CMLS, Cell. Mol. Life Sci. 61 (2004) 230–242
1420-682X/04/020230-13
DOI 10.1007/s00018-003-3386-z

CMLS Cellular and Molecular Life Sciences

Research Article

Nucleotide-binding domains of human cystic fibrosis transmembrane conductance regulator: detailed sequence analysis and three-dimensional modeling of the heterodimer

I. Callebaut[a,*], R. Eudes[a], J.-P. Mornon[a] and P. Lehn[b]

[a] Systèmes moléculaires & Biologie structurale, LMCP, CNRS UMR7590, Universités Paris 6 & Paris 7, case 115, 4 place Jussieu, 75252 Paris Cedex 05 (France), e-mail: Isabelle.Callebaut@lmcp.jussieu.fr
[b] INSERM U458, Hôpital Robert Debré, 49 blvd Sérurier, 75019 Paris (France)

Received 16 October 2003; received after revision 16 November 2003; accepted 21 November 2003

Abstract. The cystic fibrosis transmembrane conductance regulator (CFTR) protein is encoded by the gene that is defective in cystic fibrosis, the most common lethal inherited disease among the Caucasian population. CFTR belongs to the ABC transporter superfamily, whose members form macromolecular architectures composed of two membrane-spanning domains and two nucleotide-binding domains (NBDs). The experimental structures of NBDs from several ABC transporters have recently been solved, opening new avenues for understanding the structure/function relationships and the consequences of some disease-causing mutations of CFTR. Based on a detailed sequence/structure analysis, we propose here a three-dimensional model of the human CFTR NBD heterodimer. This model, which is in agreement with recent experimental data, highlights the specific features of the CFTR asymmetric active sites located at the interface between the two NBDs. Moreover, additional CFTR-specific features can be identified at the subunit interface, which may play critical roles in active site interdependence and are uncommon in other NBD dimers.

Key words. ABC transporter; cystic fibrosis; CFTR; MRP1; SUR1; disease-causing mutation; molecular modeling; hydrophobic cluster analysis.

Introduction

Cystic fibrosis (CF), the most common severe inherited disease among the Caucasian population, is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene [1, 2]. This gene encodes a protein of 1480 amino acids, which acts as an epithelial selective chloride ion channel [3, 4]. Phosphorylation of the CFTR protein regulates its channel activity, whereas its ATPase activity (ATP binding and hydrolysis) is presumed to provide the energy required for channel gating [4]. Evidence now exists that CFTR also has other functions, including the transport of other ions, such as HCO_3^- and regulation of the activity of other channels such as the epithelial sodium channel [5–7]. CFTR (ABCC7) belongs to subfamily C of the ATP-binding cassette (ABC) transporter family [8, 9]. Other ABC proteins of subfamily C are involved in human disease processes, and include multidrug resistance-associated protein 1 (MRP1/ABCC1) and the sulfonyl urea receptor 1 (SUR1/ABCC8), which are involved in multidrug resistance and persistent hyperinsulinemic hypoglycemia, respectively. A further disease-associated member of the ABC protein family (subfamily B) is the multidrug resistance (MDR) P-glycoprotein. Despite mediating the translocation of a wide variety of solutes across membranes, ABC trans-

* Corresponding author.

porters share the same basic architecture, with two nucleotide-binding domains (NBDs), which are relatively well conserved among transporters and two membrane-spanning domains (MSDs) characterized by greater sequence and structural diversity. In prokaryotes, the four domains often comprise separate polypeptide chains, whereas in some prokaryotic and most eukaryotic transporters, the different domains are contained within a single polypeptide chain. In CFTR, the four domains are provided by a single chain, with two structurally homologous halves each containing an MSD followed by an NBD. The two halves are linked by a regulatory domain (R domain) that contains phosphorylation sites [3].

The presence of ABC transporters in all kingdoms of life and their involvement in human diseases highlights the need to understand their structures and associated functions. High-resolution structures of NBDs from several ABC transporters have been solved over the last 5 years [see refs 10–12 for reviews]. The first structures reported were those of the nucleotide-binding component of the *Escherichia coli* high-affinity ribose transport system (RbsA) [13] and *Salmonella typhimurium* histidine permease (HisP) [14], which revealed the canonical fold of the ABC-type ATPase domains. This fold consists of three subdomains: an F1-ATPase-like catalytic core domain, which contains most of the residues that contact the phosphate moiety of ATP [Walker A (or P-loop), Q-loop, Walker B, and switch motifs] and two ABC-specific subdomains. The first specific subdomain corresponds to a small peripheral β sheet (ABCβ subdomain), which is involved in binding of the ribose and base of the nucleotide, whereas the second, which is mainly helical (ABCα subdomain) contains the ABC signature motif or C-loop. Several other structures (with ATP-bound, ADP-bound or empty active sites) have been solved more recently, showing in particular that γ-phosphate release during ATP hydrolysis leads to the withdrawal of the glutamine of the Q-loop from the active site and to a coupled large movement of the ABCα subdomain relative to the NBD core [15, 16]. The cooperativity in ATP hydrolysis observed in biochemical and genetic studies of ABC transporters has clearly indicated that NBDs work in pairs, either as homodimers or heterodimers [9, 17]. In the solved structures of ABC transporters, distinct dimeric arrangements were observed. The identification of biologically relevant dimers remained controversial until recently, complicated by the fact that isolated ABC NBDs appear not to form dimers in solution, without additional subunits [reviewed in refs 11, 12]. The 'head-to-tail' orientation of the Rad50 [18] and MutS dimers [19, 20], two proteins related to the ABC transporter ATPase domains, appears to be the most physiologically consistent with the observed cooperativity in ATP hydrolysis and the presence of a highly conserved ABC signature motif, with key residues of each monomer participating in ATP binding and sensing. According to such a 'head-to-tail' orientation, NBD subunits are aligned side by side in a way that juxtaposes the ABC signature motif of one subunit to the P-loop of the other, thus placing the nucleotide-binding site at the interface between the two opposing NBDs. A 'head-to-tail' orientation was also recently observed within the crystal structure of an isolated ABC ATPase domain of an archaebacterial ABC transporter (MJ0796) [21], which forms a stable but hydrolytically inactive dimer due to the mutation of the Walker B catalytic base, as well as within the crystal structure of the complete *E. coli* vitamin B12 importer BtuCD [22]. This last dimeric structure is all the more consistent as it was constrained by the presence of dimers of MSD. In addition to BtuCD, the structure of another full-length ABC transporter has also recently been solved, that of the lipid flippase MsbA from *E. coli* [23], confirming the dimeric arrangement inferred from several previous studies. The MsbA and BtuD MSDs differ, however, as they are composed of six and ten transmembrane α helices, respectively. In both structures, loops or regular secondary structures connecting the MSD transmembrane helices make contacts with the NBDs. The high resolution achieved for the BtuCD structure allowed the identification of contacts between the MSD and NBD, indicating the crucial role of the Q-loop and the first two helices of the ABC α-specific subdomain. The same NBD region appears to be involved in the MSD-NBD contacts in MsbA [24], although no high-resolution structure is available for the NBD of this last structure and the MSD regions involved in contact differ.

Despite this overwhelming amount of data, no detailed model of CFTR NBDs has yet been reported using as templates the high-resolution structures of ABC transporters described above. This clearly hampers a comprehensive analysis of the active sites and mutations of CFTR. A few predictions have, however, been made for the structural consequences of a limited number of disease-associated missense mutations [see e.g., ref 25], but none has addressed these issues in details. In addition, although NBD sequences are generally well conserved as a whole, some regions of the CFTR NBDs are difficult to align accurately with other ABC transporters, in particular within structurally variable regions. This results in ambiguity in sequence alignments and consequently hampers the making of relevant predictions and/or models. We therefore aimed at building an accurate three-dimensional model of the CFTR NBD dimer structure by taking into account the evolutionary information from ABC transporter NBDs of known three-dimensional structure. In addition to the HisP [14], BtuCD [22] and MJ0796 [15, 21] structures indicated above, we also considered all other available experimental structures, solved at atomic resolution: MJ1267 [16], an ABC transporter of *Thermotoga maritima* [R. Zhang, A. Joachimiak, A. Edwards, A. Savchenko and S. Beasley,

unpublished results], TAP1 [26] and MalK [27]. For this purpose, we took advantage of the high sensitivity of the hydrophobic cluster analysis (HCA) method [28, 29; and 30, 31 for recent typical applications] to decipher structural and functional information within distantly related sequences. First, an in-depth analysis of the aforementioned three-dimensional NBD structures combined with associated structure-based alignments allowed the identification of crucial conserved residues/motifs, which served as anchor points for the alignment of the CFTR NBD sequences. The MJ0796 NBD dimer [21] was then used for modeling the CFTR NBD heterodimer structure, as it appears particularly relevant from a physiological transporter function perspective (see above) and because its three-dimensional structure has been solved in the presence of ATP. Indeed, in MJ0796, mutation of the Walker B glutamate residue (into a glutamine) induces the formation of a stable dimer, which is, however, unable to hydrolyze the ATP molecules trapped in the active sites [21]. In addition, research in the last few years, has revealed that the CFTR NBD heterodimer is characterized by both a functional and sequence asymmetry of its nucleotide-binding sites, one site being canonical with consensus sequences, whereas the other has a highly modified, degenerated sequence [11, 32, 33]. Thus, we also undertook a detailed analysis of the asymmetric nucleotide- binding sites in order to highlight the particular properties of the unconventional site which may account for its specific biochemical properties. Here, our model allowed the identification of critical residues, among which several had not yet been described, and suggests how naturally occurring mutations in these regions may affect the structure and/or function(s) of CFTR.

Materials and methods

Structural alignments of the ABC transporter ATPase domains were taken from the FSSP database [34] and were refined after manual superimposition of the three-dimensional structures. The following structures were considered: *E. coli* BtuCD (complexed with vanadate; pdb identifier 1l7t) [22]; the Mg-ADP-bound and nucleotide-free forms of *M. jannaschii* MJ1267 (pdb1g6h and 1gaj, respectively) [16]; the branched-chain amino acid ABC transporter from *T. maritima* (in complex with ATP; pdb 1ji0) [R. Zhang et al. unpublished results]; *S. typhimurium* HisP (in complex with ATP; pdb identifier 1b0u) [14]; human TAP1 (in complex with ADP; pdb 1jj7) [26]; *T. litoralis* MalK (pdb 1G29) [27]; the Mg-ADP-bound (pdb 1f3o) [15] and Na-ATP-bound forms (pdb 1l2t) [21] of *M. jannaschii* MJ0796. The MsbA sequence was also taken into account, although no atomic resolution is available for its solved three-dimensional structure [23]. Structures of *P. furiosus* Rad50 (pdb 1f2t and 1f2u) [18] and *E. coli* MutS (pdb 1e3m) [20] were considered for comparison purposes. The human CFTR sequence [swissprot (sw) CFTR_HUMAN], as well as other members of the ABC transporter superfamily were aligned with the multiple alignment of ABC transporter ATPase domains (or NBDs) using the HCA method [28, 29]. Basically, the HCA bidimensional representation of protein sequences adds information about secondary structures to the comparison of the one-dimensional structure. Indeed, the hydrophobic clusters defined in this way are statistically centered on secondary structures, which are much more conserved than amino acid sequence [35, 36]. PSI-BLAST [37] was also used to identify some anchor points of the alignment relative to other ABC transporters. Three-dimensional modeling of the human CFTR NBD heterodimer was done with Modeller [38], using as template the experimental structure of MJ0796, in complex with ATP and solved at 1.9-Å resolution [21]. This stable ATP-bound MJ0796 dimer was obtained using a form in which the hydrolytic glutamate Glu171 has been substituted by a glutamine residue and in which the active site cofactor was Na^+ rather than Mg^{2+}. Modeling of the CFTR heterodimer was performed in the presence of ATP, cofactor ion and different numbers of water molecules (2047, 2049, 2066, 2087, 2110, 2270) within the active sites. Coordinates were checked using Verify 3D [39]. Three-dimensional structures were manipulated using Swiss-PdbViewer [40]. Solvent-accessible surfaces of residues were calculated using DSSP [41].

Results and discussion

Modeling of the human CFTR NBD dimer: general considerations

CFTR NBD sequences 1 and 2 were aligned with a structure-based sequence alignment of ABC NBDs for which three-dimensional structures were available (fig. 1). The latter alignment was deduced from automatic structure superimposition, and refined manually within the most divergent regions. Such a structure-based alignment made possible the identification of conserved residues that are crucial for the protein fold and can serve as critical ‘anchor points’ for alignment of the CFTR sequences. Amino acid locations which are critical for the fold correspond mainly to buried positions within regular secondary structures and are most often occupied by hydrophobic amino acids or residues which can substitute them (represented with a green background in fig. 1). Such hydrophobic positions, which may be termed ‘topohydrophobic’ [42], represent, in general, approximately half of the total number of hydrophobic residues in the sequence of a globular domain. A careful analysis was performed on particular regions that were difficult to align (see below). Finally, the MJ0796 NBD dimer was

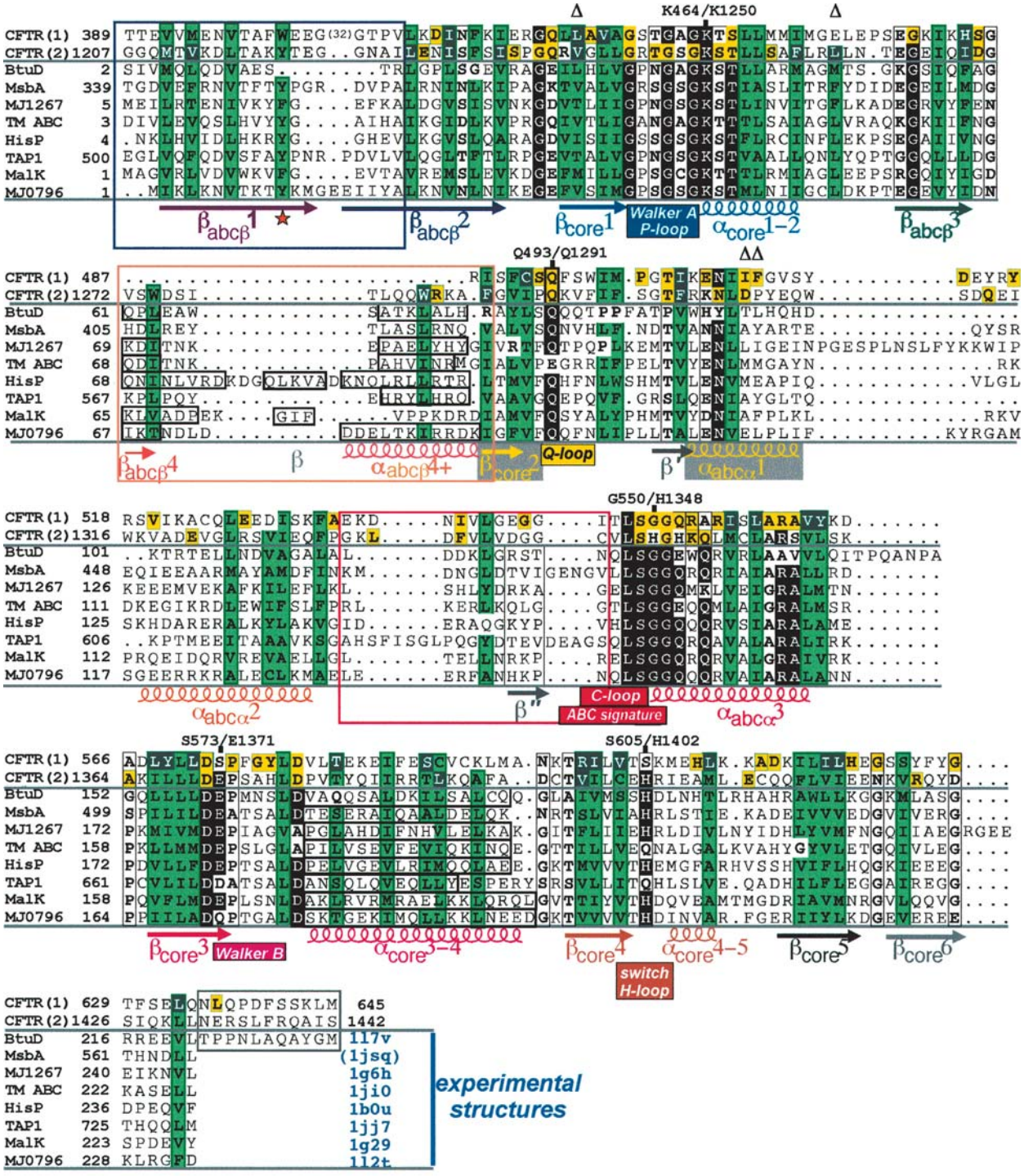

Figure 1. Alignment of the two NBD sequences of human CFTR with those of ABC transporters of known three-dimensional structure. The sequences of the ABC transporters were first aligned following three-dimensional superimposition of their structures. The two CFTR NBD sequences were subsequently aligned using the HCA method, and taking into particular account the anchor points highlighted via the structure-based alignment (residues shown with a green background, see below). The four regions which were difficult to align are boxed (boxes with dark-blue, salmon, red and gray contours) and were studied with particular attention (see text and fig. 3). The positions and labeling of the regular secondary structures observed in the MJ0796 structure are indicated below its sequence. Subscripts in the secondary structure labels indicate the concerned subdomain (e.g., $\beta_{abc\beta}1$ stands for β strand 1 from the ABC-specific β subdomain). When considerable structural variation occurred between the different sequences, the various positions of observed regular secondary structures were boxed (small boxes with white backgrounds and black contours). Positions which are always (or most often) occupied by hydrophobic amino acids (V, I, L, M, F, Y, W) or residues which are often substituted for them (such as A, C, T, S) are shaded green ('topohydrophobic' positions). These positions are often buried, as assessed by calculation of solvent accessibility for the different experimental structures. Black backgrounds indicate residues which are identical in most of the sequences analyzed. Residues mutated in CF patients are depicted white on a green background ('topohydrophobic' positions) or are shaded in yellow (other positions on the CFTR sequence). Δ indicates amino acids which have been found deleted in CF patients. The orange star indicates the position of the conserved aromatic residue involved in adenine stacking. PDB identifiers are indicated in blue at the end of each sequence. Note that the atomic coordinates of MsbA were not used, because the resolution of its crystal structure was too low.

used for modeling the CFTR heterodimer structure as indicated in Materials and methods. Such a strategy allowed us to construct a reliable three-dimensional model, which is presumed to define critical features of the CFTR NBD heterodimer and is further supported by recent experimental data (see below). However, we must stress here that, given the low levels of sequence identity in some regions, the model may be locally incorrect. Our predictions, therefore, which can provide a basis for future structure/function analysis of CFTR and related proteins, await experimental verification and refinement.

According to the alignment given in figure 1 and the resulting model shown in figure 2, the CFTR NBDs consist of an 'F1-ATPase-like' core characterized by a six-stranded β sheet {with strands labeled according to the MJ0796 labeling [21] and from one side to the other: $\beta_{core}2$ (β_c2, yellow), $\beta_{core}3$ (β_c3, pink), $\beta_{core}4$ (β_c4, brown), $\beta_{core}1$ (β_c1, light blue), $\beta_{core}5$ (β_c5 dark gray), $\beta_{core}6$ (β_c6, light gray)}, surrounded by α helices [α_{core}1-2 (light blue), α_{core}3-4 (pink), α_{core}4-5 (brown)]. A small peripheral β sheet constitutes the ABCβ subdomain, which is composed of strands $\beta_{abc\beta}2$ ($\beta_\beta2$, blue), $\beta_{abc\beta}1$ ($\beta_\beta1$, violet), $\beta_{abc\beta}3$ ($\beta_\beta3$, green), $\beta_{abc\beta}4$ ($\beta_\beta4$, salmon) and the helix $\alpha_{abc\beta}4+$ ($\alpha_\beta4+$, salmon). However, the ABCβ subdomain of the first CFTR NBD is highly modified when compared with canonical ones as it has a large insertion between strands $\beta_{abc\beta}1$ and $\beta_{abc\beta}2$ and lacks strand $\beta_{abc\beta}4$ and helix $\alpha_{abc\beta}4+$ (see below). The helical ABCα subdomain [$\alpha_{abc\alpha}1$ ($\alpha_\alpha1$, yellow), $\alpha_{abc\alpha}2$ ($\alpha_\alpha2$, orange), $\alpha_{abc\alpha}3$ ($\alpha_\alpha3$, red)], which contains the ABC signature or C loop, is predicted to interact with the MSD and the cell membrane, according to the BtuCD complex [22]. Of note, residue Phe508, whose deletion accounts for more than 70% of CF cases, is located at the end of helix $\alpha_{abc\alpha}1$, where it is clearly exposed outwards (fig. 2A, B). When superimposing the modeled human CFTR NBD1 and NDB2 with the MJ0796 NBD subunits, root mean square (rms) deviations on Cα positions of 0.48 Å and 0.51 Å were observed, respectively (204 and 217 Cα, respectively). According to the MJ0796 dimer [21], the CFTR NBDs can be aligned side by side in a way that juxtaposes the ABC signature motif of one subunit with the P-loop of the other, thus placing the nucleotide-binding sites at the interface between the two opposing NBDs (fig. 2). Superimposition of the whole CFTR heterodimer with the MJ0796 dimer resulted in rms deviation in Cα positions (421 atoms) of 0.52 Å. The CFTR dimer also fits quite well with the dimeric arrangement of the BtuD structure [22], the greater rms value (1.50 Å for only 191 Cα positions) partially accounting for the relative displacement of the ABCα subdomains with respect to the NBD cores. The buried interface between the two CFTR subunits, to which the two nucleotides make a major contribution, is similar to that observed for the MJ0796 [21] and BtuD dimers [22] and is rather small for a specific dimer contact [22, 43]. Such an arrangement provides an obvious structural basis for the observed cooperativity between the two NBDs of CFTR. In particular, one of the two nucleotide-binding sites of CFTR formed by such an arrangement (site B in fig. 2) fits well the general characteristics of canonical ABC NBD active sites, whereas the other is strikingly modified and may be termed the 'non-canonical' or 'unconventional' site (site A in fig. 2). The features of the two active sites A and B are described in detail below.

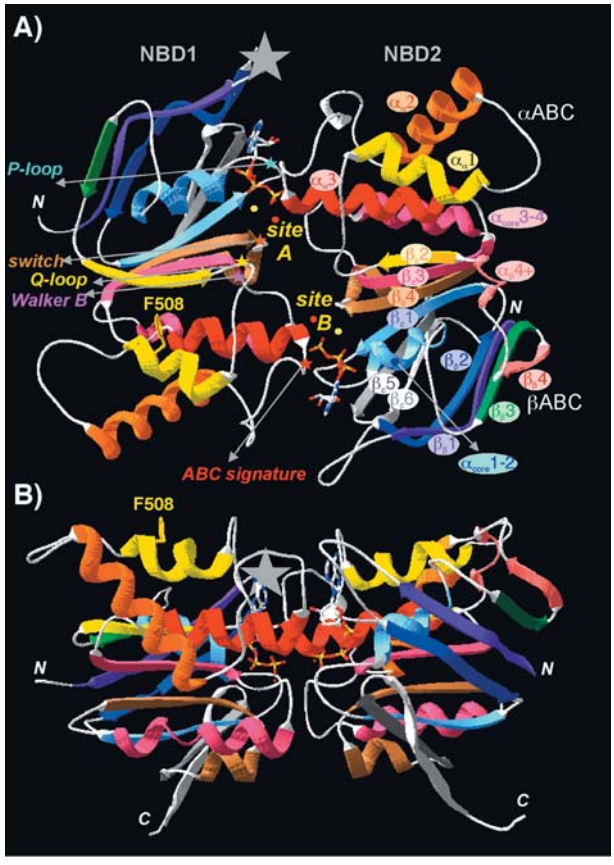

Figure 2. Three-dimensional model of the human CFTR NBD heterodimer. (*A*): Ribbon representation of the overall CFTR heterodimer structure, as modeled on the basis of both the MJ0796 experimental structure [21] and the sequence alignment shown in figure 1. The bound cation-ATP molecules are shown in space-filling representations, with the cation cofactor (Na^+, as in MJ0796) as a yellow dot. The putative hydrolytic water molecules are shown as red dots. The locations of consensus motifs (P-loop, Q-loop, Walker B, switch and ABC signature) are indicated on the left for one of the NBDs, whereas labeling of the secondary structures (colored and numbered as in fig. 1) is shown on the right. The large insertion between strands $\beta_{abc\beta}1$ and $\beta_{abc\beta}2$ of CFTR NBD1 is symbolized by a gray star. The heterodimer model is viewed from the MSDs, along the molecular pseudo two-fold axis, as predicted from the BtuCD structure [22]. (*B*) Orthogonal side view of the heterodimer, with the molecular pseudo two-fold rotation axis running vertically. The membrane bilayer, with which F508 (shown in yellow) likely interacts, is predicted to be located on top of the NBD heterodimer, according to the BtuCD structure [22],

Modeling of the human CFTR NBD dimer: critical points of the sequence alignment

Particular attention was paid to four regions, which were difficult to align and/or for which variations occur in the experimental three-dimensional structures. These regions are boxed in figure 1 and concerned mainly the ABC-specific subdomains. As these subdomains are involved in nucleotide binding, their accurate alignment is critical for a precise definition of the active sites, in particular the 'non-canonical' site largely formed by residues from the first NBD of CFTR (see below).

First, we focused on strand $\beta_{abc\beta}1$ of CFTR NBD1, which was very difficult to align relative to the NBDs of known three-dimensional structures. We thus also took into account additional sequences from eukaryotic ABC transporters to help us align the CFTR sequence with the $\beta_{abc\beta}1$ strands of known structures (fig. 3A). We used the HCA method to identify clusters, which may correspond to this strand having clear, conserved topological features. It allowed us to identify, in the CFTR NBD1 sequence, a hydrophobic cluster (ranging from amino acids 392 to 401 and located after the MSD) whose shape was very similar to that of canonical $\beta_{abc\beta}1$ strands, all of the 'topohydrophobic' positions being represented (fig. 3A). This alignment, however, implies the presence of a large insertion (loop) between strands $\beta_{abc\beta}1$ and $\beta_{abc\beta}2$ (symbolized by a star in figs 2 and 3A), which includes two hydrophobic clusters that might form regular secondary structures (α helices) outside the core of NBD1. Of note, such an insertion is also found between the same strands of the yeast bile pigment transporter BPT1 (fig. 3A). In several NBD structures, such as those of MJ0796 [21] and HisP [14] as well as for CFTR NBD2 (this study), there is a critical aromatic amino acid in strand $\beta_{abc\beta}1$ (often a tyrosine residue as in CFTR NBD2; violet background in fig. 3A) which is involved in stabilization of ATP through stacking interactions with adenine. In CFTR NBD1, residue Trp401 might play a similar role, although the presence of a large insertion just after strand $\beta_{abc\beta}1$ in proximity to the nucleotide-binding site might disturb its interaction with adenine. Worth noting here is that in the structures of the related MutS [20] and Rad50 [18] NBDs, the aromatic residues contacting the ATP adenine are located in other regions, within strand $\beta_{abc\beta}2$ and just before the ABC signature of the opposite subunit, respectively. From these different locations, one may deduce that the ATP adenine ring could have a variable orientation within the active site (data not shown). Thus also possible is that an aromatic amino acid located elsewhere in the CFTR sequence, in particular in the $\beta_{abc\beta}1$-$\beta_{abc\beta}2$ insertion which contains five phenylalanine residues, could be involved in the non-canonical ATP-binding site of human CFTR. Our alignment differs from the previous, contradictory ones performed in an automatic way with NBD sequences from resolved three-dimensional structures (e.g., MJ0796 [16, 21], HisP [14], BtuD [22]). It is, however, in agreement with experimental data from site-directed mutagenesis studies involving the three phenylalanines (F429, F430 and F433) located in the $\beta_{abc\beta}1$-$\beta_{abc\beta}2$ insertion considered as potential candidates for the critical aromatic residue. Substitution of a cysteine for these residues and its subsequent covalent modification showed that the phenylalanines were in fact not important for nucleotide binding [44]. However, no firm conclusion can be made here, because the precise location of these Phe residues remains speculative because of the uncertainties of the model in this region (no direct template for insertion modeling; data not shown). Our analysis also helps to clarify the actual boundaries of the CFTR NBDs, for which there is still controversy in the literature. Indeed, NBD1 and NBD2 of CFTR were initially defined as going from F433 to I586 and from Y1219 to R1386, respectively [1]. Refinements based on structural models constructed using NBD structures from non-ABC proteins [45–47], by functional approaches [48], and by evolutionary data [49] [reviewed in ref. 49] led to the definition of new boundaries, NBD1 extending from P439 to G646 and NBD2 going from A1225 to P1443. These limits exclude, however, sequences that we could align with strand $\beta_{abc\beta}1$ of several NBD structures from ABC transporters. Thus, our analysis clearly shows that the CFTR NBD1 N-terminal limit is located upstream of residue Val392, the first hydrophobic amino acid of the likely NBD1 strand $\beta_{abc\beta}1$. For NBD2, the corresponding residue is Met1210 (figs. 1, 3A). Accordingly, Tyr1219, which is presumed to be involved in ATP adenine binding, is present at the C-terminal end of the NBD2 strand $\beta_{abc\beta}1$. A Y1219C mutant has been shown to exhibit chloride channel activity but is more rapidly inactivated by [2-(triethylammonium)ethyl] methanethiosulfonate (covalent modification of cysteine) than wild-type CFTR [44], a finding highlighting the critical role of this tyrosine residue.

Second, as indicated in figure 1, considerable structure variation occurs within the ABC-specific β subdomain in the region between strands $\beta_{abc\beta}3$ and $\beta_{core}2$, which includes a helix (named $\alpha_{abc\beta}4+$) in most of the known three-dimensional structures (except for MalK [27] where the corresponding region is a coil). As shown in figure 3B, this highly variable region extends from strand $\beta_{abc\beta}3$ (green) to strand $\beta_{core}2$ (yellow), with the helix $\alpha_{abc\beta}4+$ represented in salmon. A short β strand ($\beta_{abc\beta}4$, shown in salmon) is also often present just after strand $\beta_{abc\beta}3$ except in TAP1 [26]. This strand is even longer in HisP [14] and MalK [27] where it is followed by an additional strand (shown in gray in figure 3B). Careful analysis of all these various structures, after their superimposition, allowed us to align with accuracy (where possible) the corresponding sequences. Here, an outstanding feature of the first NBD domain of CFTR is the absence of

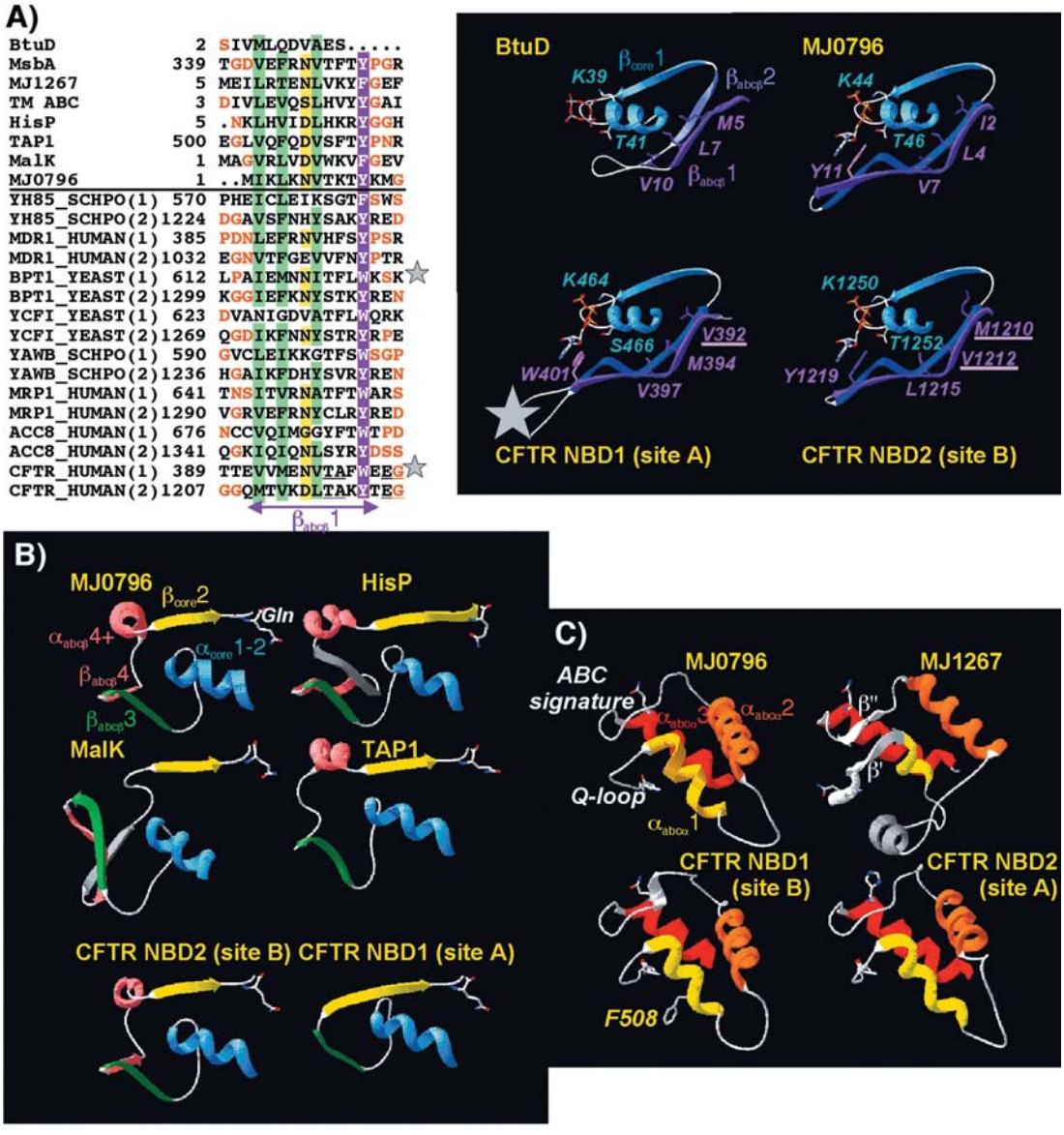

Figure 3. Critical points of the alignment. (*A*) Strand $\beta_{abc\beta}1$. To the left is shown alignment of $\beta_{abc\beta}1$ strands of the two CFTR NBDs. $\beta_{abc\beta}1$ strands of ABC NBDs from known experimental structures were aligned using the HCA method with several $\beta_{abc\beta}1$ strands of eukaryotic ABC transporter NBDs that make intermediate links between CFTR and cognate NBDs. By this route, strand $\beta_{abc\beta}1$ of the first NBD of human CFTR was found to probably correspond to a cluster located between amino acids 392 and 401, a prediction implying the presence of a large insertion between it and strand $\beta_{abc\beta}2$, symbolized here with a gray star (note: NBD1 of the yeast bile pigment transporter BPT1 also has a large insertion in this location). Hydrophobicity is conserved within 'topohydrophobic' positions (shown with a green background; see text). The aromatic amino acid involved in the stacking interaction with the ATP adenine ring in several ABC transporters (such as HisP and MJ0798, but with the exception of BtuD), is shaded violet. Amino acids with high propensities to form coil structures (P, G, D, N, S) [29] and encircling the $\beta_{abc\beta}1$ strands are indicated in orange. Sequences of experimental structures (shown at the top of the alignment) are the same as in figure 1, whereas the other sequences are designated by their swissprot identifiers. On the right are ribbon representations of the three-dimensional structures of strands $\beta_{abc\beta}1$, $\beta_{abc\beta}2$, $\beta_{core}1$ and helix $\alpha_{core}1$ of four ABC NBDs, after superimposition of their P-loops. The experimental structures of MJ0796 (bound to ATP) and BtuD (complexed with vanadate) are compared to the two CFTR NBD models. Like MJ0796, CFTR NBD2 has a tyrosine (Y1219) that contacts the ATP adenine ring. In contrast, CFTR NBD1 has a tryptophan in this position (W401) which is followed by a large insertion (symbolized as a star) which can modify the position of the adenine ring in the ATP-binding site. The lysine and threonine (serine in CFTR NBD1) residues of the P-loop are shown in atomic detail, as are the three hydrophobic amino acids which are buried within strand $\beta_{abc\beta}1$. Amino acids mutated in CF are underlined. (*B*) ABCβ subdomain. The variability within this subdomain is illustrated for four experimental structures, with which the models of the two CFTR NBDs are compared. The structures were superimposed relative to strand $\beta_{core}2$, at the end of which is located the Q-loop (the conserved glutamine being represented in atomic detail). (*C*) ABCα subdomain. The structures were superimposed relative to strand/extended segment β''. The glutamine of the Q-loop and the glutamine (histidine in CFTR NBD2) of the ABC signature are shown in a ball-and-stick representation. Phe508 in CFTR NBD1 is also shown. Note the different orientation of the α helix $\alpha_{abc\alpha}2$ (orange) in MJ1267, the structure of which has been solved in the presence of ADP, relative to that in MJ0796 (solved bound to ATP); this illustrates the rotation of the α-helical subdomain coupled to the loss of contact between the ATP γ-phosphate and the invariant glutamine of the Q-loop during ATP hydrolysis.

the short $\beta_{abc\beta}4$ β strand and of helix $\alpha_{abc\beta}4+$, strand $\beta_{abc\beta}3$ being therefore directly connected to strand $\beta_{core}2$ (fig. 3B). In contrast, NBD2 of human CFTR has here a 'standard' configuration, with a $\beta_{abc\beta}4$ strand and a $\alpha_{abc\beta}4+$ helix that can be accurately aligned with standard NBD structures (compare CFTR NBD2 to TAP1 and MJ0796 in fig. 3B).

The challenge to align was the ABC-specific α subdomain, where the greatest sequence variation is observed. This region includes helices $\alpha_{abc\alpha}1$ and $\alpha_{abc\alpha}2$ (shown in yellow and orange, respectively, in fig. 3C). In human CFTR, helix $\alpha_{abc\alpha}1$ carries Phe508, the deletion of which accounts for 70% of CF cases (http://www.genet.sickkids.on.ca/cftr), whereas, in general, the loop linking $\alpha_{abc\alpha}1$ and $\alpha_{abc\alpha}2$ is variable, sometimes with insertion of additional regular secondary structures, as in MJ1267 (fig. 3C). When considering different ABC transporters, an important feature of this region that has not yet been highlighted is the presence of two short β strands (β' and β'' – white in the MJ1267 structure in fig. 3C) or of extended regions (MJ0796 structure in fig. 3C) before $\alpha_{abc\alpha}1$ and after $\alpha_{abc\alpha}2$, respectively, which can be viewed as the 'arms of a nutcracker'. These two short regular secondary structures may play a key role in NBD signaling as they are preceded by the Q loop and followed by the ABC signature, respectively (fig. 3C). Here, use of HCA led to accurate alignment of helix $\alpha_{abc\alpha}2$ and careful consideration and superimposition of experimental three-dimensional structures encompassing strand β'' allowed us to align this strand with accuracy before undertaking the alignment of the CFTR sequences. Although characterized by strikingly different sequences, in particular within the ABC signature, the overall structures of the two CFTR ABC-specific α subdomains of NBD1 and NBD2 of human CFTR appear not to differ dramatically from each other, in contrast to what is observed for the ABC-specific β subdomains.

Finally, according to the experimental coordinates of the MJ0796 structure, our two CFTR NBD models end at strand $\beta_{core}6$ (light gray in figs 1 and 2). However, in most cases (except for MJ0796), a small helix ($\alpha_{core}6+$) follows the β strand hairpin formed by strands $\beta_{core}5$ and $\beta_{core}6$, whereas the orientation of an eventual second helix ($\alpha_{core}6++$) is highly variable (data not shown). Careful analysis of the corresponding HCA plots leads us to suggest that helix $\alpha_{core}6+$ may be present in both CFTR NBDs, and that it could be followed by an additional helix, as in BtuD (fig. 1 and data not shown). Such limits are in fact in agreement with those defined previously (see above). However, examination of the other experimental structures clearly showed that these helices do not participate in the domain cores, and consequently should have less influence on modeling. Furthermore, studies with deletion mutants of CFTR NBD2 have clearly indicated that these elements are not required for photolabeling of NBD2 by 8-azido-ATP or for NBD2-dependent channel gating [50]. Nevertheless, one cannot finally rule out that these elements may influence NBD functioning, and might act as 'caps' on the active sites.

Features of the human CFTR NBD active sites: general considerations

The dimeric arrangement observed in MJ0796 [21] as well as in BtuD [22] showed that the two nucleotide-binding sites were located at the interface between the two opposing subunits, each site being composed of residues from the two subunits (fig. 2). Examination of our CFTR heterodimer model, which was constructed with ATP and Na^+, as well as with a variable number of water molecules within the active sites (see Materials and methods), allowed us to obtain new information about the precise topological features of the nucleotide-binding sites and the dimer interface. The sequence alignment data presented above already indicated that, in the CFTR heterodimer, there is considerable sequence degeneration within one of the nucleotide-binding sites (called here site A), which is formed by NBD1 motifs and completed by the ABC signature belonging to NBD2 (fig. 2). Indeed, the NBD1 ABCβ subdomain differs strikingly from canonical ones, because it has a large insertion within the loop linking strands $\beta_{abc\beta}1$ and $\beta_{abc\beta}2$ (figs. 1, 3A) and lacks two regular secondary structures (β strand $\beta_{abc\beta}4$ and helix $\alpha_{abc\beta}4+$) (figs. 1, 3B). Sequence analysis also indicated that the hydrophobic cluster [amino acids 380–387; YKTLEYNL], typical of a β strand structure and located upstream from strand $\beta_{abc\beta}1$, might participate in the ABCβ subdomain, substituting the missing strand $\beta_{abc\beta}4$ (data not shown). Moreover, these variations are combined with sequence degeneration within consensus active-site motifs (P-loop, Walker B, switch from NBD1 and ABC signature from NBD2), leading, thereby, to a highly singular site, whose features are discussed in detail below together with some particular features of the dimer interface. In contrast, the other ATP-binding site (termed here site B), formed by NBD2 motifs and the ABC signature of NBD1, is similar to canonical ones. Thus, the two nucleotide-binding sites are clearly not equivalent, active site A having highly particular features.

Features of the human CFTR NBD active sites: contribution of NBD1 to active site A

In the NBD1 P-loop participating in the 'unconventional' active site A, Thr465 and Ser466 are located at the positions of the highly and strictly conserved Ser and Thr residues (Ser45 and Thr46 MJ0796 labeling), which make hydrogen bonds through their side chain oxygen with the β- and α-phosphates of ATP, respectively (figs. 1, 4A). In addition, in MJ0796, Ser45 makes a hydrogen bond with the cofactor ion Na^+. The NBD1 Walker B motif also differs from consensus sequences as a serine

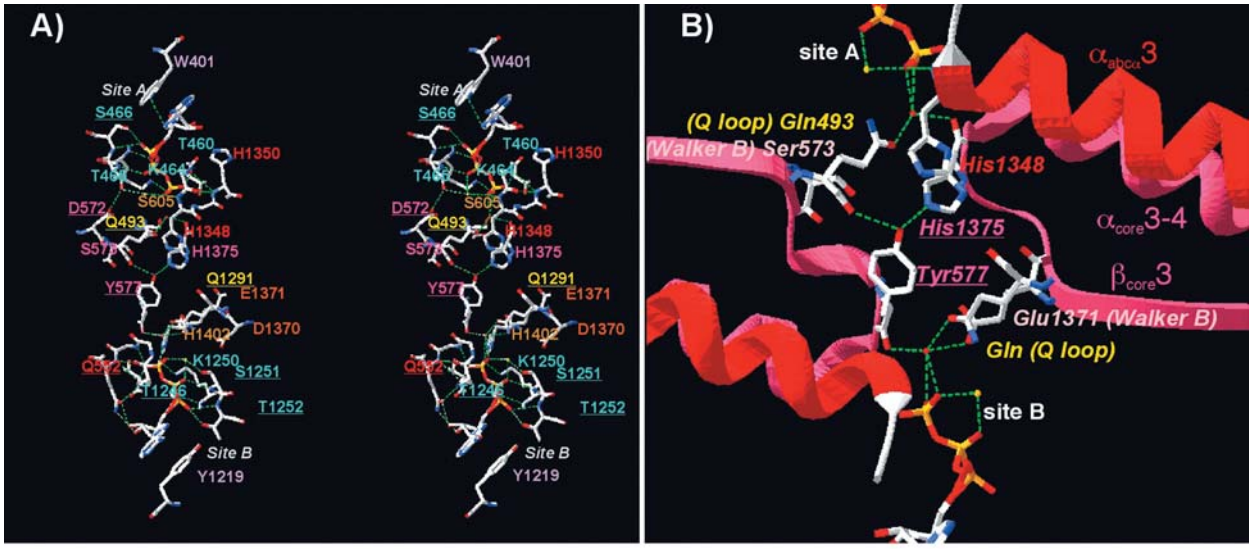

Figure 4. Nucleotide-binding sites and subunit interface of the human CFTR NBD heterodimer. (*A*) Stereo view of the models of the two asymmetric nucleotide-binding sites of the CFTR heterodimer structure. Site B can be compared to canonical ATP-binding sites, as observed in the MJ0796 homodimer. Amino acid labels are colored according to their location in the structure (see figs. 1, 2). The nucleotide-binding sites are formed by the P-loop, Q-loop, Walker B and switch sequences from one subunit and by the ABC signature and β_{core}3-α_{core}3-4 (D loop) of the opposite subunit. Critical amino acids within these sites are shown in ball-and-stick representations. Amino acids that are mutated in CF are underlined. (*B*) Partial view of the nucleotide-binding sites and heterodimer interface, highlighting the particular properties of amino acids from loops β_{core}3-α_{core}3-4 (D loop) and from the ABC signatures (see text).

(S573) replaces the highly conserved glutamic acid (mutated to Gln171 in the hydrolytically inactive MJ0796 structure) which normally forms a hydrogen bond with a water molecule interacting with the γ-phosphate of ATP, whereas the aspartic acid residue preceding it and interacting in other NBDs with the divalent cation remains unchanged. Furthermore, in the NBD1 switch, the canonical histidine (His204 in MJ0796) is substituted by a serine (Ser605). Finally noteworthy is that the glutamine of the Q-loop (Gln493) is conserved in active site A. Interestingly, interaction of the side chain oxygen of that Q-loop glutamine with the cofactor ion (Na^+ in the stable dimeric arrangement observed in MJ0798) positions the nitrogen of the same glutamine to donate an H bond to the putative hydrolytic water. Of note, as stated above, in the inactive MJ0796, a second H bond is given to the same water molecule by the side chain amide of Walker B Gln171. By contrast, in the wild-type MJ0796 (with Glu171 instead of Gln171), only the Q-loop glutamine would be able to donate an H bond, while the other partners would be acceptors, resulting in a hydrolytically active molecule [see ref. 21 for details of the proposed mechanism].

Features of the human CFTR NBD active site: contribution of the dimeric arrangement to active site A – implication for the dimer interface

All the changes occurring in the NBD1 consensus sequences (loop $\beta_{abc\beta}$1-$\beta_{abc\beta}$2, P-loop, Walker B, switch) participating in the unconventional site A appear to be correlated with modifications occuring in the ABC signature of CFTR NBD2, which faces the degenerated sequences of NBD1 and where the canonical LSGGQ motif is substituted by an LSHGH sequence. Thus, the strictly conserved glycine normally present in the middle of the ABC signature motif is replaced by a histidine (His1348), which is exposed at the interface between the two subunits, whereas the following histidine (His1350) is substituted for the highly conserved glutamine, implicated in ribose binding in conventional active sites. Another major specificity in the composition of the non-canonical active site A, which has never been highlighted before, is observed in the β_{core}3-α_{core}3-4 loop of NBD2, which in our three-dimensional model is close to the particular LSHGH signature sequence, at the interface of the two subunits (fig. 4B). Important here is that in the MJ0796 structure [21], the putative hydrolytic water makes an H bond with the backbone carbonyl of Ala175 in this loop (in addition to the bifurcated one that it makes with two oxygens atoms of the γ-phosphate), whereas the neighboring Asp177 interacts with Ser20 of the other subunit. In other ABC transporters, a small residue (Ala, Gly, Ser) is generally also located at the Ala175 position in MJ0796 (fig. 1). In contrast, in both NBDs of CFTR (and not only in NBD2 which directly participates in active site A), a bulky amino acid is located precisely at that position: a tyrosine residue (Tyr577) in NBD1 and a histidine (His1375) in NBD2. Worth noting is that, although located in the same β_{core}3-α_{core}3-4 loop, the Walker B critical glutamine and the Ala175-Asp177 couple of the

MJ0796 sequence participate in opposite nucleotide-binding sites (fig. 4B). In CFTR, the backbone carbonyl groups of histidine from the NBD2 $\beta_{core}3$-$\alpha_{core}3$-4 loop (His1375) and of tyrosine from the same loop in NBD1 (Tyr577) appear to make the expected H bond with the catalytic water molecules of sites A and B, respectively. Their side chains are also in position to interact with each other through an additional H bond (fig. 4B). In addition, the position of the imidazole ring of the NBD2 His1375 might be stabilized through interaction with the histidine of the NBD2 modified ABC signature (His1348) or, alternatively, His1348 might be in direct interaction with Tyr577. This interaction between NBD1 Tyr577 and NBD2 His1375 is likely to be crucial for the subunit interaction, as well as for the interdependence of the two nucleotide-binding sites. This model of interaction is further supported by phylogenetic analysis of CFTR sequences [49], in which substitution of His1348 by Asn in fish sequences (dogfish, salmon, and killifish) appears to be correlated to substitution of Tyr577 by His. Finally, one can hypothesize that in site A, the ATP adenine ring might be mobile, and not necessarily in the configuration shown in figures 2 and 4, owing to the influence of the large insertion within the putative loop linking strands $\beta_{abc\beta}1$ and $\beta_{abc\beta}2$. Thus, there may be no absolute constraint on the ribose-interacting residue of the ABC signature and, hence, a histidine (His1350) may be substituted for the glutamine residue of 'canonical' ABC signature motifs. The exact role of this amino acid in this context remains, however, to be determined.

Conclusions and functional implications

The 'head-to-tail' heterodimer formed by the CFTR NBDs is thus characterized by a highly unconventional site A, in which hydrolysis is likely to be severely impaired due to the absence of critical residues, coupled to a conventional site exhibiting the canonical features of ATP-binding sites. Might this sequence/structure non equivalence of the two active sites therefore be the basis for their different functional properties? CFTR chloride channel activity was investigated via electrophysiological experiments, as well as studies with non-hydrolyzable ATP analogues and NBD mutants [for reviews, see refs 32, 51]. Chloride channel activity is dependent on phosphorylation of the R domain by protein kinase A and involves coordinated action of ATP on the two NBDs. Several initial studies established that nucleotide binding and hydrolysis at both NBDs were important for channel gating, but the two NBDs were soon assumed to have diverged functionally. Hwang et al. [52] proposed that ATP hydrolysis at NBD1 is coupled to channel opening, while subsequent binding and hydrolysis at NBD2 stabilizes the open conformation and triggers channel closing. This model is in agreement with that of Zeltwanger and colleagues [51], who suggested that nucleotide binding and hydrolysis at NBD1 was coupled to channel opening, but proposed that the channel can close without nucleotide interaction. In contrast, Gunderson and Kopito [53] proposed that channel opening and closing were driven by ATP binding and hydrolysis at NBD2, respectively, while nucleotide binding and hydrolysis at NBD1 was only a prerequisite for channel opening. This last model is in agreement with the recent studies of Aleksandrov et al. [54] as well as Basso et al. [55], who performed sophisticated photolabeling experiments to investigate the nucleotide interactions underlying channel gating. Basso et al. [55] observed that the nucleotide remained tightly bound at NBD1 for many minutes, in the form of non-hydrolyzed nucleoside triphosphate, a finding suggesting that CFTR channel opening and closing results predominantly from binding and hydrolysis of ATP at NBD2. These results are in agreement with the recent work of Aleksandrov et al. [54] with the Walker A lysine mutants K464A and K1250A, with the finding that NBD2 is a site of rapid nucleotide turnover, while NBD1 is a site of stable nucleotide interaction. Thus, our three-dimensional model, with its highly unconventional nucleotide-binding site A and a standard nucleotide-binding site B (mainly contributed by NBD1 and NBD2, respectively), provides a molecular basis for these very recent experimental results: channel opening and closing may occur through ATP binding and hydrolysis, respectively, at NBD2 (canonical active site B), whereas ATP remains tightly bound, in an unhydrolyzed form, for a long period at NBD1 (unconventional active site A). These results are consistent with the observation made with two other 'asymmetric' members from subfamily ABC-C, SUR1 and MRP1. Indeed, in both cases, experiments have shown that ATP hydrolysis occurred predominantly at NBD2 [56–60]. Accordingly, sequence alignment of SUR1 and MRP1 shows that their NBD1 Walker B motifs (with Asp instead of Glu) and NBD2 ABC signature (FSQGQ and LSVGQ, respectively) are degenerated, a finding suggesting that the particular properties of unconventional sites are mainly contributed by these two motifs (fig. 5). Also interesting to note is that the glutamine residue of the NBD2 ABC signature is conserved in SUR1 and MRP1, in contrast to CFTR NBD2 where it is substituted by His, suggesting that this glutamine interacts with the ATP ribose in a 'conventional' way. This hypothesis is further supported by the correlated presence of a 'canonical' strand $\beta_{abc\beta}1$ in NBD1 of these proteins, with a 'critical' aromatic residue (Trp688 in SUR1 and Trp653 in MRP1, corresponding to Trp401 in CFTR NBD1) and without an insertion between strands $\beta_{abc\beta}1$ and $\beta_{abc\beta}2$ (fig. 5).

The present model does not take into consideration the R domain and its probable influence on NBD function. Indeed, phosphorylation of the R domain by protein kinase A enables channel activation by nucleotide interactions at the NBDs [61]. Both stimulatory and inhibitory effects of

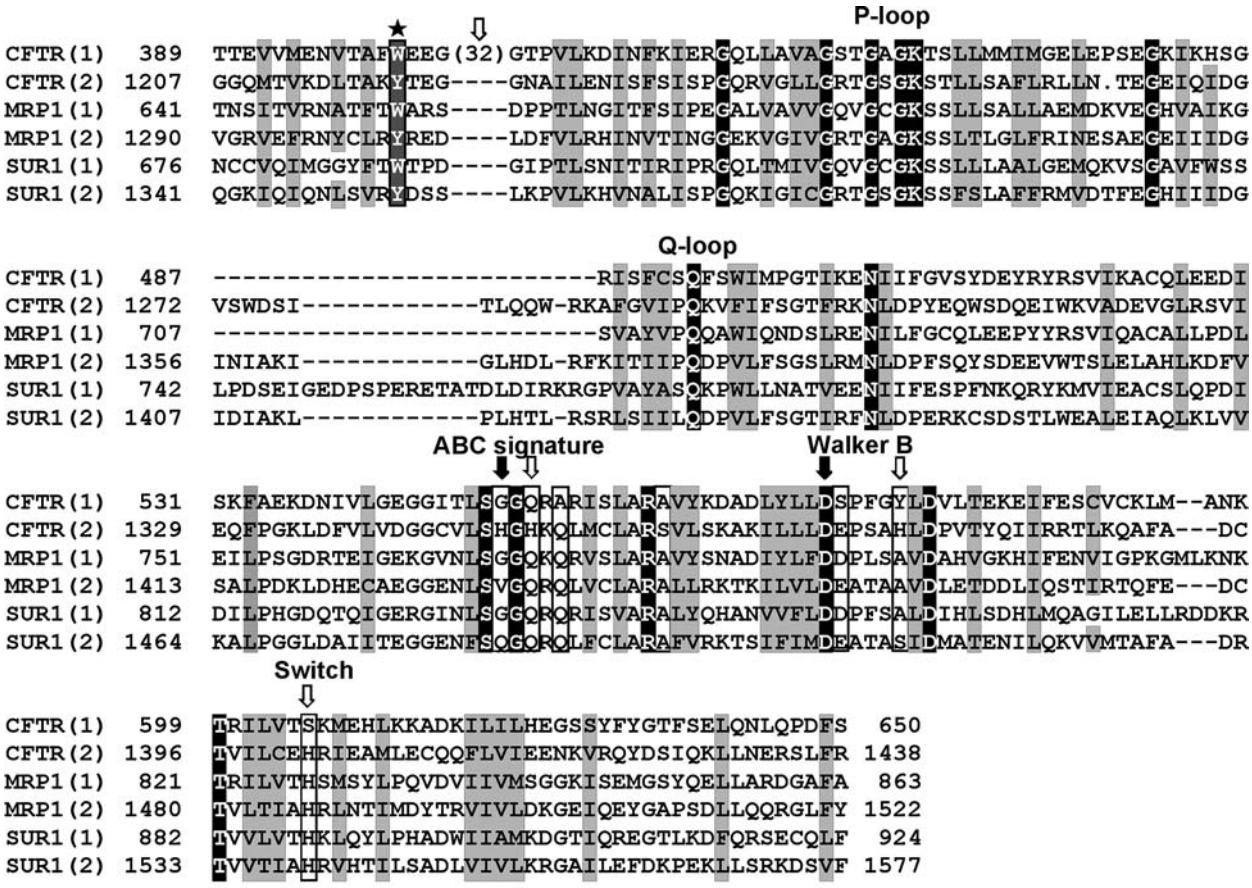

Figure 5. Alignment of the human CFTR sequences with those of MRP1 and SUR1 highlighting the properties of the 'unconventional' sites A. Amino acids which differ from consensus sequences are indicated by arrows (black when differences are found for all three sequences, white when the difference is specific to the CFTR sequence). The star indicates the position of the conserved aromatic residue predicted to be involved in adenine stacking. Swissprot accession numbers of human MRP1 and human SUR1 are P33527 and Q09428, respectively.

phosphorylated and unphosphorylated R domains, respectively, have been proposed [61, 62]. Phosphorylation was very recently suggested to act like an automobile clutch that engages the NBD events to drive gating of the transmembrane ion pore [55]. Thus, with a view to unifying the concepts, one might hypothesize that the basic organization of the NBD dimer is likely similar for all ABC transporters, which generally do not contain structures analogous to the R domain. Accordingly, the R domain might interact with the exposed surface of the NBDs and/or with the MSDs. Such a hypothesis is actually supported by experimental data obtained by use of an overlapping peptide library spanning the intracellular regions of CFTR [63].

Several mutations identified in CF patients are located within the NBD active sites, and might thus be predicted to affect either the dimerization and/or the ATP-binding/ATP-hydrolysis processes. In the absence of an experimental structure for the CFTR heterodimer, our model might permit new insights into the genotype/phenotype relationships of mutations affecting residues located directly in the active sites, as well as residues in other locations on the structure, especially class III and class IV mutations which impact chloride channel regulation/gating and conductance, respectively [64]. For example G551D, a class III mutation affecting the Gly residue located within the NBD1 ABC signature and participating in the 'conventional' active site B, should directly impair the ATP-binding properties of that site, a finding which has been demonstrated experimentally [65]. The G622D mutant (class IV) should be associated with modified properties of the ATP-binding pocket of the 'unconventional' site A. In ongoing work, we are trying to correlate the nature and location on our three-dimensional model of CFTR mutations and their resulting CF phenotype. Residue F508, the deletion of which affects protein maturation (class II mutation) and accounts for more than 70% of chromosomes in CF patients, is not located in the active site but is exposed at NBD1 surface and may therefore play an active role in the NBD interaction with the MSD, as suggested by docking of our model onto the full length ABC structures (data not shown). We should also

be able to gain interesting insights into the dynamic properties of the NBD heterodimer, which are probably crucial for signaling the catalytic events to the MSDs. Preliminary results, obtained by comparison of predicted and observed secondary structures of ABC transporter NBDs and supported by molecular dynamics simulations [66], pinpoint regions within the ABC-specific α subdomain which could be particularly mobile and which we have undertaken to identify and study. Finally, our model may be useful for the rational design of CFTR-specific drugs for pharmacological treatment of CF.

Acknowledgements. Financial support of the association 'Vaincre La Mucoviscidose' (Paris, France) is gratefully acknowledged.

1 Riordan J., Rommens J., Kerem B., Alon N., Rozhamel R., Grzeleczak Z. et al. (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science **245:** 1066–1073
2 Rommens J., Iannuzzi M., Kerem B., Drumm M., Melmer G., Dean M. et al. (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. Science **245:** 1059–1065
3 Sheppard D. and Welsh M. (1999) Structure and function of the CFTR chloride channel. Physiol. Rev. **79:** S23–S45
4 Gadsby D. C. and Nairn A. C. (1999) Control of CFTR channel gating by phosphorylation and nucleotide hydrolysis. Physiol. Rev. **79:** S77–S107
5 Stutts M., Canessa C., Olsen J., Hamrick M., Cohn J., Rossier B. et al. (1995) CFTR as a cAMP-dependent regulator of sodium channels. Science **269:** 847–850
6 Akabas M. (2000) Cystic fibrosis transmembrane conductance regulator. J. Biol. Chem. **275:** 3729–3732
7 Choi J. Y., Muallem D., Kiselyov K., Lee M. G., Thomas P. J. and Muallem S. (2001) Aberrant CFTR-dependant HCO_3^- transport in mutations associated with cystic fibrosis. Nature **410:** 94–97
8 Holland I. B. and Blight M. A. (1999) ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. J. Mol. Biol. **293:** 381–399
9 Higgins C. (2001) ABC transporters: physiology, structure and mechanism- an overview. Res. Microbiol. **152:** 205–210
10 Kerr I. (2002) Structure and association of ATP-binding cassette transporter nucleotide-binding domains. Biochem. Biophys. Acta **1561:** 47–64
11 Schmitt L. and Tampe R. (2002) Structure and mechanism of ABC transporters. Curr. Opin. Struct. Biol. **12:** 754–760
12 Hopfner K.-P. and Tainer J. (2003) Rad50/SMC proteins and ABC transporters: unifying concepts from high-resolution structures. Curr. Opin. Struct. Biol. **13:** 249–255
13 Armstrong S., Tabernero L., Zhang H., Hermodsen M. and Stauffacher C. (1998) Powering the ABC transporter: the 2.5 Å crystallographic structure of the ABC domain of RsbA. Pediatr. Pulmonol. **17:** 91–92
14 Hung L., Wang I., Nikaido K., Liu P., Ames G. and Kim S. (1998) Crystal structure of the ATP-binding subunit of an ABC transporter. Nature **396:** 703–707
15 Yuan Y.-R., Blecker S., Martsinkevich O., Millen L., Thomas P. J. and Hunt J. F. (2001) The crystal structure of the MJ0796 ATP-binding cassette. J. Biol. Chem. **276:** 32313–32321
16 Karpowich N., Martsinkevich O., Millen L., Yuan Y.-R., Dai P. L., MacVey K. et al. (2001) Crystal structures of the MJ1267 ATP binding cassette reveal an induced-fit effect of the ATPase active site of an ABC transporter. Structure **9:** 571–586
17 Jones P. and George A. (1999) Subunit interactions in ABC transporters: towards a functional architecture. FEMS Microbiol. Lett. **179:** 187–202
18 Hopfner K., Karcher A., Shin D., Craig L., Arthur L. and Carney J. (2000) Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. Cell **101:** 789–800
19 Obmolova G., Ban C., Hsieh P. and Yang W. (2000) Crystal structures of DNA mismatch repair protein MutS and its complex with a substrate DNA. Nature **407:** 703–710
20 Lamers M., Perrakis A., Enzlin J., Winterwerp H., Wind N. de and Sixma T. (2000) The crystal structure of DNA mismatch repair protein MutsS bound to a G × T mismatch. Nature **407:** 711–717
21 Smith P., Karpowich N., Millen L., Moody J., Rosen J., Thomas P. et al. (2002) ATP binding to the motor domain of an ABC transporter drives formation of a nucleotide sandwich dimer. Mol. Cell **10:** 139–149
22 Locher K. P., Lee A. T. and Rees D. C. (2002) The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. Science **296:** 1091–1098
23 Chang G. and Roth C. (2001) Structure of MsbA from *E. coli*: a homolog of the multidrug resistance ATP binding cassette (ABC) transporters. Science **293:** 1793–1800
24 Campbell J., Biggin P., Baaden M. and Sansom M. (2003) Extending the structure of an ABC transporter to atomic resolution: modeling and simulation studies of MsbA. Biochemistry **42:** 3666–3673
25 Dorwart M., Thibodeau P. and Thomas P. J. (2002) A structural view of cystic fibrosis. Cystic fibrosis Eur. Netw. **3:** pp. 1–6
26 Gaudet R. and Wiley D. C. (2001) Structure of the ABC ATPase domain of human TAP1, the transporter associated with antigen processing. EMBO J. **20:** 4964–4972
27 Diederichs K., Diez J., Greller G., Müller C., Breed J., Schnell C. et al. (2000) Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon *Thermococcus litoralis.* EMBO J. **19:** 5951–5961
28 Gaboriaud C., Bissery V., Benchetrit T. and Mornon J. P. (1987) Hydrophobic cluster analysis: an efficient new way to compare and analyse amino-acid sequences. FEBS Lett. **224:** 149–155
29 Callebaut I., Labesse G., Durand P., Poupon A., Canard L., Chomilier J. et al. (1997) Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. Cell. Mol. Life Sci. **53:** 621–645
30 Mornon J. P., Prat K., Dupuis F., Boisset N. and Callebaut I. (2002) Structural features of prions explored by sequence analysis. II. A PrP(Sc) model. Cell. Mol. Life. Sci. **59:** 2144–2154
31 Callebaut I., Curcio-Morelli C., Mornon J. P., Gereben B., Buettner C., Huang S. et al. (2003) The iodothyronine selenodeiodinases are thioredoxin-fold family proteins containing a glycoside hydrolase-clan GH-A-like structure. J. Biol. Chem. **276:** 36887–36896
32 Nagel G. (1999) Differential function of the two nucleotide binding domains on cystic fibrosis transmembrane conductance regulator. Biochem. Biophys. Acta **1461:** 263–274
33 Szabo K., Szakacs G., Hegedus T. and Sarkadi B. (1999) Nucleotide occlusion in the human cystic fibrosis transmembrane conductance regulator: different patterns in the two nucleotide-binding domains. J. Biol. Chem. **274:** 12209–12212
34 Holm L. and Sander C. (1998) Touring protein fold space with Dali/FSSP. Nucleic Acids Res. **26:** 316–319
35 Woodcock S., Mornon J. P. and Henrissat B. (1992) Detection of secondary structure elements in proteins by hydrophobic cluster analysis. Protein Eng. **5:** 629–635
36 Hennetin J., LeTuan K., Canard L., Colloc'h N., Mornon J.-P. and Callebaut I. (2003) Non intertwined binary patterns of hydrophobic/non hydrophobic amino acids are considerably bet-

ter markers of regular secondary structures than non constrained binary patterns. Proteins **51:** 236–244
37 Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W. et al. (1997) Gapped-BLAST and PSI-BLAST : a new generation of protein database search programs. Nucleic Acids Res. **25:** 3389–3402
38 Sali A., Potterton L., Yuan F., Vlijmen H. van and Karplus M. (1995) Evaluation of comparative protein modelling by MODELLER. Proteins **23:** 318–326
39 Eisenberg D., Luthy R. and Bowie J. (1997) VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. **277:** 396–404
40 Guex N. and Peitsch M. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis **18:** 2714–2723
41 Kabsch W. and Sander C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers **22:** 2577–2637
42 Poupon A. and Mornon J. P. (1998) Populations of hydrophobic amino acids within protein globular domains: identification of conserved 'topohydrophobic' positions. Proteins **33:** 329–342
43 Bahadur R., Chakrabarti P., Rodier F. and Janin J. (2003) Dissecting subunit interfaces in homodimeric proteins. Proteins **53:** 708–719
44 Berger A. L. and Welsh M. J. (2000) Differences between cystic fibrosis transmembrane conductance regulator and HisP in the adenine ring of ATP. J. Biol. Chem. **275:** 29407–29412
45 Annereau J., Wulbrand U., Vankeerberghen A., Cuppens H., Bontems F., Tümmler B. et al. (1997) A novel model for the first nucleotide binding domain of the cystic fibrosis transmembrane conductance regulator. FEBS Lett. **407:** 303–308
46 Bianchet M., Ko Y., Amzel L. and Pedersen P. (1997) Modeling of nucleotide binding domains of ABC transporter proteins based on a F1-ATPase/RecA topology: structural model of the nucleotide binding domains of the cystic fibrosis transmembrane conductance regulator (CFTR). J. Bioenerg. Biomembr. **29:** 503–524
47 Hoedemaeker F., Davidson A. and Rose D. (1998) A model for the nucleotide-binding domains of ABC transporters based on the large domain of aspartate aminotransferase. Proteins **30:** 275–286
48 Chan K. W., Csanady L., Seto-Young D., Nairn A. C. and Gadsby D. C. (2000) Severed molecules functionally define the boundaries of the cystic fibrosis transmembrane conductance regulator's NH(2)-terminal nucleotide binding domain. J. Gen. Physiol. **116:** 163–180
49 Chen J., Cutler C., Jacques C., Boeuf G., Denamur E., Lecointre G., Mercier B., Cramb G. and Ferec C. (2001) A combined analysis of the cystic fibrosis transmembrane conductance regulator: implications for structure and disease models. Mol. Biol. Evol. **18:** 1771–1788
50 Gentzsch M., Aleksandrov A., Aleksandrov L. and Riordan J. R. (2002) Functional analysis of the C-terminal boundary of the second nucleotide binding domain of the cystic fibrosis transmembrane conductance regulator and structural implications. Biochem. J. **366:** 541–548
51 Zeltwanger S., Wang F., Wang G.-T., Gillis K. and Hwang T.-C. (1999) Gating of cystic fibrosis transmembrane conductance regulator chloride channels by adenosine triphosphate hydrolysis: quantitative analysis of a cyclic gating scheme. J. Gen. Physiol. **113:** 541–554
52 Hwang T.-C., Nagel G., Narin A. and Gadsby D. C. (1994) Regulation of gating of CFTR Cl channels by phosphorylation and hydrolysis. Proc. Natl. Acad. Sci. USA **91:** 4698–4702
53 Gunderson K. and Kopito R. (1994) Conformational states of CFTR associated with channel gating: the role of ATP binding and hydrolysis. Cell **82:** 231–239
54 Aleksandrov L., Aleksandrov A., Chang X.-B. and Riordan J. R. (2002) The first nucleotide binding domain of cystic fibrosis transmembrane conductance is a site of stable nucleotide interaction, whereas the second is a site of rapid turnover. J. Biol. Chem. **277:** 15419–15425
55 Basso C., Vergani P., Nairn A. and Gadsby D. (2003) Prolonged nonhydrolytic interaction of nucleotide with CFTR's NH2-terminal nucleotide binding domain and its role in channel gating. J. Gen. Physiol. **122:** 333–348
56 Ueda K., Inagaki N. and Seino S. (1997) MgADP antagonism to Mg2+ independent ATP binding of the sulfonylurea receptor SUR1. J. Biol. Chem. **272:** 22983–22986
57 Matsuo M., Kioki N., Amachi T. and Ueda K. (1999) ATP-binding properties of the nucleotide-binding folds of SUR1. J. Biol. Chem. **274:** 37479–37482
58 Matsuo M., Tanabe K., Kioki N., Amachi T. and Ueda K. (2000) Different binding properties and affinities for ATP and ADP among sulfonylurea receptor subtypes, SUR1, SUR2A, and SUR2B. J. Biol. Chem. **275:** 21111–21114
59 Gao M., Cui H., Loe D., Grant C., Almquist K., Cole S. et al. (2000) Comparison of the functional characteristics of the nucleotide binding domains of multidrug resistance protein 1. J. Biol. Chem. **275:** 13098–13108
60 Hou Y., Cui L., Riordan J. R. and Chang X.-B. (2000) Allosteric interactions between the two non-equivalent nucleotide binding domains of multidrug resistance protein MRP1. J. Biol. Chem. **275:** 20280–20287
61 Ostedgaard L., Baldursson O. and Welsh M. (2001) Regulation of the cystic fibrosis transmembrane conductance regulator Cl^- channel by its R domain. J. Biol. Chem. **276:** 7689–7692
62 Winter M. and Welsh M. (1997) Stimulation of CFTR activity by its phosphorylated R domain. Nature **389:** 294–296
63 Wang W., He Z., O'Shaughnessy T., Rux J. and Reenstra W. (2002) Domain-domain associations in cystic fibrosis transmembrane conductance regulator. Am. J. Physiol. Cell Physiol. **282:** C1170–1180
64 Rowntree R. and Harris A. (2003) The phenotypic consequences of CFTR mutations. Ann. Hum. Genet. **67:** 471–485
65 Logan J., Hiestand D., Daram P., Huang Z., Muccio D., Hartman J. et al. (1994) Cystic fibrosis transmembrane conductance regulator mutations that disrupt nucleotide binding. J. Clin. Invest. **94:** 228–236
66 Jones P. and George A. (2002) Mechanism of ABC transporters: a molecular dynamics simulation of a well characterized nucleotide-binding subunit. Proc. Natl. Acad. Sci. USA **99:** 12639–12644