ELSEVIER

Topological model of membrane domain of the cystic fibrosis transmembrane conductance regulator

Xavier Gallet,* Franck Festy,† Philippe Ducarme,* Robert Brasseur,* and Annick Thomas-Soumarmon†

**Faculté des Sciences Agronomiques, Centre de Biophysique Moléculaire Numérique, Gembloux, Belgium*
†Hôpital Bichat-Claude Bernard, INSERM U10, Paris, France

The cystic fibrosis transmembrane conductance regulator is a cAMP-regulated chloride channel. We used molecular modelling to predict 3-D models for the CFTR membrane domain. Hydropathy and residue conservation in all CFTRs as well as in other proteins suggested that the membrane domain is a 12-helix bundle. If the domain is enclosing a channel for chloride, it could be made of five helices. We propose two structural models in which both lumenal and cytoplasmic entrances to the chloride pore have a ring of positively charged residues. The inner surface of the channel is covered with neutral polar plus one or two charged residues. Helices that are not directly involved in the chloride channel could organise to form a second channel; a dimeric symmetrical structure is proposed. Analysis raised interest for helix 5: this hydrophobic fragment is conserved in all CFTRs and aligns with segments present in several different ion channels and transporters. The existence of an FFXXFFXXF motif is proposed. Helix 5 could be an important domain of CFTRs. The models agree with available data from pathological mutations but does not account for the membrane insertion of a hydrophilic fragment of NBD1. © *1998 by Elsevier Science Inc.*

Keywords: channel protein, homology, molecular modelling, sequence alignment

INTRODUCTION

Cystic fibrosis (CF) is a lethal genetic disease[1] caused by mutations of the CF gene,[2] which encodes the cystic fibrosis transmembrane conductance regulator (CFTR).[3,4] The CFTR is a large glycoprotein containing two repetitive regions, each formed by a membrane-spanning domain (MSD) plus a nucleotide-binding domain (NBD). The two symmetrical parts (MSD1–NBD1/MSD2–NBD2) are separated by a regulatory (R) domain containing several sites potentially phosphorylated by protein kinases A and C.[5,6] The symmetry suggests that CFTRs were produced by duplication of an ancestral gene. More than 500 different mutations have been reported[1] and the most frequent (70%) is the deletion of the phenylalanine codon (508) in NBD1 (ΔF508). This deletion affects the maturation and the translocation of the CFTR to the apical membrane of epithelial cells. The mutated protein (ΔF508-CFTR) is blocked in the endoplasmic reticulum compartment, where it is retained by chaperones.[7]

The CFTR functions as a cyclic AMP-activated chloride channel[8] regulated by direct ATP hydrolysis and kinase-mediated phosphorylations. Reconstitution of short synthetic fragments of MSD1 (helices 2 and 6) in bilayers generates a chloride conductance, suggesting they are involved in the functioning of the chloride channel. Similar results also implicate NBD1.[9] More recent studies have suggested that the CFTR also excretes ATP, which, by an autocrine mechanism, activates the outwardly rectifying chloride channels (ORCC) via a P_{2U} purinergic receptor and a GTP-binding protein.[10–12] Thus, the CFTR may be an ATP/Cl^- channel and, through its ATP channel activity, a regulator of cell ionic conductance. However, the ATP transport function remains controversial.[13,14] A protein–protein interaction mechanism has been suggested to explain the activation of amiloride-sensitive epithelial Na^+ channels (ENaCs) when the apical membrane lacks CFTRs, and the Na^+ reabsorption in CF patients.[15] Other hypotheses have been proposed to explain how the CFTR might regulate Ca^{2+}-activated Cl^- channels[16] or K^+ channels.[17]

NBD sequences share similarities with domains of yeast transporters, bacterial permeases, P-glycoproteins, and adenylate kinase. Thus, the CFTR belongs to the ATP-binding cassette (ABC) family. By homology with adenylate kinase, 3-D models of the NBD have been proposed.[18,19]

To understand the properties of the CFTR, a 3-D model of the membrane domain would be valuable. However, experi-

Color Plates for this article are on pages 97–98.

Address reprint requests to: X. Gallet, Faculté des Sciences Agronomiques, Centre de Biophysique Moléculaire Numérique, 5030 Gembloux, Belgium.

Journal of Molecular Graphics and Modelling 16, 72–82, 1998

655 Avenue of the Americas, New York, NY 10010

1093-3263/98/$19.00
PII S1093-3263(98)00015-1

mental analysis of CFTR structure is difficult because of its large size, hydrophobicity, and weak expression in tissues. As a consequence, little structural information is available. In previous reports, hydropathy and residue variability were used to develop models of cation channels,[20] receptors,[21] and P-ATPase.[22] In this study, we report molecular modelling of the 3-D topology of the MSD1 and MSD2 domains of the CFTR. Twelve transmembrane helices were selected according to the hydropathy profile of the primary sequence. To determine the relative position and orientation of helices in the 3-D model and to identify which helices could be involved in chloride or other putative substrate transport, protein sequences in the NBRF databank were searched for similarity with CFTR sequences. Significant identities were found between all CFTRs and between human CFTR and chloride channels. Those data support the hypothesis that the chloride channel could involve five helices. Other similarities were found with ATP/ADP carriers (e.g., nucleoside and amino acid transporters). We propose 3-D models describing a chloride channel with hydrophilic inner surfaces; we discuss the existence of a separate ATP channel. The models agree with most available data, including results from mutagenesis and patch-clamp investigations.

MATERIALS AND METHODS

Prediction of the membrane-spanning segments

1. The algorithm of Kyte and Doolittle[23] was used to calculate the hydropathy profile of the CFTR sequence. Average hydrophobicities $\langle H_i \rangle$ are calculated along the sequence, using windows of 11 residues, and then assigned to the amino acid i, in the centre of the window. Segments are selected when $\langle H_i \rangle > 1$, using the Kyte and Doolittle hydrophobicity scale. $\langle H_i \rangle$ values and window size were calibrated as those providing the best detection of transmembrane domains of bacteriorhodopsin and subunits of the bacterial photosynthetic reaction centre.

2. The algorithm of Eisenberg[24,25] was used to plot the hydrophobic moment $\langle \mu \rangle$ versus hydrophobicity $\langle H \rangle$, with the Eisenberg consensus hydrophobicity scale and a window of 11 residues ($N = 11$):

$$\langle H_i \rangle = 1/n\left(\sum_{n=1}^{N} h_n \right)$$

$$\langle \mu_i \rangle = \left\{ \left[\sum_{n=1}^{N} h_n \sin(n\delta) \right]^2 + \left[\sum_{n=1}^{N} h_n \cos(n\delta) \right]^2 \right\}^{1/2}$$

h is the hydrophobicity value of the amino acid n according to the Eisenberg's consensus hydrophobicity scale. The average values $\langle H_i \rangle$ and $\langle \mu_i \rangle$ are assigned to the amino acid i in the centre of the window. The gyration angle δ is 100° to calculate the hydrophobic moment μ for an α-helical structure. Transmembrane hydrophobic or amphiphilic segments were selected for $\langle H_i \rangle > 0.5$; $\langle \mu_i \rangle < 0.4$ and $\langle H_i \rangle > 0$; $\langle \mu_i \rangle > 0.3$, respectively.

Calculation of the structures of transmembrane segments and loops

The sequence corresponding to two close transmembrane helices and the loop in between was selected. In that fragment, the secondary structure of each transmembrane helix was assigned by attributing ϕ and ψ values of a right-handed α helix ($\phi = -65°$, $\psi = -42°$) to all residues and helices were then energy refined using the Simplex procedure.[26] This program uses an energy equation that accounts for interactions with solvent.[27,28] A hydrophobic solvent (methyl group) was used to mimic the hydrophobicity of the membrane. Then, the structures of the short stretch between the two helices was calculated using a Monte Carlo procedure,[29] starting from an initial coiled structure ($\phi = 0°$, $\psi = 180°$). Ten thousand steps of random 5° rotations of all dihedral angles (ϕ, ψ) of the loop residues were calculated and the five conformations of minimal energy were recorded and examined. The structure that matched the helices anti-parallel, such that they can both traverse the membrane at best, was saved. When no structure fitted this requirement, another sequence of loop was tested either by shifting one residue left or right or by adding one more residue. When several conformations fitted, the most hydrophilic loop was preferred.

Orientation of α helices in the membrane

The orientation of each helix in the membrane complex was predicted according to two criteria: mutations and hydrophobicity.

Mutations

1. Six sequences of the CFTR (human, bovine, rabbit, mouse, frog, and dogfish) were aligned using CLUSTAL.[30] Conserved amino acids of membrane domains were mapped on the helix structure. A *conservation vector* indicating the most conserved side of a transmembrane segment was calculated as described below.

2. The CFTR transmembrane segments were compared with all sequences in the National Biomedical Research Foundation (NBRF) databank, using FASTA.[31] The 700 most similar fragments of sequences were selected and biological activities of the proteins were taken into account: those eliciting transport functions for ions, nucleosides, peptides, sugars, and organic compounds were retained. Conserved residues were mapped on the helix structure.

3. The transmembrane segments were aligned using CLUSTAL, with the sequences of proteins of the ABC family (human MDR1 and -3; *Plasmodium falciparum* MDR1 and -2; and human MRP-1, STE6, OppB/C, and HisM/Q).

Hydrophobicity

Molecular hydrophobic potentials (MHPs)[32] were calculated to visualise the surfaces of hydrophilic and hydrophobic isopotentials around the molecule. Hydrophilic faces of helices were candidates for a central pore location.

Conservation and amphiphilicity vectors were calculated by using the Fourier transform FT[33]:

$$\mathrm{FT} = \left\{ \left[\sum_{i=1}^{N} A_i \sin(i\delta) \right]^2 + \left[\sum_{i=1}^{N} A_i \cos(i\delta) \right]^2 \right\}^{1/2}$$

For the *amphiphilicity vector*, A_i is the hydrophobicity value of the amino acid i according to the Fauchère scale.[34] For the conservation vector, A_i is the number of mutations of the amino acid i in the six CFTR sequences. Both vectors were calculated assuming an α-helical structure ($\delta = 100°$) and a window of 17 residues ($N = 17$) centred on the transmembrane sequence.

Calculation of the 3-D model

Helices and pairs of helices were matched together at best to make a pore, as specified in Results and Discussion. A matched structure with no contact was prepared and then energy minimised by calculating the energies of 2 500 positions close to the initial position, but assuming 2.5° rotations and 2.5-Å vertical and horizontal translations, respectively, using the empirical equation previously published.[27,28] The final structure corresponded to the largest gain in energy due to interactions between helices, assuming that the bathing medium was hydrophobic (methyl group).

All calculations and molecule visualisations were performed on PC 586 microcomputers, using PC-PROT+ (Protein Computing Programs), PC-TAMMO+ (Theoretical Analysis of Molecular Membrane Organisation)[26] and WinMGM software[35] from Ab Initio Technology (Obernai, France).

RESULTS AND DISCUSSION

Prediction of the 2-D model

Hydrophobic fragments (8 to 41 residues) were selected using the Kyte and Doolittle algorithm. Since channel proteins can have a central hydrophilic pore, the method of Eisenberg was used to further detect amphipathic hydrophobic stretches. Those segments overlap and complete previous predictions. The membrane-associated regions map in two domains of the protein. The first, MSD1, contains approximately residues R80 to V350 and the second, MSD2, contains L850 to I1150.

We then predicted the secondary structures of transmembrane fragments. Only a few membrane protein structures have been resolved. Some, such as bacteriorhodopsin,[36] are helix bundles, whereas porin is made of 16 β strands forming a β barrel.[37] The list is insufficient to assume that we know all possible intramembranous motifs.[38] For instance, mixed α/β structures have been proposed for the nicotinic acetylcholine receptor[39] and for the glucose transporter GLUT1,[40] but those predictions are not yet experimentally confirmed.

α Helices and β strands require about 18 to 20 and 8 to 10 residues to span a lipid bilayer, respectively. Therefore, membranous α helices are easier to predict from sequences because they are larger hydrophobic clusters, more easily identifiable in hydropathy plots. Helix backbones self-stabilise by internal hydrogen bonds and outer surfaces of helices are hydrophobic when the side chains of the constituting amino acids are hydrophobic. In contrast, a β strand should be unstable in a lipid bilayer because of polar carbonyl and amine groups protruding on the side of the strand. They will have to form hydrogen bonds with neighbouring strands to make sheets and thus become stabilised in a hydrophobic environment. Therefore, predicting a transmembrane β strand should, according to our present knowledge of the structure of membrane proteins, imply proposing several of them in order to make a β-barrel structure.

Hydrophobic clusters of CFTR residues are large; therefore we have assumed that the secondary structure is helical. Other results supporting that choice will be reported later on. Some clusters are long enough to form a hairpin, with two helices traversing the membrane and a loop in between. The minimal sequences of loops required to match both helices anti-parallel are W216–Q220 and Y325–K329 in MSD1, and A1009–P1013 and E1124–R1128 in MSD2. Those loops should all be lumenal, in agreement with the glycosylation of adequate sites inserted at those levels.[41] Their short sizes suggest that they should remain buried among polar heads of lipids. The two last lumenal loops should be longer: one, between helices 7 and 8 (W882–S911), has two glycosylation sites (N894 and N900). The last (P99–R117), between the two first helices, is a "receptor-binding site" according to De Loof et al.[42] and should be a good antigenic site.[43,44] All cytoplasmic loops are much longer, from 62 to 511 residues (NBD1 plus R).

In the 2-D model (Color Plate 1), each membrane-spanning domain (MSD) has six transmembrane helices, the lengths of which vary from 17 (H2, H5, H6, H10, and H12) to 24 (H7) amino acids. The equivalent number of membrane helices in MSD1 and MSD2 and, more generally, the symmetry of the structure MSD1–NBD1/MSD2–NBD2 support gene duplication. Eighty percent of the CFTR residues are cytoplasmic, 15% are intramembranous, and less than 5% are extracellular.

To propose a model of the membrane domain of the CFTR, we analysed three properties of transmembrane helices: conservation in CFTRs, amphipathy, and homology with other transporters.

Conservation in CFTRs The sequences of six CFTRs were aligned to identify conserved membrane residues. Lipid-facing residues do not need to be as well conserved as those facing other helices, where steric and charge complementation are required for a compact structure. Furthermore, the inner-channel residues should be the most conserved since they may be functionally involved.

When all CFTRs are compared, the most conserved transmembrane helix is H5, with 15 residues of 17 identical. (Table 1, Figure 1A). This suggests that helix 5 has an

Table 1. Numeric description of membranes helices[a]

	1	2	3	4	5	6	7	8	9	10	11	12
Size (aa)	18	17	21	21	17	17	24	21	20	17	21	17
Number of mutations (aa)	5	5	6	10	2	5	15	9	9	10	10	4
Conservation ratio	0.72	0.70	0.71	0.52	0.88	0.70	0.40	0.57	0.55	0.41	0.52	0.71
Conservation vector magnitude	0.18	0.22	0.15	0.29	0.21	0.29	0.23	0.15	0.51	0.09	0.13	0.14
Hydrophobicity vector magnitude	1.66	0.35	0.28	0.64	1.01	1.62	1.92	1.12	1.83	0.86	1.56	1.77

[a] Size refers to the number of residues in the membrane; *number of mutations* refers to the number of mutated residues after alignments of CFTR sequences. The conservation ratio plots the number of mutations relative to size. The magnitude of conservation and amphiphilicity vectors are obtained as described in Materials and Methods.

important role in CFTRs. Unexpectedly, if the CFTR sequence is a gene duplication, helix 11, the fifth helix of MSD2, has only 11 of 21 residues conserved. Helix 12 has a good conservation score, with 13 of 17 residues identical. Its corresponding helix in MSD1, helix 6, has 12 of 17 residues identical. On the other hand, helices 7 and 10 are the most mutated fragments, with only 9 and 7 conserved residues out of 24 and 17, respectively. However, mutated residues are often replaced by amino acids with similar physicochemical properties. Owing to their lower level of conservation, we suggest that helices 7 and 10 should be the first choice for interacting with lipids in the 3-D model.

For many helices, conserved and variable positions are on opposite sides: this partition is evident in helix H9 (Figure 1B) but exists also in helices H1, H2, H3, H5, H6, H8, and H12. This is consistent with α-helix secondary structures with the conserved sides pointing inside the protein. Helices H4 and H7 have a patch of mutations in the middle of conserved clusters.

Amphipathy Ion channels can be hydrophilic holes crossing through membrane bilayers. The most hydrophilic transmembrane sequence is helix 6, with several polar residues (R334, K335, T338, T339, and S341). The most amphipathic helices, i.e., those eliciting the best partition of hydrophobicity over the helical structure, are helices H1, H6, H7, H9, H11, and H12 (Table 1). An *amphiphilicity vector* was calculated to point out the less hydrophobic side.

In most cases (helices H1, H2, H4, H5, H8, H9, and H10), conservation and amphiphilicity vectors point in the same direction, diverging by less than 30°. For helix 12, the two vectors point in almost opposite directions. In helix 6, they diverge by 80°. In both cases, the low levels of mutations could be responsible for the divergence. Thus, since the conservation vectors of helices 6 and 12 have low amplitude, orientations of the amphiphilicity vectors were considered as more significant in building model II.

Sequence homologies We compared each transmembrane helix with sequences of other chloride channels in protein databanks. Significant identities were found between helices H1, H2, H5, H6, H10, H12 and fragments of chloride channels, band 3 anion transport proteins, $GABA_A$, and glycine receptors (Table 2A).

All together, data on sequence conservation and amphipathy profiles support the possible involvement of helices H1, H5, H6 and H12 in a chloride channel.

During the homology search of sequences, identities were also found between helices H2, H4, H5, H7, H8, and H12 and ATP and nucleoside transporters (Table 2B). Helices H4, H7, and H8 have homologies with one or several fragments of ATP and nucleoside transporters but none with chloride channels. Therefore, if the CFTR transports ATP,[10,11] the ATP channel could be separate from the chloride channel.

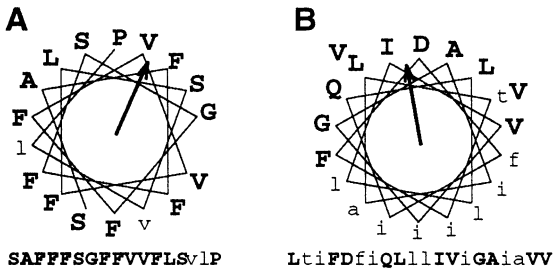

Figure 1. Helical wheel plots of the CFTR-conserved residues (boldface) in the H5 (A) and H9 (B) helices. Below, the corresponding sequences are given. (A) The helical wheel plot of H5 residues shows only two mutations (V322, L323) in CFTRs with respect to the human sequence. (B) For the helical wheel plot of H9 residues, mutated and conserved residues are clearly separated. The arrows show the orientation of the conservation vector (Table 1).

Features of the H5 helix

Since the conserved residues of helix 5 of the CFTR are located all around the helix, we suggest that it could be centrally located in the 3-D structure. Helix 5 also has residues in common with fragments of both chloride channels and ATP and nucleoside transporters. Those residues are on the same side of the helix wheel. Therefore if a separate ATP channel exists in the CFTR, a rotation movement should be necessary to switch conserved residues from the chloride to the ATP pore. This supports a sequential rather than simultaneous functioning of both forms of anion transport. It also supports the central location of helix 5 in the 3-D structure. Among all ABC proteins, the H5 sequence is similar to hMDR1, hMDR3, and HisM sequences but only weakly conserved relative to pfMDR1, pfMDR2, STE6, OppC/D, and HisQ sequences.

On the other hand, the conservation of helix 5 could reflect other properties. Helix 5 shares similarities with fragments of many other transporters (Table 3): 29 transporters or ion channels have fragments with more than 35% identity in a FASTA screening. This is the largest score compared with all other helices. Similarities with transmembrane helices of bovine cytochrome *c* oxidase (third helix of subunit 3)[45] and photosynthetic reaction centre from *Rhodobacter sphaeroides* (first helix of the L chain),[46] whose helical structures are known, support the helical secondary structure that we propose. In the cytochrome *c* oxidase from *Paracoccus denitrificans*, the helix 5-like segment is in helix 3 of subunit III (seven transmembrane helices). Helix 3 is specifically bound to a phospholipid and the cleft of cytochrome *c* oxidase subunit III is suggested to be a docking site for other membrane proteins.[47] Helix 5 is short (17 residues), but encloses 6 phenylalanines, 5 of which are located on the same side of the helix wheel. That FFXXFFXXF signature, tested on the nonredundant OWL bank of 165 129 sequences, extracted 177 proteins, among which were all of the CFTRs and 41 cytochrome *c* oxidases. Twenty-six nonredundant fragments of sequence are reported in Table 4.

Description of the 3-D models

Several pairs of helices (H3/H4, H5/H6, H9/H10, and H11/H12) were matched because they are separated by short extracytoplasmic loops and are therefore presumably close in the membrane. Helices 1 and 2 were also matched even if the loop in between is 18 residues. Helix couples were then associated. We propose two models: I and II.

Model I (Figure 2) is a model of the chloride channel

Table 2. Sequence alignments of chloride channels, ATP/ADP carriers, and nucleoside transporter proteins with the CFTR transmembrane helices[a]

Helices/proteins of bank	Sequence alignment	Identity (%)
Helix 1	W R F M F Y G I F L Y L G E V T K A V Q P L	id
A42497 AE3 Cl-/HCO3- exchanger - rat	P L A V L F G I F L Y M G V T S L S G I Q L	33 ✓
S59861 band 3 anion transport protein - trout	P M T A L F G I F L Y M G I T S L S G I Q M	33
Helix 2	E R S I A I Y L G I G L C L L F I V R T L	
S12382 glycine receptor alpha-1 chain - human	S Y V K A I D I W M A V C L L F V F S A L	35 ✓
S34469 GABA receptor A beta chain - snail	S Y T K A I D V W M S T C L V F V F A A L	29
Helix 5	N S S A F F F S G F F V V F L S V L P Y A	
A38079 GABA receptor rho-2 chain - human	A V D I Y L W V S F V F V F L S V L E Y A	41 ✓
JH0794 GABAa receptor - fruit fly	S R L A C E I Q C F V M V F A S L L E Y A	35
S21086 anion exchange protein 2 - hum.	Y L V G R V W I G F W L V F L A L L M V A	35
A41145 GABA receptor A - fruit fly	S I D V Y L G T C F V M V F A S L L E Y A	29
A36218 band 3 anion transport protein - human	Y I V G R V W I G F W L I L L V V L V V A	29
Helix 6	I K G I I L R K I F T T I S F C I V L R M	
S24318 band 3 anion transport protein - trout	I L T I P L R M A I T G T L F T D K E M K	29 ✓
Helix 10	Q P Y I F V A T V P V I V A F I M L R A Y	
A42497 AE3 Cl-/HCO3- exchanger - rat	L E S V L E V P V P V R F L F V M L G P S	35
A30816 band 3 anion transport protein - chicken	I W M M F A S P V P A L L V F I L I F L E	29
Helix 12	G R V G I I L T L A M N I M S T L Q W A V	
JH0824 GABAa receptor gamma 1 - chicken	G R V L Y T L R L T I N A E C Y L Q L H N	35
JH0316 GABAa receptor gamma 2 - mouse	G R V L Y T L R L T I D A E C Q L Q L H N	29 ✓
B		
Helix 2	E R S I A I Y L G I G L C L L F I V R T L	
S30259 ADP,ATP carrier protein - C. reinhardtii	G R L A S P Y K G I G E C F V R T V R E E	41
Helix 4	L Q A S A F C G L G F L I V L A L F Q A G L G R M	
A54892 Na+-dependent nucleoside transport prot.- rat	S W R A V S W G L G L Q F V L G L F V I R T E P G	33
GRBYCP cytosine/purine transport protein - yeast	I I F F N I M G L I F V A F F S V F G A E L G L R	33
A56382 purine permease - E. nidulans	W G S A E F I G L G F L V F A T I I L C E R F G S	28
Helix 5	N S S A F F F S G F F V V F L S V L P Y A	
JQ0026 ADP,ATP carrier protein - R .prowazekii	Y V I T S F F L G Y F A L F A F V L Y P Y	41 ✓
D64111 uracil permease - H. influenzae	W A I A I S F C G K V G A F L S T I P T I	41
S45776 uracil permease - yeast	N L T N F E F M C F M V F W V A C L P F L	35
GRBYCP cytosine/purine transport protein - yeast	N I M G L I F V A F F S V F G A E L G L R	35
S45877 uracil transport protein - yeast	N A T T Y E F M C F F I F W A A S L P F L	29
Helix 7	H K S L I F V L I W C L V I F L A E V A A S L V V L W L	
S38893 uracil transport protein - B. caldolyticus	M L S L I P V L V G I V V G Y L Y A L A V G L V D L S K	46
A57532 Na+-dependent nucleoside transport prot.- rat	R R A L A L F V I T C L V I F I L A C H F L K K F F A K	29
S45877 uracil transport protein - yeast	L A A Y L C G V A P C L P G F I A E V G A P A I K V S D	29 ✓
Helix 8	T S S Y Y V F Y I Y V G V A D T L L A M G F F R G	
S22661 cytosine transport prot. - E. coli	G L S Y H D F F L A V L I G N L L L G I Y T S F L	24 ✓
Helix 12	G R V G I I L T L A M N I M S T L Q W A V	
A56265 uracil transport prot. - E. coli	V L V G Y A L S F A M G I V D T T P I I N	41 ✓

[a] Predicted CFTR transmembrane sequences are boxed. Other sequences are extracted from the NBRF databank by FASTA analysis. Databank accession numbers are presented at the beginning of each line. Identical residues are shown by black boxes. Amino acids with similar properties are lightly stippled. Percentages of identity (id) are reported on the right. The check marks after identity values indicate sequences used to report conserved amino acids in Figures 2 and 3.

constructed using the following hypothesis. From the preceding data, we suggest that helices H1, H5, H6, and H12 could be involved in the chloride channel. Measurements of chloride permeability after insertion of synthetic helices in lipid bilayers suggest that helix 6 is directly involved,[48] as is helix 2; this did not emerge from our analysis. Model I was thus made by energy minimisation of the matching of three helix couples: H1/H2, H5/H6, and H11/H12. Helix couples were manually

Table 3. Alignment of Helix 5 with sequences from the PDB databank[a]

Proteins of bank	Sequence alignment	Identity (%)
Helix 5	N S S A F F F S G F F V V F L S V L P Y A	id
1. A42970 H+-transporting ATPase - yeast	I Q I A F G F R G F V G V F M T V A L F A	47
PWECBK H+/K+-exchanging ATPase -E. coli	I L S A V I F N A L I I V F L I P L A L K	41
I54168 calcium channel protein alpha-1 chain - human	I I I A F F M M N I F V G F V I V T F Q E	41
E64245 periplasmic phosphate permease - M. genitalium	S A F A L G I S K R E V I F K I V L P S A	41
A41594 Na+/H+ antiporter - Bacillus firmus	V S G A F F A S T L G V D V I D Y L P F A	41
A64072 ATP synthase F0 subunit a (atpB) - H. influenza	T L Q A F I F M M L T V V Y L S I A Y N K	41
S15795 vacuolar H+ pump - C. elegans	E I F Q T F F G G R Y V I F L M G A F S I	35
A53257 H+-transporting ATP synthase - N. fowleri	D R T K M F F P F F F Y L F L F I C L S N	35
S14209 H+-transporting ATP synthase - starfish	I I T T V F I L L F S V N V L G L L P Y A	35
S17719 H+-transporting ATP synthase - B. firmus	V Q F L S F F A G L S V S F L N L W T T Y	35
S42658 H+-transporting ATP synthase - radish	N V K Q M F F P C I L V T F L F L L F C N	35
GRECNK nitrate transport protein narK - E. coli	D G Q G G S F M A F F A V F L A L F L T A	35
2. S47892 neutral amino acid permease - N. crassa	I C S A L F I S G F S F Y F P A L M Y F K	41
S50711 basic-amino-acid permease - yeast	N A S D F F T A Y I S V I L F V V L W V G	35
3. S25009 monosaccharid transport protein STP4 - A. thaliana	K F G L F F F F A F F V V I M T I F I Y L	41
C41853 hexose phosphate transport protein - S. typhimurium	Q A A C F F T T G F F V F G P Q M L I G M	41
S25015 monosaccharide transport protein MST1 - tobacco	K F G L F L F F A F F V V I M T V F I Y F	41
S06920 glucose transport protein, hepatic - mouse	M I G M F F C T I F M S V G L V L L D K F	35
S12042 glucose transport protein STP1 - A.thaliana	K F G L F L V F A F F V V V M S I F V Y I	35
GREC lactose permease - E. coli	L V C F C F F K Q L A M I F M S V L A G N	35
4. A55630 aquaporin-5, salivary gland - rat	V C S L A F F K A V F A E F L A T L I F V	41
A43319 chromaffin granule amine transporter - rat	G Y H I P M F V G F M I M F L S T L M F A	41
B41044 octopine permease protein M - A. tumefaciens	P F D P A F L W Q T F V A L L S G I P L A	35
E64112 bicyclomycin resistance protein - H. influenza	F T A L R F V Q G F F G A A P V V L S G A	35
5. DVHU1 multidrug resistance protein 1 - human	Q S M A T F F T G F I V G F T R G W K L T	41
DVHU3 multidrug resistance protein 3 - human	Q A V A T F F A G F I V G F I R G W K L T	41
MMEBMT Histidine permease M - S. typhimurium	A A R A Y G F S S F K M Y R C I I L P S A	35
6. 1OCC cytochrome c oxidase (chain L) - bovine	M M T L F F G S G F A A P F F I V R H Q L	41
1PRC photosynthetic reaction center (chain L) - Rh. viridis	F V G F F G V S A I F F I F L G V S L I G	35

[a] (1) Ion channels; (2) peptide/amino acid transporters; (3) sugar transporters; (4) organic substrates transporters; (5) ABC proteins; (6) membrane proteins from the PDB databank. The sequence of H5 is boxed. Sequences are from the NBRF databank, except for the photosynthetic reaction centre and the cytochrome *c* oxidase, which are from the PDB databank. Accession numbers are listed at the beginning of each line. Identical residues are shown by black boxes. Residues with similar properties are lightly stippled. Percentages of identity (id) are reported on the right.

arranged so that (1) selected helices made the channel, (2) conservation and amphiphilicity vectors were oriented towards the channel centre, and (3) helices were close but not in contact. The energy of the construct was then minimised by shifting and tilting helices to generate interactions. When no interaction was created, no minimisation occurred thus no solution was found. Several manual arrangements were tested and the construct with the largest energy minimisation was selected. This procedure does not give an exhaustive overview of all possibilities since it starts from restrictive hypotheses. However, according to our limited knowledge of membrane protein structures, and according to the few experimental data on CFTR structure, it would seem unrealistic to exhaustively test all possibilities.

Model II (Figure 3) accounts for all helices of both MSDs. The model is based on the assumption that (1) the CFTR evolved by gene duplication: MSD1 and MSD2 structures must be symmetrical; (2) since most helices involved in the chloride channel are from MSD1,[49] then the chloride channel must be in MSD1. Some helices of MSD2 share homology with nucleoside transporters: the ATP channel, if it exists, must be in MSD2. In practice, helices were oriented to join the conserved residues inside each pore, conservation being with respect to chloride channels and nucleoside transporters for MSD1 and MSD2, respectively. The most conserved helix, H5, was positioned first, in the centre. Its symmetrical helix of MSD2 (H11) was added. Helices of MSD1 were then organised to build the chloride channel as in model I. The ATP channel was symmetrically adapted. Helices H3/H4 and H9/H10, which are not suggested to be involved in ion transport, were placed so as to attain the most compact model, with the most conserved face oriented towards the pore. The final structure was energy refined as described for model I.

An enlarged view of the inside of each chloride channel (Color Plate 2A and B) shows numerous positive charges gathered at the extracellular (R117, K329 for models I and II plus R1128 for model I) and cytoplasmic (R80, R134, and R347 for models I and II plus R1102 in model II) ends of the pores. Those charges could fulfill several functions, including

Table 4. Alignment of the FFXXFFXXF signature[a]

CFTR (F311-F319)	311	F	F	S	G	F	F	V	V	F
A45612 ATPase 6 - Sauroleishmania tarentolae	166	F	F	I	I	F	F	L	F	F
ACCOX3G A.crispulus cox3 gene - Anthoceros crispulus	78	F	F	L	A	F	F	W	A	F
ASP_PLAFS aspartic acid-rich protein precuror - P. Falciparum	8	F	F	F	F	F	F	F	F	F
ATP6_NAEFO ATP synthase A chain (protein 6) - N. Fowleri	29	F	F	P	F	F	F	Y	L	F
BPCDNA6PT mRNA for BpcDNA6 protein - Brugia pahangi.	135	F	F	S	L	F	F	S	E	F
CCCOX3 mitochondrial cox3 gene - Chara corallina	78	F	F	V	A	F	F	W	A	F
CDPMTCO mitochondrial cytochrome oxidase subunit III (COIII) gene - Crocodylus porosus	5	F	F	L	G	F	F	W	A	F
COX3_HORSE cytochrome c oxidase polypeptide III - horse	93	F	F	S	G	F	F	W	A	F
DIHINTA13 integrase gene; virulence-associated protein genes - Dichelobacter nodosus	11	F	F	L	R	F	F	F	L	F
F40201 artifact-warning sequence (translated ALU class F) - human	457	F	F	F	F	F	F	A	L	F
GLR_RAT glucacon receptor precursor (GL-R). - rat	384	F	F	D	L	F	F	S	S	F
I30010 NADH dehydrogenase (ubiquinone) chain 5 - S. Tarentolae (SGC6)	16	F	F	F	M	F	F	F	M	F
LPF_ECOLI Phe leader peptide (Attenuator peptide). - E. Coli	6	F	F	F	A	F	F	F	T	F
MEU50217 cytochrome c oxidase subunit III, mitochondrial protein - blue mussel	81	F	F	F	S	F	F	W	A	F
MPOMTCG34 Marchantia polymorpha mitochondrion DNA - Marchantia polymorpha	114	F	F	L	F	F	F	V	T	F
MTU50218 cytochrome c oxidase subunit III, mitochondrial protein - north Pacific mussel	81	F	F	F	S	F	F	W	T	F
MUSPAFR Mouse gene for platelet activationg factor receptor - mouse	195	F	F	L	V	F	F	L	I	F
OPC1_RAT opioid binding protein/cell adhesion molecule prec.- rat	337	F	F	A	H	F	F	I	K	F
PRRMTCG6 complete mitochondrial genome DNA - Protopterus dolloi	93	F	F	F	G	F	F	W	A	F
PSBN_CHLRE photosystem II reaction centre N protein. - C. Reinhardtii	6	F	F	F	T	F	F	L	W	F
S43101 SED5 protein - fruit fly	300	F	F	F	L	F	F	V	V	F
TBU018491 EATRO 164 kinetoplast (CR4) mRNA - Trypanosoma brucei	93	F	F	D	L	F	F	M	L	F
TCU38184 kinetoplast minicircle ATPase subunit 6 mRNA - Trypanosoma cruzi	167	F	F	V	F	F	F	I	F	F
TECOX3 mitochondrial cox3 gene - Tmesipteris elongata	78	F	F	S	A	F	F	R	A	F
TPCOX3 mitochondrial cox3 gene - Tetraphis pellucida	78	F	F	L	A	F	F	W	G	F
VCU390682 Vibrio cholerae pathogenicity island - Vibrio cholerae	112	F	F	I	G	F	F	W	I	F

[a] 177 sequences were extracted from the OWL databank. The number before the first phenylalanine gives the position of the first residue in the protein sequence. Only non-redundant fragments of sequences are shown. Although different, sequences of several cytochrome *c* oxidase polypeptide III (COX3), probable or hypothetical proteins are removed for clarity.

creating a high local anionic concentration and acting as a selectivity barrier. The central part of the chloride channel in model I carries one positively charged residue, R334 (H6), and one negatively charged residue, E92 (H1). Model II has R334 and a different negatively charged residue; E1104 (H11). Several polar groups are also present; Y89, Y122, S313, S321, S341, and T1142 for model I; Y89, S308, T338, T341, T1115, S1118, and T1122 for model II.

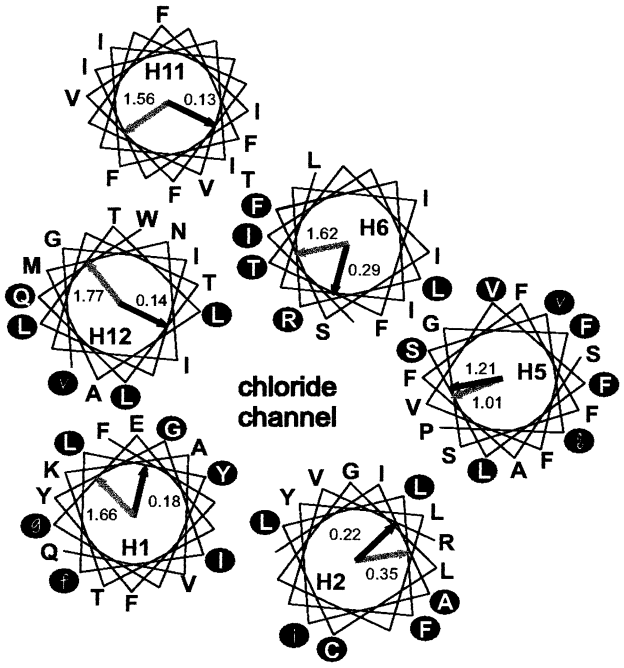

Figure 2. Top extracellular view of model I. Numbers denote the running number of the transmembrane helices in the primary sequence. Black arrows are conservation vectors; grey arrows are amphiphilicity vectors. Only the CFTR-conserved residues are written in capital letters. Those residues (indicated by both capital and lower-case letters) in black circles are conserved with respect to chloride channel sequences.

The topology of channels suggests that chloride ions bind at the entrance. Hydrated anions then traverse the channel via successive hydrogen bonds with polar residues. The positive residue of the pore centre, R334, could bind the anions and act as a local energetic barrier releasing chloride upon structural changes. This is in agreement with experiments by Sheppard et al., that report that mutation R334W modifies the conductance.[50] On the other hand, mutation K335E affects channel selectivity.[51] We propose that K335 could be implicated in a salt bridge with E873 (H7) in model I and with E92 (H1) in model II, since those acidic residues are located at the same level in the membrane. A mutation of K335 to a glutamic residue could thus alter helix matching.

The functioning of the CFTR is not completely elucidated and in model II, the pore is wide and could fulfill the cotransport of chloride with another substrate. It could also fulfill a function such as that described for K^+, Na^+, and Ca^{2+} channels, wherein a fragment of the lumenal domain inserts in the membrane to complete the pore. Such a mechanism has at present not been validated but should be explored in future since the first extracellular loop,[52] and the R and NBD domains[53] regulate chloride transport. More recently a short fragment, G404–N432 of NBD1, was described as traversing the membrane.[54,55] This fragment is hydrophilic and thus has little chance of traversing the membrane outside a surrounding protein.

There is no real agreement on the existence of a second pore in CFTRs. However, gene duplication is a reasonable hypothesis and duplicated sequences are generally described as symmetric structures. This occurred for sodium and calcium channels.[56] Since our models are made to trigger experimental assays, and since no experimental assays involve helices of

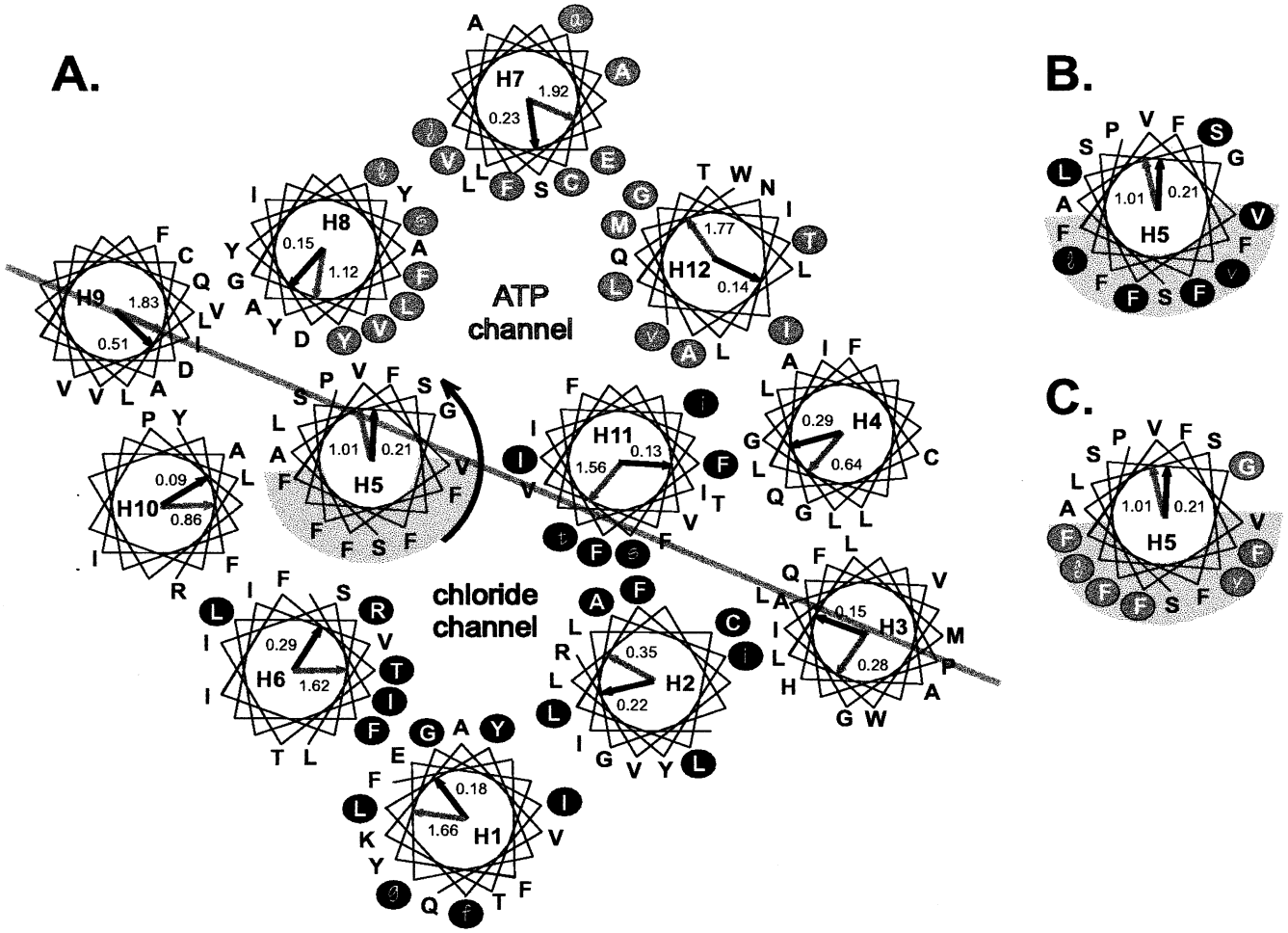

Figure 3. Top extracellular view of model II (A). Running numbers of the transmembrane helices (H1 to H12) are indicated inside each helix in the primary sequence. Black arrows are conservation vectors; grey arrows are amphiphilicity vectors. Only the CFTR-conserved residues are written in capital letters. Those residues (indicated by both capital and lower-case letters) in black and grey circles are conserved with respect to chloride channel and nucleoside transporter sequences, respectively. The symmetry axis is plotted. For helix 5, the grey area plotted around H5 is the conserved side with respect to the GABA receptor (B) and the ATP/ADP carrier (C).

MSD2 in the chloride channel, we suggest that the CFTR has two structurally symmetric, but functionally asymmetric, pores.

We suggest that the second pore involves MSD2, and that helices H5 (centrally located) H7, H8, and H12 are similar to membrane fragments of nucleoside transporters. We propose an ATP channel, made of helices H5, H7, H8 H11, and H12 (Figure 3). This ATP channel carries only one negatively charged residue (E873 in H7), which could interact with the ribose moiety of ATP (Color Plate 2C). The pore-lining residues are less hydrophilic than those in the chloride pore and several phenylalanines may be noted.

CONCLUSIONS

Our approach starts with the primary sequence of the human CFTR to give a 3-D topology of the membrane domain. The models are designed for experimental investigations of molecular mechanisms of ATP and chloride transport. Previously, only a 2-D model of the membrane domain of the CFTR had been proposed.[2] Overall, we agree with that model but some extracellular loops were too short to be structurally valid. Only one residue (K329) connected H5 and H6 and only two amino acids (Q1012 and P1013) connected H9 and H10. In our calculations the shortest loops are five residues long, to match the backward and upward (antiparallel) helices.

We postulate that all transmembrane fragments fold as α helices. Work investigating the accessibility of amino acids in H1 supports that proposal.[57] Moreover, it should be a good choice for H1, H6, H7, H9, H11, and H12 because the helices are amphipathic and the segregation of hydrophobic and hydrophilic amino acids on opposite sides effectively minimises energy. Moreover, the opposite distribution of variable and conserved positions in H9, and the similarity of helix 5 sequence with that of the transmembrane helices of the photosynthetic reaction centre or cytochrome *c* oxidase, further support our prediction. The last argument concerns the ratios of isoleucine, leucine, and valine (27% in H1 to 58% in H7), all described as helix stabilisers in membranes.[58]

Comparisons of sequences led to two conclusions. The high level of conservation of fragments in all CFTRs supports an overall functional importance. The similarity with membrane fragments of chloride and ATP transporters leads us to suggest (1) that several MSD1 fragments could be involved in the chloride channel, in agreement with the results of Sheppard et al.[49]; and (2) that if an ATP channel exists, it should be distinct from the chloride channel. There is experimental evidence of ATP transport by the CFTR,[11] but this evidence is disputed. Studies on Calu3 cells support the position that the CFTR does not transport ATP.[13] The overall biochemical data are controversial; our analysis suggests that the CFTR could be an ATP channel, but does not prove it. The only proof should be experimental. ATP transport has been involved in the regulation of ORCC channels. To our knowledge, no mutation causing a defect in the regulation of ORCC channels has yet been described. Nevertheless, one paper demonstrates that mutant D836X encodes a truncated protein (up to amino acid D836, i.e., including MSD1, NBD1, and the R domain but not MSD2 and NBD2). The mutant has a functional chloride channel[49] but it should not transport ATP since most helices of the ATP channel are absent (H7, H8, H11, and H12). Therefore, model II predicts that this mutant could be unable to regulate ORCC channels.

Mutations associated with mild forms of cystic fibrosis (R117H, R334W, and R347P) implicate three of our inner pore residues in the chloride conductance.[50] In other studies, basic amino acids of membrane helices were replaced by acidic residues (K95D, K335E, R347E, and R1030E). This modifies anion selectivity.[51] K95 and R1030 have no definite role in our models, K335 is implicated in helix matching, and R347 is at the channel entrance (Color Plate 2). McDonough et al.[59] suggest that diphenylamine-2-carboxylic acid interacts with S341 (H6) and T1134 (H12) residues: S341 is inside the chloride channel (Color Plate 2). Synthetic peptides corresponding to H2, H6, or the heterodimer H2/H6 induce channel conductance when inserted in bilayers.[48] In contrast, the four other membrane fragments of MSD1 fail to make active chloride channels. This suggests that not all helices are equally involved. Close examination of the models suggests that H6 should be a selectivity filter with K329 and R347, and should regulate chloride conductance with R334. Helix H2 should be a selectivity filter with R117 and R134 (Color Plate 2).

Our sequence analysis highlights the importance of helix 5. However, there is no experimental counterpart to this suggestion and no pathological support. This is not contradictory per se, because if the conservation of helix 5 is crucial for CFTR function, spontaneous mutations could be too deleterious. We would therefore encourage testing experimental mutations of the FFXXFFXXF motif at the levels of CFTR expression and function.

The predicted structures were discussed in the context of both chloride and ATP transport by the CFTR, but other kinds of substrates, such as neutral amino acids,[60] bicarbonate,[61] and adenosine 3′-phosphate 5′-phosphosulfate,[62] were ignored. Also ignored were data suggesting that hydrophilic fragments of NBD1 traverse the membrane.[54,55] The latter point will be examined in a forthcoming work on the modelling of "soluble" domains of the CFTR by homology.

ACKNOWLEDGMENTS

The Association Française de Lutte contre la Mucoviscidose (AFLM) supported X. Gallet, and Franco-Belge exchanges were funded by INSERM-CFB. R. Brasseur is Research Director at the National Fund for Scientific Research of Belgium (FNRS). This work was supported by the "Interuniversity Poles of Attraction Programme-Belgian State, Prime Minister's Office-Federal Office for Scientific, Technical and Cultural Affairs" contract P.4/03. The authors acknowledge Mr. Jean-Jacques Garnier and Hewlett-Packard France for computer donations.

REFERENCES

1 Welsh, M.J., Tsui, L.C., Boat, T.F., and Beaudet, A.L. Cystic fibrosis. In: *The Metabolic Basis of Inherited Disease* (Scriver, C.R., Beaudet, A.L., Sly, W.S., and Valle, D., eds.), 7th Ed. McGraw-Hill, New York, 1995, pp. 3799–3876

2 Riordan, J.R., Rommens, J.M., Kerem, B.S., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J.L., Drumm, M.L., Iannuzzi, M.C., Collins, F.S., and Tsui, L.C. Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Science* 1989, **245,** 1066–1073

3 Kerem, B.S., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M., and Tsui, L.C. Identification of the cystic fibrosis gene: Genetic analysis. *Science* 1989, **245,** 1073–1080

4 Rommens, J.M., Iannuzi, M.C., Kerem, B.S., Drumm, M.L., Melmer, G., Dean, M., Rozmahel, R., Cole, J.L., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M., Riordan, J.R., Tsui, L.C., and Collins, F.S. Identification of the cystic fibrosis gene: Chromosome walking and jumping. *Science* 1989, **245,** 1059–1065

5 Picciotto, M.R., Cohn, J.A., Bertuzzi, G., Greengard, P., and Nairn, A.C. Phosphorylation of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* 1992, **267,** 12742–12752

6 Berger, H.A., Travis, S.M., and Welsh, M.J. Regulation of the cystic fibrosis transmembrane conductance regulator Cl^- channel by specific protein kinases and protein phosphatases. *J. Biol. Chem.* 1993, **268,** 2037–2047

7 Pind, S., Riordan, J.R., and Williams, D.B. Participation of the endoplasmic reticulum chaperone calnexin (p88, IP90) in the biogenesis of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* 1994, **269,** 12784–12788

8 Bear, C.E., Li, C.H., Kartner, N., Bridges, R.J., Jensen, T.J., Ramjeesingh, M., and Riordan, J.R. Purification and functional reconstitution of the cystic fibrosis transmembrane conductance regulator (CFTR). *Cell* 1992, **68,** 809–818

9 Arispe, N., Rojas, E., Hartman, J., Sorscher, E.J., and Pollard, H.B. Intrinsic anion channel activity of the recombinant first nucleotide binding fold domain of the cystic fibrosis transmembrane regulator protein. *Proc. Natl. Acad. Sci. U.S.A.* 1992, **89,** 1539–1543

10 Reisin, I.L., Prat, A.G., Abraham, E.H., Amara, J.F., Gregory, R.J., Ausiello, D.A., and Cantiello, H.F. The cystic fibrosis transmembrane conductance regulator is a dual ATP and chloride channel. *J. Biol. Chem.* 1994, **269,** 20584–20591

11 Schwiebert, E.M., Egan, M.E., Hwang, T.H., Fulmer, S.B., Allen, S.S., Cutting, G.R., and Guggino, W.B. CFTR regulates outwardly rectifying chloride channels through an autocrine mechanism involving ATP. *Cell* 1995, **81,** 1063–1073
12 Devidas, S. and Guggino, W.B. The cystic fibrosis transmembrane conductance regulator and ATP. *Curr. Opin. Cell. Biol.* 1997, **9,** 547–552
13 Reddy, M.M., Quinton, P.M., Haws, C., Wine, J.J., Grygorczyk, R., Tabcharani, J.A., Hanrahan, J.W., Gunderson, K.L., and Kopito, R.R. Failure of the cystic fibrosis transmembrane conductance regulator to conduct ATP. *Science* 1996, **271,** 1876–1879
14 Li, C., Ramjeesingh, M., and Bear, C.E. Purified cystic fibrosis transmembrane conductance regulator (CFTR) does not function as an ATP channel. *J. Biol. Chem.* 1996, **271,** 11623–11626
15 Stutts, M.J., Canessa, C.M., Olsen, J.C., Hamrick, M., Cohn, J.A., Rossier, B.C., and Boucher, R.C. CFTR as a cAMP-dependent regulator of sodium channels. *Science* 1995, **269,** 847–850
16 Grubb, B.R., Vick, R.N., and Boucher, R.C. Hyperabsorption of Na^+ and raised Ca^{2+}-mediated Cl^- secretion in nasal epithelia of CF mice. *Am. J. Physiol.* 1994, **266,** C1478–C1483
17 Loussouarn, G., Demolombe, S., Mohammed-Panah, R., Escande, D., and Baró, I. Expression of CFTR controls cAMP-dependent activation of epithelial K^+ currents. *Am. J. Physiol.* 1996, **271,** C1565–C1573
18 Hyde, S.C., Emsley, P., Hartshorn, M.J., Mimmack, M.M., Gileadi, U., Pearce, S.R., Gallager, M.P., Gill, D.R., Hubbard, R.E., and Higgins, C.F. Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. *Nature (London)* 1990, **346,** 362–365
19 Mimura, C.S., Holbrook, S.R., and Ames, G.F. Structural model of the nucleotide-binding conserved component of periplasmic permeases. *Proc. Natl. Acad. Sci. U.S.A.* 1991, **88,** 84–88
20 Guy, H.R. Models of voltage- and transmitter-activated channels based on their amino acids sequences. In: *Monovalent Cations in Biological Systems* (Pasternak, C.A., ed.). CRC Press, Boca Raton, Florida, 1990, pp. 31–58
21 Baldwin, J.M. The probable arrangement of the helices in G protein-coupled receptors. *EMBO J.* 1993, **12,** 1693–1703
22 Stokes, D.L., Taylor, W.R., and Green, N.M. Structure, transmembrane topology and helix packing of P-type ion pumps. *FEBS Lett.* 1994, **346,** 32–38
23 Kyte, J. and Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 1982, **157,** 105–132
24 Eisenberg, D., Schwarz, E., Komaomy, M., and Wall, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* 1984, **179,** 125–142
25 Eisenberg, D. Three-dimensional structure of membrane and surface proteins. *Annu. Rev. Biochem.* 1984, **53,** 595–623
26 Brasseur, R. Calculation of the three-dimensional structure of *Saccharomyces cerevisiae* cytochrome *b* inserted in a lipid matrix. *J. Biol. Chem.* 1988, **263,** 12571–12575
27 Lins, L. and Brasseur, R. The hydrophobic effect in protein folding. *FASEB J.* 1995, **9,** 535–540
28 Brasseur, R. Simulating the folding of small proteins by use of the local minimum energy and the free solvation energy yields native-like structures. *J. Mol. Graphics* 1995, **13,** 312–322
29 Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., and Teller, A.H. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 1953, **21,** 1087–1092
30 Higgins, D.G. and Sharp, P.M. CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene* 1988, **73,** 237–244
31 Pearson, W.R. and Lipman, D.J. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A.* 1988, **85,** 2444–2448
32 Brasseur, R. Differentiation of lipid-associating helices by use of three-dimensional molecular hydrophobicity potential calculations. *J. Biol. Chem.* 1991, **266,** 16120–16127
33 Donnelly, D., Overington, J.P., Ruffle, S.V., Nugent, J.H., and Blundell, T.L. Modeling alpha-helical transmembrane domains: The calculation and use of substitution tables for lipid-facing residues. *Protein Sci.* 1993, **2,** 55–70
34 Fauchère, J.L. How hydrophobic is tryptophan? *Trends Biochem. Sci.* 1985, **10,** 268
35 Rahman, M. and Brasseur, R. WinMGM: A fast CPK molecular graphics program for analyzing molecular structure. *J. Mol. Graphics* 1994, **12,** 212–218
36 Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E., and Downing, K.H. Model for the structure of bacteriorhodopsin based on high-resolution electron cryomicroscopy. *J. Mol. Biol.* 1990, **213,** 899–929
37 Cowan, S.W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Paupitt, R.A., Jansonius, J.N., and Rosenbusch, J.P. Crystal structures explain functional properties of two *E. coli* porins. *Nature (London)* 1992, **358,** 727–733
38 Montal, M. Proteins folds in channel structure. *Curr. Opin. Struct. Biol.* 1996, **6,** 499–510
39 Unwin, N. Acetylcholine receptor channel imaged in the open state. *Nature (London)* 1995, **373,** 37–43
40 Ducarme, P., Rahman, M., Lins, L., and Brasseur R. The erythrocyte/brain glucose transporter (GLUT1) may adopt a two-channel transmembrane α/β structure. *J. Mol. Model.* 1996, **2,** 27–45
41 Chang, X.B., Hou, Y.X., Jensen, T.J., and Riordan, J.R. Mapping of cystic fibrosis transmembrane conductance regulator membrane topology by glycosylation site insertion. *J. Biol. Chem.* 1994, **269,** 18572–18575
42 De Loof, H., Rosseneu, M., Brasseur, R., and Ruysschaert, J.M. Hydrophobicity profiles for detection of receptor binding domains on apoliprotein E and the low density lipoprotein apoliprotein(B–E) receptor. *Proc. Natl. Acad. Sci. U.S.A.* 1986, **83,** 2295–2299
43 Gallet, X., Benhabiles, N., Lewin, M., Brasseur, R., and Thomas-Soumarmon, A. Prediction of the antigenic sites of the cystic fibrosis transmembrane conductance regulator protein by molecular modelling. *Protein Eng.* 1995, **8,** 829–834
44 Demolombe, S., Baró, I., Laurent, M., Hongre, A.S., Pavirani, A., and Escande, D. Abnormal subcellular

localization of mutated CFTR protein in a cystic fibrosis epithelial cell line. *Eur. J. Cell. Biol.* 1994, **65,** 214–219
45 Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R., and Yoshikawa, S. The whole structure of the 13-subunit oxidized cytochrome *c* oxidase at 2.8Å. *Science* 1996, **272,** 1136–1144
46 Chang, C.H., El-Kabbani, O., Tiede, D., Norris, J., and Schiffer, M. Structure of the membrane-bound protein photosynthetic reaction center from *Rhodobacter sphaeroides. Biochemistry* 1991, **30,** 5352–5360
47 Iwata, S., Ostermeier, C., Ludwig, B., and Michel, H. Structure at 2.8Å resolution of cytochrome *c* oxidase from *Paracoccus denitrificans. Nature (London)* 1995, **376,** 660–669
48 Oblatt-Montal, M., Reddy, G.L., Iwamoto, T., Tomich, J.M., and Montal, M. Identification of an ion channel-forming motif in the primary structure of CFTR, the cystic fibrosis chloride channel. *Proc. Nat. Acad. Sci. U.S.A.* 1994, **91,** 1495–1499
49 Sheppard, D.N., Ostedgaard, L.S., Rich, D.P., and Welsh, M.J. The amino-terminal portion of CFTR forms a regulated Cl^- channel. *Cell* 1994, **76,** 1091–1098
50 Sheppard, D.N., Rich, D.P., Ostedgaard, L.S., Gregory, R.J., Smith, A.E., and Welsh, M.J. Mutations in CFTR associated with mild-disease-form Cl^- channels with altered pore properties. *Nature (London)* 1993, **362,** 160–164
51 Anderson, M.P., Gregory, R.J., Thompson, S., Souza, D.W., Paul, S., Mulligan, R.C., Smith, A.E., and Welsh, M.J. Demonstration that CFTR is a chloride channel by alteration of its anion selectivity. *Science* 1991, **253,** 202–205
52 Price, M.P., Ishihara, H., Sheppard, D.N., and Welsh, M.J. Function of *Xenopus* cystic fibrosis transmembrane conductance regulator (CFTR) Cl^- channels and use of human–*Xenopus* chimeras to investigate the pore properties of CFTR. *J. Biol. Chem.* 1996, **271,** 25184–25191
53 Hwang, T.C., Nagel, G., Nairn, A.C., and Gadsby, D.C. Regulation of the gating of cystic fibrosis transmembrane conductance regulator Cl channels by phosphorylation and ATP hydrolysis. *Proc. Natl. Acad. Sci. U.S.A.* 1994, **91,** 4698–4702
54 Ko, Y.H., Delannoy, M., and Pedersen, P.L. Cystic fibrosis transmembrane conductance regulator: The first nucleotide binding fold targets the membrane with retention of its ATP binding function. *Biochemistry* 1997, **36,** 5053–5064
55 Gruis, D.B. and Price, E.M. The nucleotide binding folds of cystic fibrosis transmembrane conductance regulator are extracellularly accessible. *Biochemistry* 1997, **36,** 7739–7745
56 Guy, H.R. and Durell, S.R. Using sequence homology to analyse the structure and function of voltage-gated ion channel proteins. In: *Molecular Evolution of Physiological Processes.* The Rockefeller University Press, NY. 1994, pp. 197–212
57 Akabas, M.H., Kaufmann, C., Cook, T.A., and Archdeacon, P. Amino acid residues lining the chloride channel of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* 1994, **269,** 14865–14868
58 Li, S.C. and Deber, C.M. A mesure of the helical propensity for amino acids in membrane environments. *Nature Struct. Biol.* 1994, **1,** 368–373
59 McDonough, S., Davidson, N., Lester, H.A., and McCarty, N.A. Novel pore-lining residues in CFTR that govern permeation and open-channel block. *Neuron* 1994, **13,** 623–634
60 Rotoli, B.M., Bussolati, O., Sironi, M., Cabrini, G., and Gazzola, G.C. CFTR protein is involved in the efflux of neutral amino acids. *Biochem. Biophys. Res. Commun.* 1994, **204,** 653–658
61 Poulsen, J.H., Fischer, H., Illek, B., and Machen, T.E. Bicarbonate conductance and pH regulatory capability of cystic fibrosis transmembrane conductance regulator. *Proc. Natl. Acad. Sci. U.S.A.* 1994, **91,** 5340–5344
62 Pasyk, E.A. and Foskett, J.K. Cystic fibrosis transmembrane conductance regulator-associated ATP and adenosine 3′-phosphate 5′-phosphosulfate channels in endoplasmic reticulum and plasma membranes. *J. Biol. Chem.* 1997, **272,** 7746–7751