Supporting Information

Jordan *et al.* 10.1073/pnas.0806306105

SI Text

Functional Analysis. For single-channel recording, pipettes were pulled from borosilicate glass (Sutter Instrument) and were filled with chloride-containing pipette solution (in mM): 150 NMDG-Cl, 5 $MgCl_2$, 10 TES (pH 7.5). Channels were activated by excision into intracellular solution containing (in mM): 150 NMDG-Cl, 1.1 $MgCl_2$, 2 Tris-EGTA, 10 TES, 1 MgATP, and 50 Units/ml PKA (pH 7.5). CFTR currents were measured with an Axopatch 200B amplifier (Axon Instruments) and were recorded at 10 kHz to DAT tape. For subsequent analysis, records were played back and filtered with a four-pole Bessel filter (Warner Instruments) at a corner frequency of 100 Hz and acquired by using a Digidata 1322A interface (Axon) and computer at 500 Hz with pClamp 8.2. For display, single channel records were filtered digitally to 70 Hz.

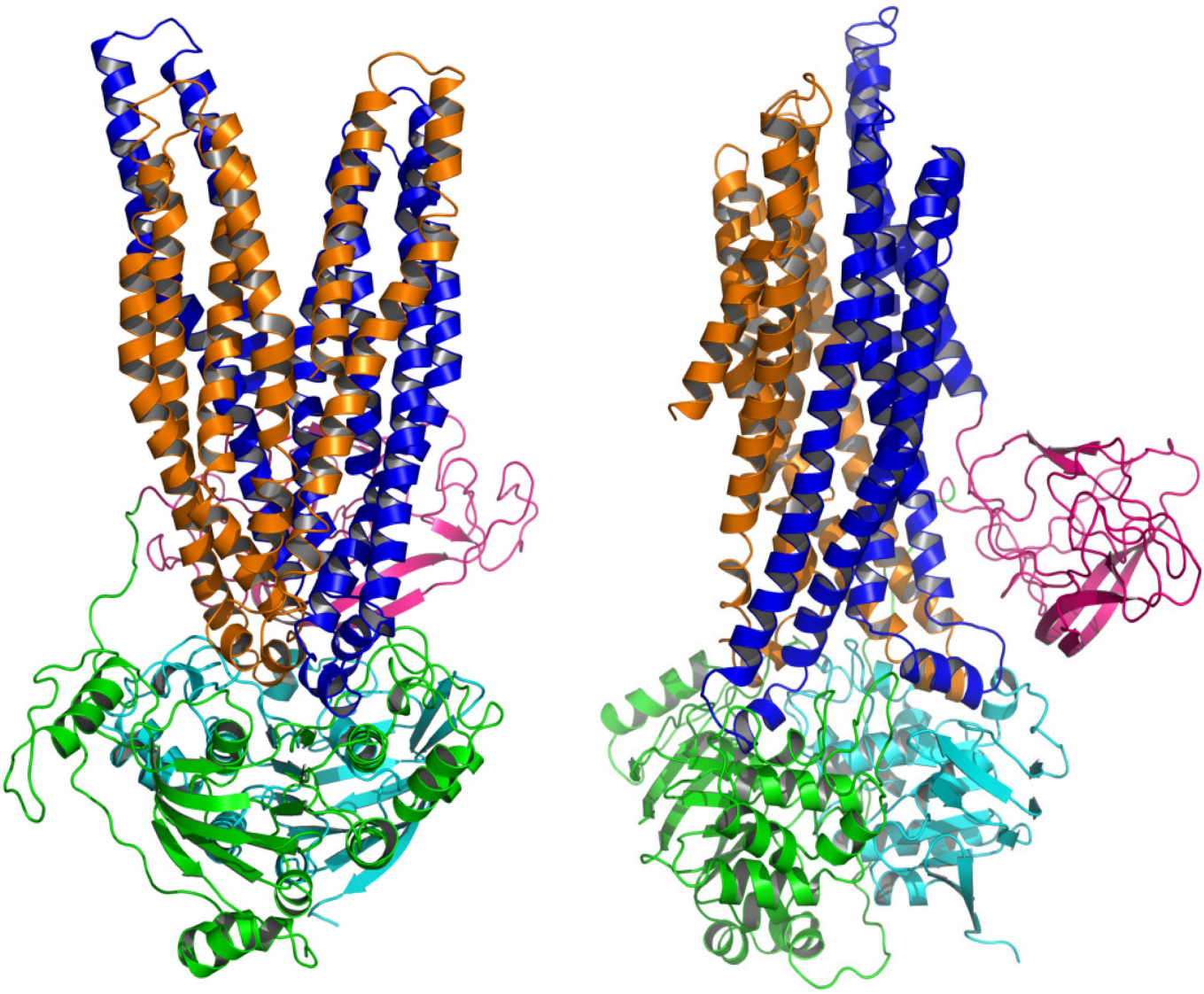


Fig. S1. Three-dimensional domain architecture of CFTR shown from the front (*Left*) and rotated 90° (*Right*). TMD1 is shown in orange, TMD2 is blue. The cytosolic domains are colored as: ABC1, green; ABC2, cyan; R-domain, pink.

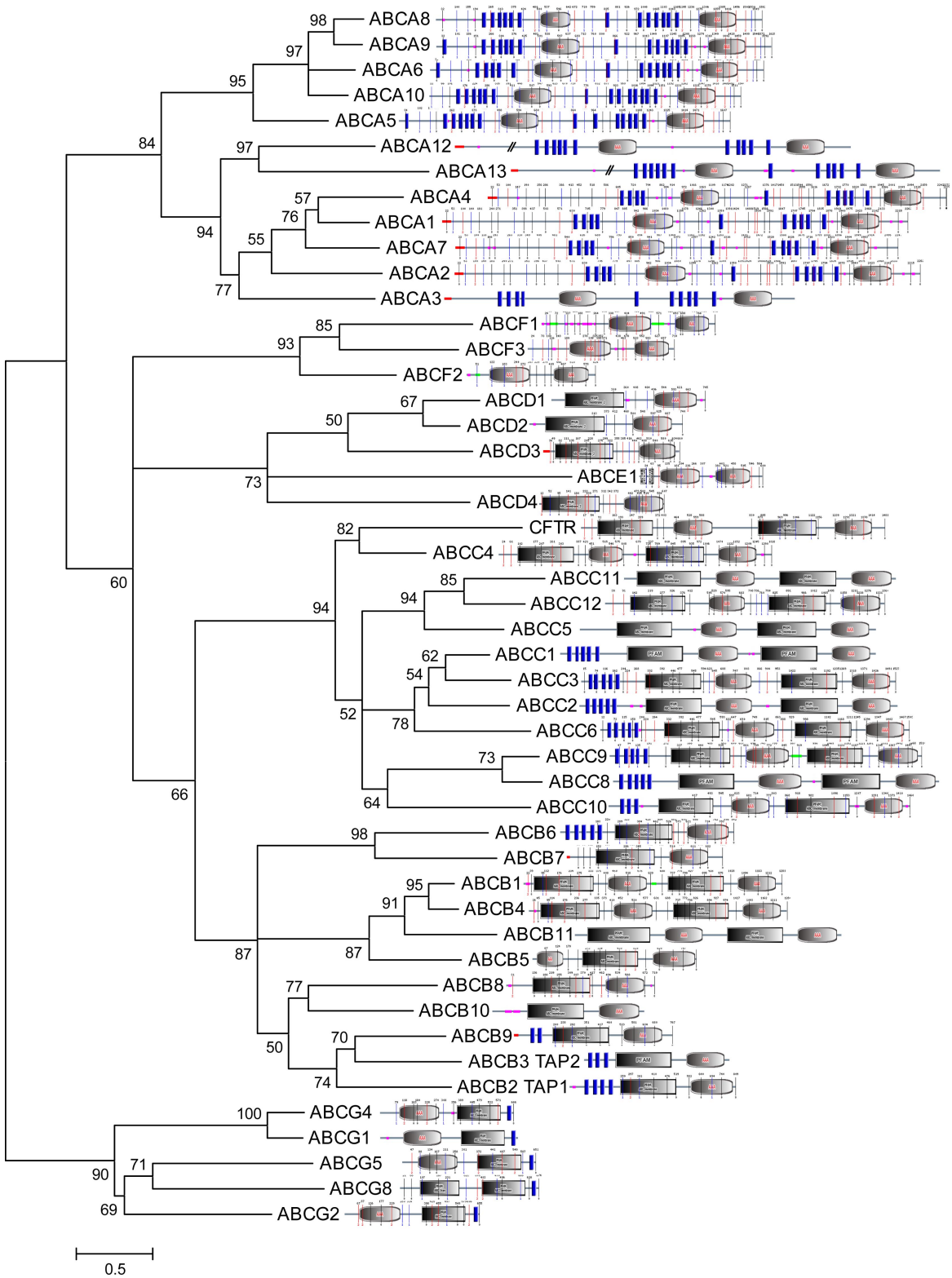


Fig. S2. Phylogenetic relationships and domain architectures of all human ABC proteins. Domain architectures were characterized with the Simple Modular Architecture Research Tool (see *Methods*). Locations of SMART and/or PFAM domains along the sequence are shown. Predicted signal peptides are shown in red and predicted transmembrane helices are shown in blue. A maximum likelihood phylogeny was reconstructed by using the program PUZZLE with Dayhoff distances computed from a protein sequence alignment. Bootstrap values showing support for internal nodes are shown above the branches.

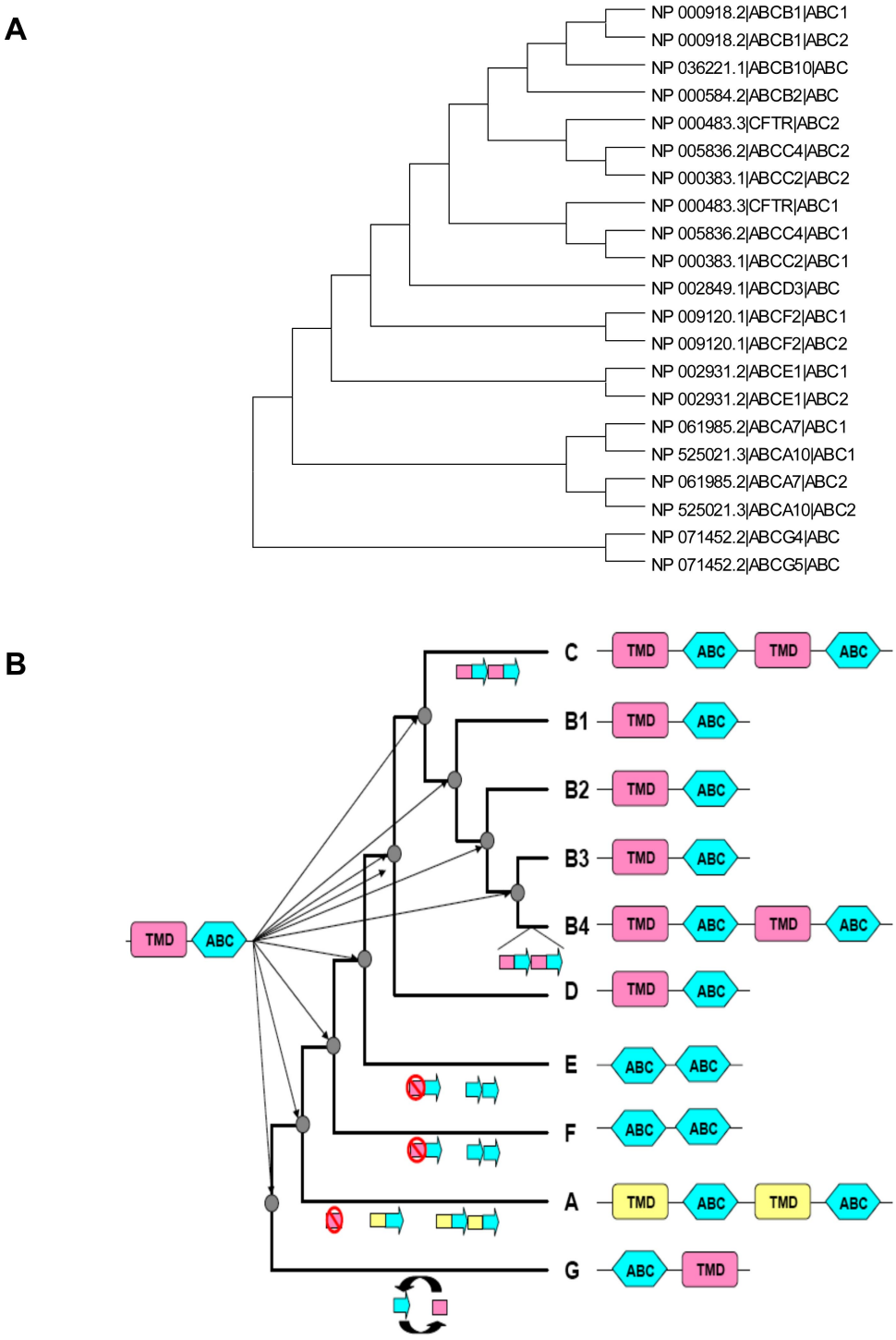


Fig. S3. Evolution of ABC family domain architecture. (*A*) ABC domain phylogenetic analysis, which was used to reconstruct the history of CFTR domain architecture evolution (*B*). In *A*, individual ABC domains (ABC1 and ABC2) were aligned and their evolutionary histories were reconstructed. An internal duplication leading to a domain architecture with two ABC domains would yield a phylogeny where both ABC (1 and 2) domains from a single protein are sister taxa (e.g., ABCB1). Ancient duplication followed by divergence would yield a phylogeny where the same ABC domains from different proteins are sister taxa (e.g., ABCA7 and ABCA10). In *B*, the scheme shown is based on the phylogenies shown in Fig. S2 and Fig. S3*A*. Domain architectures, that is, linear organizations of TMDs and ABC domains, are shown for human ABC A-G family proteins. Members of ABC subfamily B show two different domain architectures. The comparative method was used to infer the ancestral ABC family domain architecture and to reconstruct the specific domain architecture changes along evolutionary lineages leading to the extant subfamilies, or in the case of subfamily B the individual members. The ancestral state and domain architecture changes are inferred as those that entail the fewest number of changes along the tree, that is, the most parsimonious scenario. Specific domain architecture changes are shown below the branches where they are inferred to have occurred.

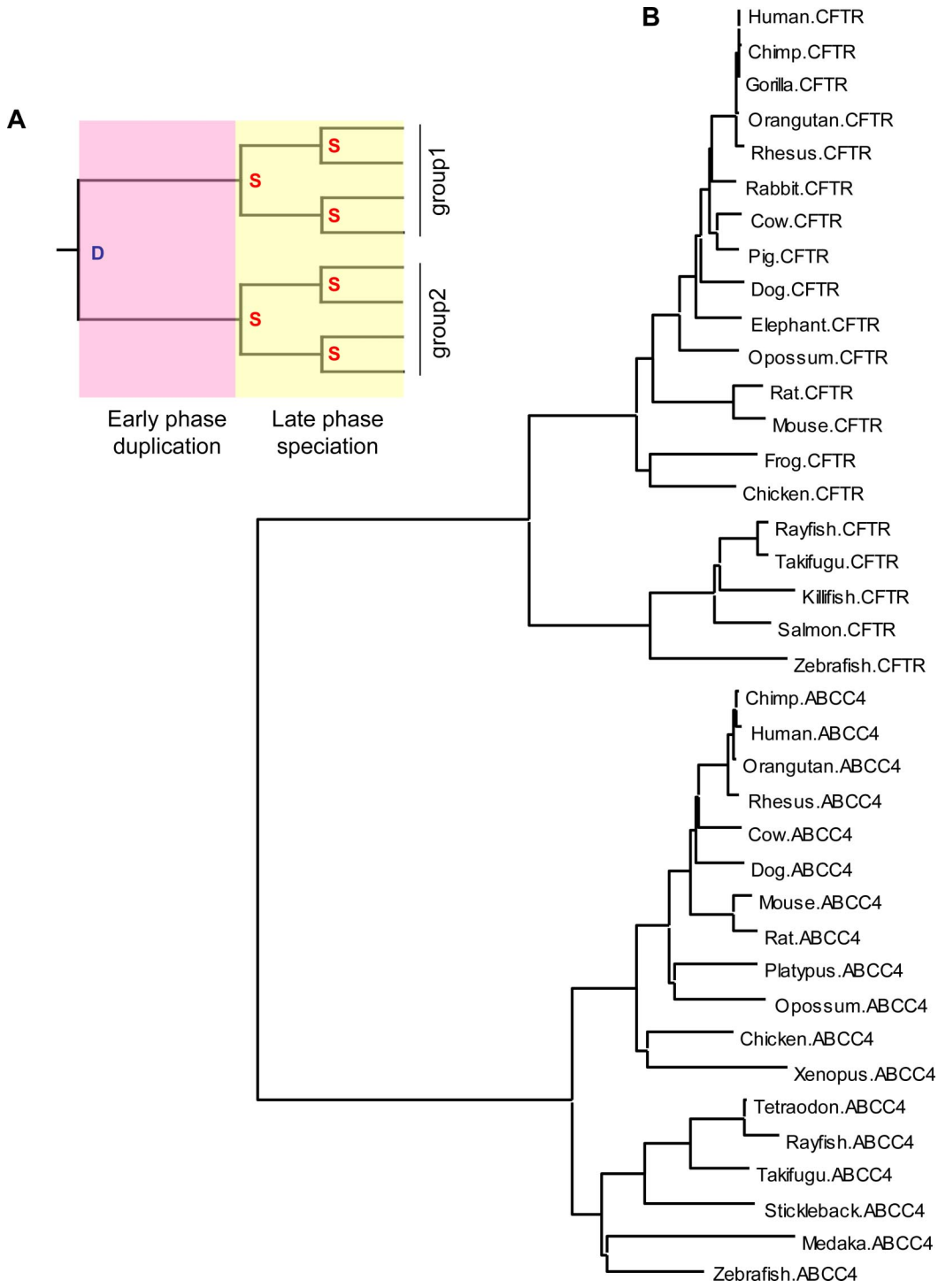


Fig. S4. Rationale behind the type II divergence analysis. (*A*) An ancient gene duplication yields two paralogous groups of related proteins, such as CFTR and ABCC4. Changes in functional constraints between paralogous groups occur in two phases: an early phase (pink) just after duplication (D) and a later phase (yellow) during speciation (S) within each group. Early phase changes are characterized by relaxation of functional constraint along one or both duplicate lineages. Sites that are responsible for group-specific functions diverge rapidly during this phase. Once paralog-specific functions that distinguish groups are encoded, late phase changes occur during speciation within groups. Sites that have group-specific critical functions do not change within groups during the late phase. Sites that have group-specific critical functions and are responsible for functional differences between groups, the so-called type II sites, also do not change. This is why type II divergence sites are maximally divergent between groups and absolutely conserved within groups. Sites that are functionally critical to both groups remain absolutely conserved after duplication and speciation, and such sites do not reveal anything about functional differences between groups. (*B*) Phylogeny of 20 CFTR and 18 ABCC4 proteins used in the functional divergence analysis.

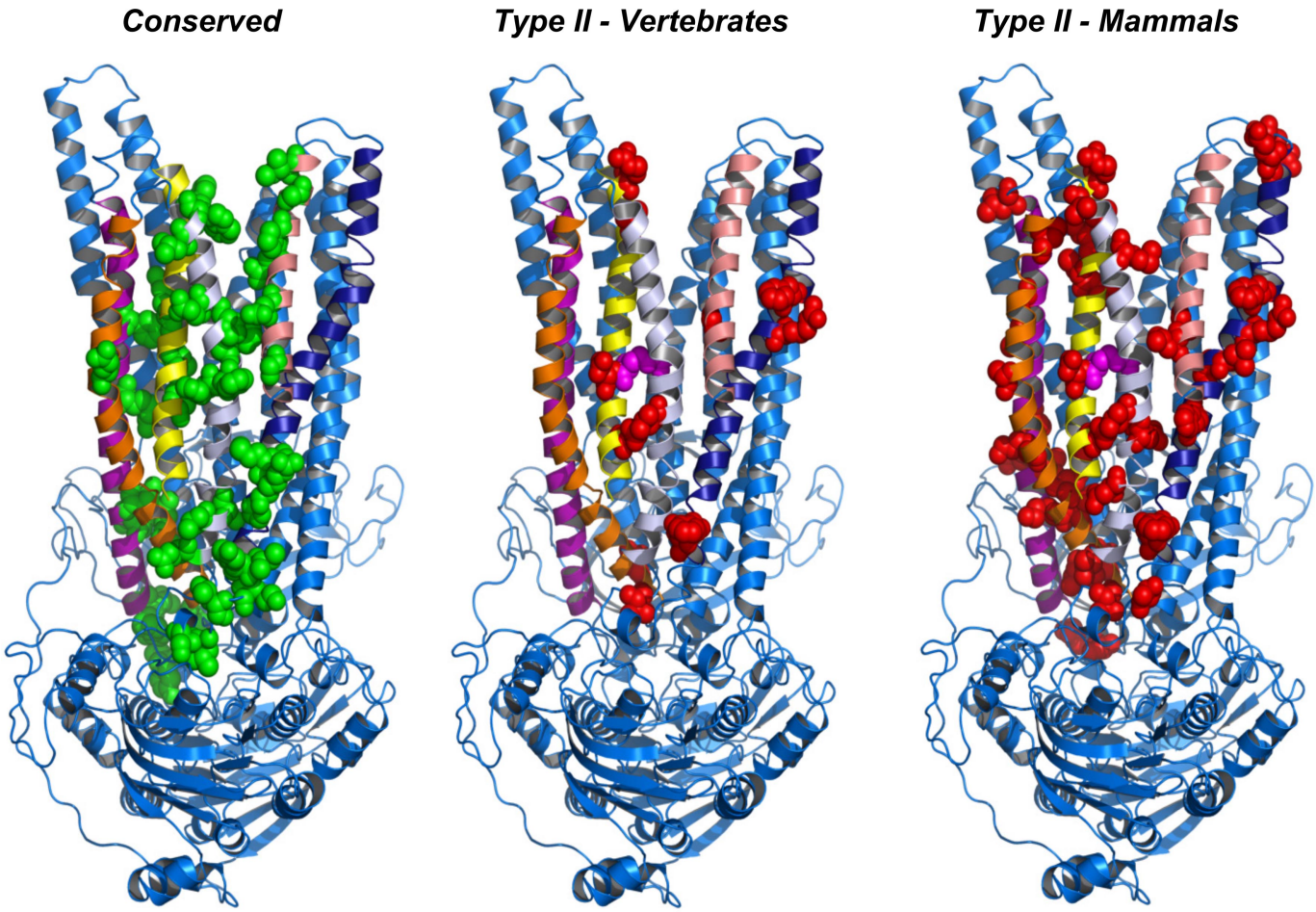
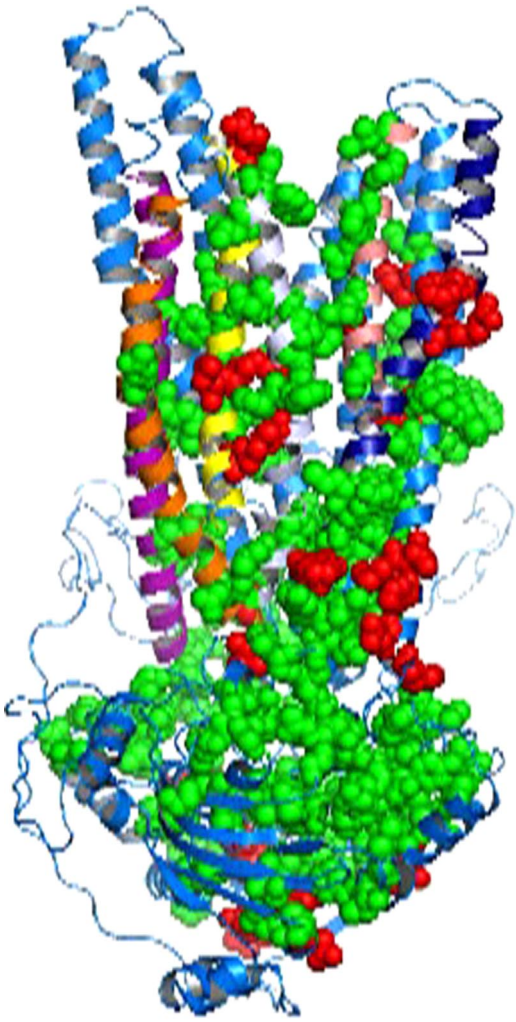
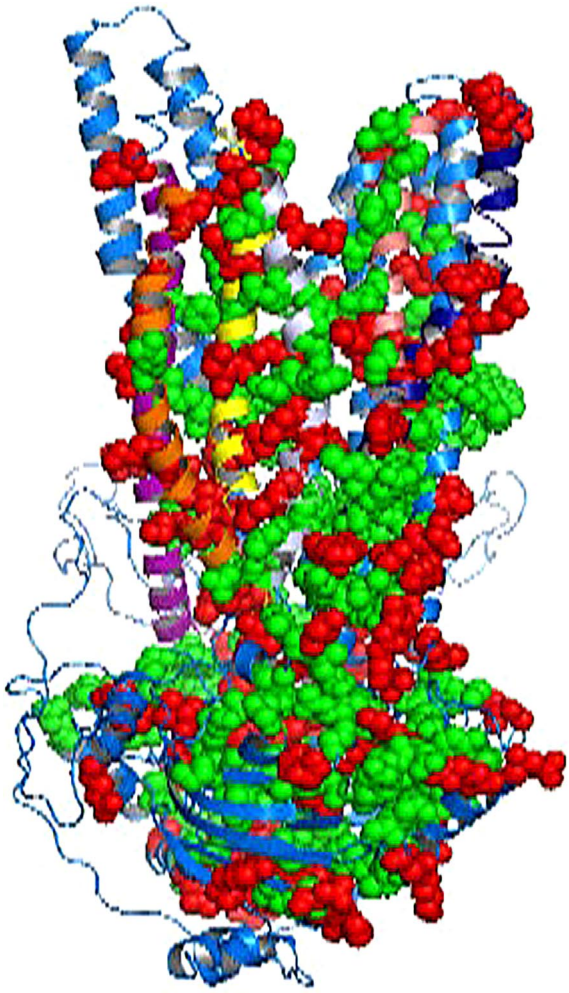


Fig. S5. Structural environment of conserved (*Left*) and type II divergent (*Center* and *Right*) CFTR residues in TMD1 shown within the context of the entire CFTR protein. Helices are colored as follows: TM1, salmon; TM2, dark blue; TM3, silver; TM4, orange; TM5, purple; and TM6, yellow. TMD2, both ABC domains, and the R-domain are shown in blue. This clearly demonstrates the location of both conserved and type II divergent sites located within the interconnecting domains that are involved in transmitting structural rearrangements from the ABC domains to the TMDs.



Movie S1. Conserved and divergent sites in all four domains: Vertebrate proteins. Conserved and type II divergent residues are shown as green and red spheres, respectively, for the entire CFTR protein. CFTR is rotated about the *y*-axis to illustrate how widely dispersed both conserved and type II divergent sites are within the protein. The helices of TMD1 are colored as follows: TM1, salmon; TM2, dark blue; TM3, silver; TM4, orange; TM5, purple; and TM6, yellow. The remainder of CFTR is colored blue.

Movie S1 (MPG)



Movie S2. Conserved and divergent sites in all four domains: Mammalian proteins. Conserved and type II divergent residues are shown as green and red spheres, respectively, for the entire CFTR protein. CFTR is rotated about the *y*-axis to illustrate how widely dispersed both conserved and type II divergent sites are within the protein. The helices of TMD1 are colored as follows: TM1, salmon; TM2, dark blue; TM3, silver; TM4, orange; TM5, purple; and TM6, yellow. The remainder of CFTR is colored blue.

Movie S2 (MPG)

Table S1. Sites of known functional effects of mutations in TMD1 of human CFTR, used as input for analysis shown in Figure 2

Amino acid residue	TM helix	Role in CFTR or association with CF disease and categories of effects of mutations* and Refs.
M82	1	A 1
Y84	1	A 1
G85	1	A 1
F87	1	A 1
L88	1	A 1
Y89	1	A 1
L90	1	A 1
G91	1	A 1, C 2
E92	1	A 1
K95	1	C 3, B 4
A96	1	A 1
Q98	1	A 1
P99	1	A 1, C 5
L101	1	A 1
I105	1	A 1
A107	1	A 1
S108	1	A 1
Y109	Extracellular Loop 1	A 1
D110	Extracellular Loop 1	A 1
P111	Extracellular Loop 1	A 1
N113	Extracellular Loop 1	A 1
E116	2	A 1
R117	2	A 1, C 6, S 6
I119	2	A 1
A120	2	A 1
Y122	2	A 1
I125	2	A 1
G126	2	A 1
L137	2	A 1
H139	2	A 1, P 7
P140	2	A 1
A141	2	A 1
L145	2	A 1
H146	2	A 1
I148	2	A 1
G149	2	A 1, P 7
Q151	2	A 1
M152	2	A 1
A155	2	A 1
S158	2	A 1
L159	2	A 1
Y161	Intracellular Loop 1	A 1
K162	Intracellular Loop 1	A 1
L165	Intracellular Loop 1	A 1
R170	Intracellular Loop 1	A 1
I175	Intracellular Loop 1	A 1
I177	Intracellular Loop 1	A 1
G178	Intracellular Loop 1	A 1
Q179	3	A 1
L183	3	A 1
N186	3	A 1
N187	3	A 1
N189	3	A 1
D192	3	A 1, P 7
E193	3	A 1, S 7
G194	3	A 1
A196	3	A 1
H199	3	A 1
F200	3	A 1
V201	3	A 1
W202	3	A 1
I203	3	A 1
P205	3	A 1, C 5

Amino acid residue	TM helix	Role in CFTR or association with CF disease and categories of effects of mutations* and Refs.
L206	3	A 1
A209	3	A 1
L210	3	A 1
G213	3	A 1
E217	3	A 1
Q220	Extracellular Loop 2	A 1
C225	4	A 1
L227	4	A 1
V232	4	A 1
Q237	4	A 1
A238	4	A 1
G239	4	A 1
G241	4	A 1
M243	4	A 1
M244	4	A 1
R248	4	A1
R258	4	A 1, P 7
M265	4	A 1
W277	Intracellular Loop 2	A 1
E279	5	A 1
M281	5	A 1
I285	5	A 1
N287	5	A 1
E292	5	A 1
L293	5	A 1
R297	5	A 1
A299	5	A 1
Y301	5	A 1
R303	5	C 8
F305	5	A 1
S307	5	A 1
A309	5	A 1
F311	5	A 1
S313	5	A 1
G314	5	A 1, C 2
F316	5	A 1
V317	5	G 9
L320	5	A 1
S321	5	A 1
V322	5	A 1
P324	Extracellular Loop 3	C 5
L327	Extracellular Loop 3	A 1
I331	Extracellular Loop 3	S 10
L333	6	S 10
R334	6	A 1, C 6,11,12,13, B 13,14, S 13,15
K335	6	C 2,3,11,16,18,19, B 14,17, S 10
I336	6	A 1
F337	6	C 11,20,22, B 14,21
T338	6	A 1, C 11,19,20,22,23,24, B 17,21, S 25
T339	6	C 23,24
I340	6	A 1
S341	6	A 1, C 11,19, B 14,17
I344	6	C 22, B 14
L346	6	A 1
R347	6	A 1, C 6,16,18,27,28,30,31, B 14,16,26,32, S 27, N 29
M348	6	A 1, B 14
A349	6	A 1
V350	6	B 14
R352	6	A 1, C 8,11,30,33,34, B 14,34, S 34, N 34, G 34
Q353	6	A 1
P355	6	A 1
W356	6	A 1
Q359	6	A 1
T360	6	A 1

Amino acid residue	TM helix	Role in CFTR or association with CF disease and categories of effects of mutations* and Refs.
W361	6	A 1

*Categories include: A, Disease-associated; B, Affects block of the pore by drugs or large anions; C, Affects conductance, selectivity, or current rectification; G, Affects macroscopic gating; N, Has nonspecific effects such as inducing appearance of sensitivity to sulfhydryl-modifying reagents; P, Affects processing of new CFTR polypeptide; S, Affects stability of the open channel state or open probability.

1. Tsui LC, Zielenski J, O'Brien A (2008) Cystic Fibrosis Mutation Database, http://www.genet.sickkids.on.ca/cftr/Home.html.
2. Mansoura MK, *et al.* (1998) *Biophys J* 74:1320–1332.
3. Anderson MP, *et al.* (1991) *Science* 253:202–205.
4. Linsdell P. (2005) *J Biol Chem* 280:8945–8950.
5. Sheppard DN, Travis SM, Ishihara H, Welsh MJ (1996) *J Biol Chem* 271:14995–15001.
6. Sheppard DN, *et al.* (1993) *Nature* 362:160–164.
7. Seibert FS, *et al.* (1997) *Biochemistry* 36:11966–11974.
8. St.Aubin CN, Linsdell P (2006) *J Gen Physiol* 128:535–545.
9. Zhang ZR, Zeltwanger S, Smith SS, Dawson DC, McCarty NA (2001) *Am J Physiol* 282:L135–L145.
10. Beck EJ, Yang Y, Yaemsiri S, Raghuram V (2008) *J Biol Chem* 283:4957–4966.
11. Fatehi M, Linsdell P (2008) *J Biol Chem* 283:6102–6109.
12. Gong X, Linsdell P (2003) *J Physiol* 549:387–397.
13. Smith SS, *et al.* (2001) *J Gen Physiol* 118:407–431.
14. Cui G, Zhang ZR, Song B, McCarty NA (*unpublished observations*)
15. Zhang ZR, Song B, McCarty NA (2005) *J Biol Chem* 280:41997–42003.
16. Linsdell P, Hanrahan JW (1998) *J Gen Physiol* 111:601–614.
17. McDonough S, Davidson N, Lester HA, McCarty NA (1994) *Neuron* 13:623–634.
18. Tabcharani JA, *et al.* (1993) *Nature* 366:79–82.
19. McCarty NA, Zhang ZR (2001) *Am J Physiol* 281:L852–L867.
20. Linsdell P (2001) *J Physiol* 531:51–66.
21. Gupta J, Linsdell P (2002) *Pflügers Arch* 443:739–747.
22. Linsdell P, Evagelidis A, Hanrahan JW (2000) *Biophys J* 78:2973–2982.
23. Linsdell P, *et al.* (1997) *J Gen Physiol* 110:355–364.
24. Linsdell P, Zheng SX, Hanrahan JW (1998) *J Physiol* 512:1–16.
25. Liu X, *et al.* (2004) *Biophys J* 87:3826–3841.
26. Wang W, Li G, Clancy JP, Kirk KL (2005) *J Biol Chem* 280:23622–23630.
27. Cotten JF, Welsh MJ (1999) *J Biol Chem* 274:5429–5435.
28. Kogan I, *et al.* (2003) *EMBO J* 22:1981–1989.
29. Kogan I, Ramjeesingh M, Huan LJ, Wang Y, Bear CE (2001) *J Biol Chem* 276:11575–11581.
30. Cheung M, Akabas MH (1997) *J Gen Physiol* 109:289–299.
31. Tabcharani JA, Linsdell P, Hanrahan JW (1997) *J Gen Physiol* 110:341–351.
32. Linsdell P, Hanrahan JW (1996) *J Physiol (Cambr)* 496:687–693.
33. Guinamard R, Akabas MH (1999) *Biochemistry* 38:5528–5537.
34. Cui G, Zhang ZR, O'Brien ARW, Song B, McCarty NA (2008) *J Membr Biol* 222:91–106.

Table S2. Type-II divergence sites for TMD1 across all vertebrates

Position	CFTR	ABCC4	Property Change
91	G	E	Hydrophilic vs. Acidic
131	F	T	Hydrophobic vs. Hydrophilic
134	R	L	Basic vs. Hydrophobic
157	F	C	Hydrophobic vs. Hydrophilic
158	S	H	Polar uncharged vs. Basic
175	I	T	Hydrophobic vs. Hydrophilic
195	L	T	Hydrophobic vs. Hydrophilic
262	T	M	Hydrophilic vs. Hydrophobic
334	R	S	Basic vs. Hydrophilic
352	R	L	Basic vs. Hydrophobic
353	Q	F	Hydrophilic vs. Hydrophobic
359	Q	E	Hydrophilic vs. Acidic

Table S3. Divergence site bins for TMD1

Bin	Between groups	Within groups	Number of sites	Average Posterior Probability
Conserved	No change or conserved change	0–3 changes	108	0.11
Intermediate Class 1	Conserved change	0–3 changes	100	2.53
Intermediate Class 2	Conserved or Radical change	4–9 AA changes	44	7.16
Intermediate Class 3	Radical change	1–3 AA changes	17	12.54
Type-II	Radical change	No change	12	29.74

Table S4. Accession numbers and names for human ABC sequences analyzed in this study

Refseq ID	Protein name
NP_000483.3	CFTR
NP_005493.2	ABCA1
NP_001597.2	ABCA2
NP_001080.2	ABCA3
NP_000341.2	ABCA4
NP_061142.2	ABCA5
NP_525023.2	ABCA6
NP_061985.2	ABCA7
NP_009099.1	ABCA8
NP_525022.2	ABCA9
NP_525021.3	ABCA10
NP_775099.2	ABCA12
NP_689914.2	ABCA13
NP_000918.2	ABCB1
NP_000584.2	TAP1
NP_000535.3	TAP2
NP_000434.1	ABCB4
NP_848654.3	ABCB5
NP_005680.1	ABCB6
NP_004290.2	ABCB7
NP_009119.2	ABCB8
NP_062571.1	ABCB9
NP_036221.1	ABCB10
NP_003733.2	ABCB11
NP_004987.2	ABCC1
NP_000383.1	ABCC2
NP_003777.2	ABCC3
NP_005836.2	ABCC4
NP_005679.2	ABCC5
NP_001162.3	ABCC6
NP_000343.2	ABCC8
NP_064694.2	ABCC9
NP_258261.2	ABCC10
NP_149163.2	ABCC11
NP_150229.2	ABCC12
NP_000024.2	ABCD1
NP_005155.1	ABCD2
NP_002849.1	ABCD3
NP_005041.1	ABCD4
NP_002931.2	ABCE1
NP_001020262.1	ABCF1
NP_009120.1	ABCF2
NP_060828.2	ABCF3
NP_997510.1	ABCG1
NP_004818.2	ABCG2
NP_071452.2	ABCG4
NP_071881.1	ABCG5
NP_071882.1	ABCG8