the membrane only when translated together. By contrast, D3 or D4 could not insert into the membrane individually. Indeed, neighboring domains act synergistically during membrane insertion in several other proteins (27, 37, 48–50). The results from both protease protection (construct 3PrI) and cysteine scanning suggest that D3 crosses the membrane and that its C-terminal part is accessible from the exterior. Thus, there is strong cumulative evidence that D3 does indeed span the lipid bilayer, but that D4 does not (Fig. 1). This is consistent with our previous suggestion (51), which was based on a comparative analysis of several CLC proteins. While D4 presents as a peak of moderate hydrophobicity in Kyte–Doolittle plots of CIC-0 and CIC-1 and is conserved between these proteins, there is nearly no sequence conservation to more distantly related CLC proteins like CIC-5 or CIC-7. The latter channels even lack an hydrophobic stretch in that segment.

Glycosylation and protease protection assays consistently show that domains D6 to D8 are transmembrane domains. Compatible with the "positive-inside" rule, this places the D7–D8 linker into the cytosol. Though poorly conserved by sequence, the linker is strongly positively charged in all known CLC proteins. The extracellular localization of the D8–D9 linker is consistent with its glycosylation in CIC-0 and CIC-K (15, 16), and most mammalian CLC proteins have glycosylation consensus sites in this loop.

The broad hydrophobic region D9 to D12 (Fig. 1) is found in all CLC proteins. It is interrupted by a hydrophilic stretch of variable length, as seen in hydrophobicity plots of CIC-0 and CIC-2 through CIC-5. Protease protection experiments indicate that this stretch is cytoplasmic. The hydrophobic regions before and after this stretch are broad enough to span the membrane at least twice. However, we felt that it would be inappropriate to probe these regions of uninterrupted hydrophobicity by glycosylation scanning or protease protection assays. From studies on CIC-0 we knew that mutations in a highly conserved region in the middle of D11/D12 affect gating and rectification (M. Pusch, U. Ludewig, and T.J.J., unpublished data). Currents mediated by the L549C mutant situated in that region were sensitive to modification with MTSES, showing that this residue is accessible from the exterior. Because the stretches before D11 and after D12 are cytoplasmic, D11/D12 must span the membrane twice, or at least dip deeply into the lipid bilayer. We have not yet performed a similar analysis in the D9–D10 stretch, as we do not have any clues from mutagenesis studies as to which amino acid to choose. As the D8–D9 linker is extracellular and the segment after D10 intracellular, this segment must cross the membrane either once or three times. This implies that CLC channels have 10 (or 12) transmembrane spans in total.

The large hydrophilic C terminus is poorly conserved within the gene family and varies in length. Both sides of the weakly hydrophobic, conserved region D13 are cytoplasmic. No glycosylation occurs at the engineered site after D12 in construct 12TAC nor at the naturally occurring site after D13 [as shown for CIC-K1 and CIC-K2 (15)], and prolactin epitopes are not protected when fused either before or behind D13.

It is interesting to compare our experimental results with predictions from algorithms that use alignments of several members of a gene family as input. We used two such programs [TMAP (52) and PREDICTPROTEIN (53–55)] with CIC-0 to CIC-7, CIC-Ka, *Saccharomyces cerevisiae* scCIC, and *Escherichia coli* ecCIC as inputs. Both programs correctly indicate that D4 does not cross the membrane, but only TMAP correctly predicts that D3 does so. PREDICTPROTEIN is consistent with four transmembrane domains in the D9–D12 region, whereas TMAP lumps D9 and D10 together into one domain. Thus, only the predictions by TMAP are consistent with the experimental data.

The topology determined here is consistent with several electrophysiological findings. In CIC-0, mutation K519E next to the end of D12 causes an outward rectification consistent

with a charge reversal close to the cytoplasmic end of the pore (9). The cytoplasmic localization of this residue was confirmed by cysteine modification experiments (19). In addition, indirect evidence obtained with CIC-2 (13) is compatible with the cytoplasmic localization of D13 shown here. CIC-2 gating depends on an N-terminal inactivation domain. It could be transplanted to the C terminus either before or behind D13 without loss of function. Recent experiments have identified the D7–D8 linker as a putative receptor for this inactivation domain (22). Thus, all four regions must be on the same (cytoplasmic) side of the membrane, which is fully consistent with the present work.

However, the topology determined here contradicts some other reports. It was reported (56) that a variant of CIC-K2 lacking D2 induced currents identical to WT CIC-K2, suggesting—in clear contrast to the present data—that D2 does not span the bilayer (16). A deletion of an upstream transmembrane domain will grossly change the topology of a multispan membrane protein, predicting a loss of function. Indeed, deleting D2 in CIC-0 and CIC-5 abolishes currents (6, 12), and deleting D2 in CIC-5 causes Dent's disease (12). Thus, CIC-K currents [which could not be reproduced by others (15, 57)] should be reexamined. Additionally, a serum against a C-terminal fusion protein of a rabbit CIC-K protein was reported to inhibit currents in renal membrane vesicles when applied from the extracellular face (57). This is difficult to reconcile with the cytoplasmic location of the C terminus shown here.

In summary, we have shown experimentally that CIC-1 (and probably all CLC proteins) has 10 (or 12) transmembrane domains. Both its N and C termini reside in the cytoplasm. In contrast to a previous model based solely on hydrophathy analysis according to Kyte and Doolittle (5), D4 does not cross the membrane, but is accessible at either end from the exterior.

This work was supported by the Deutsche Forschungsgemeinschaft, the Muscular Dystrophy Association, and the Fonds der Chemischen Industrie.

- Fujita, N., Mori, H., Yura, T. & Ishihama, A. (1994) *Nucleic Acids Res.* **22**, 1637–1639.
- Greene, J. R., Brown, N. H., DiDomenico, B. J., Kaplan, J. & Eide, D. J. (1993) *Mol. Gen. Genet.* **241**, 542–553.
- Hechenberger, M., Schwappach, B., Fischer, W. N., Frommer, W. B., Jentsch, T. J. & Steinmeyer, K. (1996) *J. Biol. Chem.* **271**, 33632–33638.
- Steinmeyer, K., Ortland, C. & Jentsch, T. J. (1991) *Nature (London)* **354**, 301–304.
- Jentsch, T. J., Steinmeyer, K. & Schwarz, G. (1990) *Nature (London)* **348**, 510–514.
- Jentsch, T. J., Günther, W., Pusch, M. & Schwappach, B. (1995) *J. Physiol. (London)* **482P**, 19S–26S.
- George, A. L., Crackower, M. A., Abdalla, J. A., Hudson, A. J. & Ebers, G. C. (1993) *Nat. Genet.* **3**, 305–309.
- Koch, M. C., Steinmeyer, K., Lorenz, C., Ricker, K., Wolf, F., Otto, M., Zoll, B., Lehmann Horn, F., Grzeschik, K. H. & Jentsch, T. J. (1992) *Science* **257**, 797–800.
- Pusch, M., Steinmeyer, K., Koch, M. C. & Jentsch, T. J. (1995) *Neuron* **15**, 1–20.
- Steinmeyer, K., Klocke, R., Ortland, C., Gronemeier, M., Jockusch, H., Gründer, S. & Jentsch, T. J. (1991) *Nature (London)* **354**, 304–308.
- Steinmeyer, K., Lorenz, C., Pusch, M., Koch, M. C. & Jentsch, T. J. (1994) *EMBO J.* **13**, 737–743.
- Lloyd, S. E., Pearce, S. H., Fisher, S. E., Steinmeyer, K., Schwappach, B., Scheinman, S. J., Harding, B., Bolino, A., Devoto, M., Goodyer, P., Rigden, S. P., Wrong, O., Jentsch, T. J., Craig, I. W. & Thakker, R. V. (1996) *Nature (London)* **379**, 445–449.
- Gründer, S., Thiemann, A., Pusch, M. & Jentsch, T. J. (1992) *Nature (London)* **360**, 759–762.
- Staley, K. J., Smith, R., Schaak, J., Wilcox, C. & Jentsch, T. J. (1996) *Neuron* **17**, 543–551.
- Kieferle, S., Fong, P., Bens, M., Vandewalle, A. & Jentsch, T. J. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 6943–6947.
- Middleton, R. E., Pheasant, D. J. & Miller, C. (1994) *Biochemistry* **33**, 13189–13198.
- Lorenz, C., Pusch, M. & Jentsch, T. J. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 13362–13366.
- Ludewig, U., Pusch, M. & Jentsch, T. J. (1996) *Nature (London)* **383**, 340–343.
- Middleton, R. E., Pheasant, D. J. & Miller, C. (1996) *Nature (London)* **383**, 337–340.
- Fahlke, C., Knittel, T., Gurnett, C. A., Campbell, K. P. & George, A. L. (1997) *J. Gen. Physiol.* **109**, 93–104.
- Pusch, M., Ludewig, U., Rehfeldt, A. & Jentsch, T. J. (1995) *Nature (London)* **373**, 527–531.
- Jordt, S. E. & Jentsch, T. J. (1997) *EMBO J.* **16**, 1582–1592.
- Chavez, R. A. & Hall, Z. W. (1991) *J. Biol. Chem.* **266**, 15532–15538.
- Chang, X. B., Hou, Y. X., Jensen, T. J. & Riordan, J. R. (1994) *J. Biol. Chem.* **269**, 18572–18575.
- Turk, E., Kerner, C. J., Lostao, M. P. & Wright, E. M. (1996) *J. Biol. Chem.* **271**, 1925–1934.
- Chavez, R. A. & Hall, Z. W. (1992) *J. Cell Biol.* **116**, 385–393.
- Skach, W. R. & Lingappa, V. R. (1994) *Cancer Res.* **54**, 3202–3209.
- Akabas, M. H., Kaufmann, C., Cook, T. A. & Archdeacon, P. (1994) *J. Biol. Chem.* **269**, 14865–14868.
- Kürz, L. L., Zühlke, R. D., Zhang, H. J. & Joho, R. H. (1995) *Biophys. J.* **68**, 900–905.
- Perez-Garcia, M. T., Chiamvimonvat, N., Marban, E. & Tomaselli, G. F. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 300–304.
- Higuchi, R. (1990) in *PCR Protocols*, eds. Innis, M. A., Gelfand, D. H., Sinsky, J. J. & White, T. J. (Academic, San Diego), pp. 177–183.
- Nilsson, I. M. & von Heijne, G. (1993) *J. Biol. Chem.* **268**, 5798–5801.
- Landolt-Marticorena, C. & Reithmeier, R. A. (1994) *Biochem. J.* **302**, 253–260.
- Hresko, R. C., Kruse, M., Strube, M. & Mueckler, M. (1994) *J. Biol. Chem.* **269**, 20482–20488.
- Perara, E. & Lingappa, V. R. (1985) *J. Cell Biol.* **101**, 2292–2301.
- Rothman, R. E., Andrews, D. W., Calayag, M. C. & Lingappa, V. R. (1988) *J. Biol. Chem.* **263**, 10470–10480.
- Skach, W. R. & Lingappa, V. R. (1993) *J. Biol. Chem.* **268**, 23552–23561.
- Akabas, M. H., Stauffer, D. A., Xu, M. & Karlin, A. (1992) *Science* **258**, 307–310.
- Kuner, T., Wollmuth, L. P., Karlin, A., Seeburg, P. H. & Sakmann, B. (1996) *Neuron* **17**, 343–352.
- Pascual, J. M., Shieh, C. C., Kirsch, G. E. & Brown, A. M. (1995) *Neuron* **14**, 1055–1063.
- Akabas, M. H. & Karlin, A. (1995) *Biochemistry* **34**, 12496–12500.
- Cheung, M. & Akabas, M. H. (1996) *Biophys. J.* **70**, 2688–2695.
- Loo, T. W. & Clarke, D. M. (1995) *J. Biol. Chem.* **270**, 843–848.
- Traxler, B., Boyd, D. & Beckwith, J. (1993) *J. Membr. Biol.* **132**, 1–11.
- Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132.
- von Heijne, G. & Gavel, Y. (1988) *Eur. J. Biochem.* **174**, 671–678.
- Sipos, L. & von Heijne, G. (1993) *Eur. J. Biochem.* **213**, 1333–40.
- Skach, W. R., Shi, L. B., Calayag, M. C., Frigeri, A., Lingappa, V. R. & Verkman, A. S. (1994) *J. Cell Biol.* **125**, 803–815.
- Henn, D. K., Baumann, A. & Kaupp, U. B. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 7425–7429.
- Lin, J. & Addison, R. (1995) *J. Biol. Chem.* **270**, 6935–6941.
- Jentsch, T. J., Günther, W., Pusch, M. & Schwappach, B. (1995) *J. Physiol. (London)* **482**, 19S–25S.
- Persson, B. & Argos, P. (1994) *J. Mol. Biol.* **237**, 182–192.
- Rost, B., Casadio, R., Fariselli, P. & Sander, C. (1995) *Protein Sci.* **4**, 521–533.
- Rost, B. (1996) *Methods Enzymol.* **266**, 525–539.
- Rost, B., Fariselli, B. & Casadio, R. (1996) *Protein Sci.* **7**, 1704–1718.
- Adachi, S., Uchida, S., Ito, H., Hata, M., Hiroe, M., Marumo, F. & Sasaki, S. (1994) *J. Biol. Chem.* **269**, 17677–17683.
- Zimniak, L., Winters, C. J., Reeves, W. B. & Andreoli, T. E. (1996) *Am. J. Physiol.* **270**, F1066–F1072.