Published in final edited form as:

Nat Struct Mol Biol. 2007 August ; 14(8): 738–745. doi:10.1038/nsmb1278.

CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices

Jennifer M R Baker^{1,2}, Rhea P Hudson¹, Voula Kanelis¹, Wing-Yiu Choy^{2,4}, Patrick H Thibodeau³, Philip J Thomas³, and Julie D Forman-Kay^{1,2}

¹Program in Molecular Structure and Function, The Hospital for Sick Children, 555 University Ave., Toronto, Ontario M5G 1X8, Canada.

²Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada.

³Department of Physiology, 6001 Forest Park Lane, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9040, USA.

Abstract

The regulatory (R) region of the cystic fibrosis transmembrane conductance regulator (CFTR) is intrinsically disordered and must be phosphorylated at multiple sites for full CFTR channel activity, with no one specific phosphorylation site required. In addition, nucleotide binding and hydrolysis at the nucleotide-binding domains (NBDs) of CFTR are required for channel gating. We report NMR studies in the absence and presence of NBD1 that provide structural details for the isolated R region and its interaction with NBD1 at residue-level resolution. Several sites in the R region with measured fractional helical propensity mediate interactions with NBD1. Phosphorylation reduces the helicity of many R-region sites and reduces their NBD1 interactions. This evidence for a dynamic complex with NBD1 that transiently engages different sites of the R region suggests a structural explanation for the dependence of CFTR activity on multiple PKA phosphorylation sites.

The CFTR chloride channel, the protein mutated in cystic fibrosis, is a member of the ATP-binding cassette (ABC) superfamily of proteins¹. Like other members of the superfamily, CFTR has two membrane-spanning domains (MSD1 and MSD2) and two nucleotide-binding domains (NBD1 and NBD2). Intracellular regions between the transmembrane segments probably adopt helical structures that extend from the MSDs². Unique to CFTR is the cytoplasmic intrinsically disordered^{3,4} R region, of approximately 200 residues, which we refer to as a region rather than as a domain to reflect its lack of a stable, folded globular structure.

© 2007 Nature Publishing Group

Correspondence should be addressed to J.D.F.-K. (forman@sickkids.ca)..

⁴Present address: Department of Biochemistry, The University of Western Ontario, London, Ontario N6A 5C1, Canada.

AUTHOR CONTRIBUTIONS

J.M.R.B. designed experiments, expressed and purified R region, performed NMR experiments, analyzed data and wrote the manuscript; R.P.H. expressed and purified NBD1; V.K. analyzed data; W.-Y.C. analyzed NMR data; P.H.T. generated the NBD1 construct and advised on NBD1 purification; P.J.T. contributed new reagents and analyzed data; J.D.F.-K. directed the study, designed experiments and wrote the manuscript. All the authors edited the manuscript.

Accession codes. Biological Magnetic Resonance Bank: Chemical shift assignments have been deposited with accession codes 15336 (nonphosphorylated R region) and 15340 (phosphorylated R region).

Supplementary information is available on the Nature Structural & Molecular Biology website.

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

Modulation of normal CFTR channel function involves ATP binding and hydrolysis at the NBDs in combination with phosphorylation by protein kinase A (PKA). Two ATP-binding sites are formed at the interface of the proposed NBD1-NBD2 dimer, as shown by structures of bacterial NBD homodimers⁵, with each composite site comprised of residues from both NBDs. PKA phosphorylation sites are found primarily in the R region, as well as at a single position in the regulatory insertion within the NBD1 sequence⁶. These multiple phosphorylation sites generally act additively to control CFTR channel opening without a requirement for phosphorylation at any one specific site⁷⁻⁹.

Functional studies of full-length CFTR indicate that the R region has multiple effects on CFTR channel activity. Missense Ser→Ala mutations in PKA consensus motifs, introduced singly or in combination, demonstrate various relative contributions of these positions to channel activity^{10,11}. Whereas most phosphorylation sites stimulate channel activity, Ser737 and Ser768 are inhibitory sites. Substitutions at these residues result in increased channel conductance¹⁰, and removal of residues 760–783 or 817–838 produces active channels that open independently of phosphorylation^{12,13}. Coexpression of CFTR₁₋₈₃₅ and CFTR₈₃₇₋₁₄₈₀ also produces low levels of constitutive Cl⁻ channel activity, which is further stimulated with PKA¹⁴. The R region may have an additional stimulatory role, as shown by CFTR channels lacking much of the R region (Δ 708–835/S660A), which gate independently of PKA yet are further stimulated by the addition *in trans* of phosphorylated R region (residues 645–835)¹⁵. This complex regulation of CFTR by its unique R region may be required because CFTR is a channel, unlike other ABC transporters.

One model to explain this mode of regulation involves the binding of multiple phosphorylation sites, potentially with different affinities and effects, to multiple CFTR-binding surfaces, so that increased phosphorylation leads to increased channel activity³. Understanding the dynamic, multisite interactions involved in this model at the molecular level requires specific consideration of the R region as an intrinsically disordered protein segment. Intrinsically disordered protein segments such as the R region exist as an ensemble of rapidly interconverting heterogeneous conformations, rather than having the relatively stable conformation of a folded protein¹⁶. Individual conformations can contain elements of secondary structure, but the population-weighted average conformation has only fractional secondary structure. These protein sequences have lower complexity than folded proteins and are enriched in arginine, lysine, glutamate, proline and serine, with fewer cysteine, tyrosine, tryptophan, isoleucine and valine residues¹⁷, as observed for the R region, with 30% charged residues. Prediction of intrinsically disordered proteins in complete genomes shows an increasing proportion of disordered proteins with increasing organism complexity: up to 14% of archaeal, 21% of bacterial and 41% of eukaryotic proteins are predicted to contain stretches of more than 50 disordered residues¹⁸. Key regulatory cell-signaling proteins and human cancer-associated proteins show greater amounts of intrinsic disorder than proteins involved in metabolism, biosynthesis or degradation¹⁹, and proteins classified as hub proteins (more than ten interaction partners) are enriched in disordered regions compared with end proteins (only one interaction partner)²⁰. The intrinsic plasticity of disordered proteins may facilitate their regulated binding to a variety of binding partners¹⁶. Any stabilization of interacting structural elements upon binding results in entropic penalties that reduce the affinity of the interaction, providing reversibility that is important for proteins involved in inducible events, such as phosphorylation, and for those with multiple interaction partners¹⁶.

As channel gating is controlled by both PKA phosphorylation of the R region and ATP binding and hydrolysis at the NBDs, we hypothesized that the R region has direct phosphorylation-dependent interactions with the NBDs²¹. To test this hypothesis, we have performed NMR studies of the structural properties of the isolated R region in both the

nonphosphorylated and PKA-phosphorylated states, and of the interaction between the R region and NBD1. NMR is the only technique available to examine these properties at the level of individual residues, as crystallization of disordered proteins is not possible and crystal structures including fragments of the R region give only a static picture of one conformation in a dynamic ensemble. The resulting residue-specific structural and binding information demonstrates direct, phosphorylation-sensitive interactions between the isolated disordered R region and NBD1, suggesting a molecular mechanism for channel regulation in full-length CFTR.

RESULTS

The R region is disordered independent of phosphorylation

To examine changes in the structural properties of the isolated R region upon phosphorylation, we phosphorylated purified R region (residues 654–838) *in vitro* with PKA catalytic subunit (phosphorylation sites are shown in **Fig. 1a**). Analysis by mass spectrometry indicated that phosphorylation was achieved at eight or nine sites of the R region. NMR ^1H - ^{15}N correlation spectra for the nonphosphorylated R region and the highly phosphorylated R region (superimposed in **Fig. 1b**) show sharp peaks, with dispersion for backbone amide proton resonances limited to values between approximately 8 and 8.7 p.p.m. in the proton dimension. This limited dispersion is diagnostic of disorder (in contrast to a dispersion of ~7 to 10 p.p.m. for folded proteins) and reflects the rapid interconversion between heterogeneous conformations in disordered proteins, in which all nuclei experience similar average chemical environments²². Phosphorylation produces marked downfield chemical shift changes for phosphoserine residues²³ owing to the addition of a charged phosphate, and there are smaller chemical shift changes for many other residues (see ^1H chemical shift changes in **Fig. 1c**). Phosphorylation does not, however, induce a global folding event; in both the nonphosphorylated and phosphorylated states, the observed ensemble averaged chemical shift values reflect a predominantly disordered state.

Phosphorylation reduces R-region helicity

Assignments were obtained for 97% of the ^1H , ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$ and C' resonances of the nonphosphorylated R-region sequence and 99% of the resonances of the phosphorylated sequence, using a variety of triple-resonance experiments (see Methods). Residues 718–722 in the nonphosphorylated R region lack assignments because of resonance broadening (that is, loss of intensity accompanied by an increase in NMR resonance linewidth), probably resulting from millisecond- to microsecond-timescale sampling of a small population of stabilized conformations for these residues. Analysis of resonance intensities in NMR spectra indicates that the phosphorylated R-region sample was completely phosphorylated at the previously reported PKA phosphorylation sites 660, 700, 712, 737, 753, 768, 795 and 813 (refs. 24,25), with partial phosphorylation (approximately 60%) at Ser670 (ref. 6) and ~15% phosphorylation at the previously uncharacterized PKA phosphorylation site Thr788.

The availability of resonance assignments permits analysis of the fractional population of secondary elements for individual residues in the R region, which was done using the Secondary Structure Propensity (SSP) program²⁶ with $\text{C}\alpha$ and $\text{C}\beta$ chemical shifts as inputs (**Fig. 2a**; for details, see Methods). In folded proteins, consecutive residues have SSP values of +1 for α -helical structure and –1 for β -strand structure, respectively. In disordered proteins, consecutive residues with positive or negative values reflect the local fractional population in α -helical or β -strand conformations as a weighted average of the conformers present. In the nonphosphorylated R region, residues 654–668, 759–764, 766–776 and 801–817 all have a greater than 5% α -helical population, with values above 20% for some residues, whereas 744–753 have a greater than 5% β -strand population, with some values

above 20%. Phosphorylation produces a global decrease in helical content, consistent with that observed in circular dichroism experiments⁴. Phosphoserine (pSer) at the N-terminal cap or in the first three positions of a helix (for example, at position 768) stabilizes the helix, whereas pSer within the helix or at the C terminus (for example, at positions 700, 737 and 813) destabilizes it²⁷. Ser660 is within a long helical segment comprising residues 654–670, which shortens to 660–667 upon phosphorylation so that pSer660 becomes the N-terminal cap residue. This observation is consistent with crystal structures of NBD1 that include R-region residues 654–673, termed the regulatory extension (RE). In these structures, residues 655–668 form a helix in the nonphosphorylated ATP-bound state, but the helix extends only from residue 658 to 668 in the phosphorylated state⁶.

Phosphorylation reduces R-region structural contacts

We next examined the dynamic properties of residues within the R region to provide information on fluctuating structural contacts. Motions on fast timescales were quantified using R_1 , $R_{1\rho}$ and heteronuclear NOE relaxation experiments²⁸ (**Fig. 2b,c**, **Supplementary Table 1** and **Supplementary Table 2** online), and R_2 relaxation rates calculated from R_1 and $R_{1\rho}$ relaxation rates (see Methods). Rates could not be calculated for some residues because of the substantial spectral overlap, as the experiments were recorded as 2D ^1H - ^{15}N correlations. Low R_2 relaxation rates reflect rapid motion (on the picosecond to nanosecond timescale) such as conformer interconversion of disordered proteins, whereas higher relaxation rates are indicative of restricted mobility or slower fluctuations between conformational states (on the microsecond to millisecond timescale). In both its nonphosphorylated and phosphorylated states, the R region has R_2 relaxation rates generally in the range of 4–8 s⁻¹, consistent with the values expected for a disordered protein and lower than those expected for a folded protein with the same number of residues. Uniformly low R_2 relaxation rates throughout the sequence, with further reduced rates at the termini, are expected for a model disordered protein with no structural contacts²⁹. In contrast, the R_2 relaxation rates for the R region are not uniform throughout the sequence: higher rates were observed for a number of residues, including His667, Glu733, Ser768 and Ser809, that are within regions having >5% helical population and close to phosphorylation sites. This suggests that the restriction of motion by secondary structural contacts contributes to higher R_2 relaxation rates, with potential additional contributions from the exchange of individual conformers between different conformational states, tertiary contacts or hydrophobic clusters²⁹. Phosphorylation generally decreases R_2 relaxation rates (see difference between **Fig. 2b** and **Fig. 2c**, shown in **Supplementary Fig. 1** online).

NBD1 predominantly binds nonphosphorylated R region

We added unlabeled NBD1 to ^{15}N , ^{13}C -labeled R-region samples to test for direct interactions between the isolated R region and NBD1, and examined the modulation of these interactions by phosphorylation and ATP binding, key regulatory events for full-length CFTR. The addition of NBD1 leads primarily to reduced peak intensity for specific resonances in the R-region spectrum as well as to small chemical shift changes (**Fig. 3a**). NMR lineshapes reflect the motion of the bond vector measured in the experiment and are sensitive to overall protein tumbling, as they are dependent on the R_2 relaxation rate described above. Disordered proteins have characteristically narrow lineshapes (low R_2 relaxation rates), reflecting rapid conformational averaging in the sample and fairly independent motion of each peptide plane. Changes in peak intensity upon interaction reflect contributions from a number of factors, with the largest contributor being direct binding to the target protein, which causes the interacting residues of the R region to tumble more slowly and leads to broadening of the peak. R-region residues not directly at the interaction interface may also appear to be tumbling more slowly owing to interactions with other R-region residues that bind NBD1 directly. Changes in interactions within the R region and

exchange between free and bound conformations also contribute to the observed line-shape. Examples of effects on the NMR resonances upon binding of NBD1 are shown for Ser790 (**Fig. 3b**) and Thr760 (**Fig. 3c**). Although it is unknown whether we completely saturated the binding reaction (that is, whether all of the R region in the sample interacted with NBD1), we refer to the R-region samples in the presence of the added NBD1 as 'bound'.

Figure 4 plots the ratios of the peak intensities of NBD1-bound R-region resonances to those of free R-region resonances, for nonphosphorylated and phosphorylated R region and for NBD1 in the absence and presence of ATP. If there were no interactions, the ratio would be 1 for all residues. A relatively uniform and low ratio, with only minor deviations, would be expected for all residues if the R region became ordered upon binding, as the entire R region would largely tumble slowly as a unit in complex with NBD1 rather than nearly independently. If very short stretches of extended structure were involved in the interaction, the ratios would dip with sharp minima. Instead, what we observed was ratios varying from 0 to 1.3, with broad rather than sharp minima, reflective of longer stretches of residues interacting with varying affinities and no global disorder-to-order transition. Several segments of the R region appear to bind NBD1, to various degrees, implying dynamic exchange of several R-region binding segments on and off NBD1.

In the nonphosphorylated R region, several segments of the sequence are involved in interactions with apo-NBD1 and ATP-bound NBD1, including the major interaction site surrounding Ser768, where some resonances are broadened to an intensity of 0 (**Fig. 4a,c**). The majority of the interaction sites either include the known *in vivo* phosphorylation sites at residues 660, 700, 737, 795 and 813 (refs. 8,30,31) or are adjacent to them on the C-terminal side. The interaction of residues 661–671 with NBD1 is consistent with crystal structures of NBD1 containing these residues as part of the regulatory extension that contacts the NBD1 core^{6,32,33}. In the context of full-length CFTR, in which NBD1 and the R region are covalently linked, the affinities of the individual interacting segments would be expected to be even higher than those shown here with NBD1 and the R region as separate polypeptides. Phosphorylation of the R region, however, abrogates most binding to NBD1 (**Fig. 4b,d**), with the exception of weaker binding in the vicinity of Ser768 when ATP is present (**Fig. 4d**). Although ATP binding by NBD1 causes few changes in the interaction of NBD1 and the nonphosphorylated R region, ATP binding enhances the interaction of the Ser768 site when the R region is phosphorylated. In the presence of ATP, maintenance of or an increase in peak intensity near Ser790, compared with other residues, is observed for both phosphorylation states (**Fig. 4c,d**). This indicates interactions involving Ser790 in the free R region that are released upon interaction with NBD1.

DISCUSSION

The R region regulates CFTR in a complex manner that is dependent on multiple PKA sites and sensitive to a variety of interaction partners. To aid in understanding of the molecular basis for this mode of regulation, we have characterized the structural properties of the isolated R region and its interaction with NBD1, paying particular attention to the unique properties of disordered protein segments necessary to explain its behavior. The isolated R region is globally disordered, meaning that it samples multiple, heterogeneous conformations having varying degrees of compactness and secondary structural features, with rapid interconversion between conformers. Distinct segments of the R region have fractionally populated local helical structure, reflecting a bias of these residues toward helical conformations in the pool of conformers (**Fig. 2a**). These helices fluctuate within the ensemble, with some conformers having helices that span over 20 residues and others having helical conformations only at residues with the highest fractional secondary structure. The nonuniform R_2 relaxation rates we observed reflect restriction of mobility

caused by the formation of contacts within and potentially between these fluctuating secondary structural elements (**Fig. 2b,c**). Phosphorylation reduces the bias toward helical conformations and decreases many structural contacts within the R region, except in the vicinity of Ser768. The decreases in helicity and R_2 relaxation rates on the N-terminal side of Ser737 are consistent with the marked gel shift previously observed upon phosphorylation at this site, which results from a conformational change³⁴ and may also reflect changes in nonlocal structure.

Two predominant mechanisms have been suggested for the function of intrinsically disordered protein regions in protein recognition. There are many examples of disordered proteins that undergo a disorder-to-order transition upon binding, stabilizing a structural domain within a previously disordered region (for a review, see ref. 35). In an alternative mechanism, sequence motifs (approximately 5–10 residues) that are targets for a modular binding domain, such as the SH2, SH3 or WW domain, are located largely within disordered regions of proteins, and their interactions involve an extended chain (or turn) binding to the modular domain³⁶. The presence of fractional helical structure in the free state of the R region and changes in resonance intensity over several stretches of 10–15 residues upon binding to NBD1 provide evidence that R-region interactions are mediated by stabilization of fluctuating helical structural elements. However, there was no complete disorder-to-order transition upon binding, so our observations represent an intermediate between the above two mechanisms^{37,38}. If a complete disorder-to-order transition occurred, all the resonances of an R region with substantial tertiary contacts would be affected to a similar degree upon binding to the larger NBD1, and the values in **Figure 4a–d** would be similar. Instead, various segments of the R-region sequence show different effects upon addition of NBD1, indicating that several R-region segments bind and dissociate from NBD1 dynamically, so that the intensities observed report the population-weighted average degree of association and relative affinity of each segment. Binding of residues 763–777 is tightest in the nonphosphorylated R region, with almost complete disappearance of the resonances upon binding to NBD1. These residues partially overlap with the segment 771–779, whose sequence matches that of an α -helical molecular recognition element (α -MoRE)³⁸, which is disordered in the free state and undergoes helix stabilization in the bound state.

The coincidence of fractional helical structure in the free R region (indicated by yellow bars) and NBD1-interacting regions is highlighted in **Figure 4**. Transient helices seem to be stabilized upon NBD1 binding, extending the lengths of the helices and their populations, as evidenced by the broad minima observed. The inexact correspondence between free-state helical structure and binding reflects this structural adjustment upon binding as well as electrostatic and other energetic contributions to the interaction. The correspondence between reduction in helical structure for many segments of the R region (**Fig. 2a**) and loss of most NBD1 binding (**Fig. 4**) upon phosphorylation is also indicative of a role for helices in the interaction. Additional support comes from the interaction of the helical regulatory extension with NBD1 in crystal structures⁶ and from the presence of an α -MoRE sequence and the highest R_2 relaxation rates (**Fig. 2b,c**) at residues 763–777, the most stabilized helix³⁸. Together, these data provide evidence for a dynamic complex in which multiple fluctuating helices in the free R region bind NBD1, are transiently stabilized at the interaction interface and are subsequently released.

The model in which multiple R-region segments can bind NBD1 in the context of isolated constructs may explain the lack of requirement for specific phosphorylation sites in the regulation of full-length CFTR^{8,9}. The PKA phosphorylation sites are contained within segments of the R region that interact with NBD1 (**Fig. 4**), so mutation of a subset of phosphorylation sites in the context of full-length CFTR would probably primarily effect CFTR regulation by their coincident interacting segments. The remaining wild-type

phosphorylation sites and their surrounding interaction segments could still act, allowing channel regulation by PKA, consistent with the observed partial channel activity. Milder phenotypes are seen for many cystic fibrosis-causing CFTR missense mutations within the R region, consistent with this multisite behavior, and the majority of these mutations are at the PKA recognition and phosphorylation sites (R709N, S712C, R735K, S737F, V754M, R766M, R810G and S813P; <http://www.genet.sickkids.on.ca/cftr/app>).

A possible interaction surface for NBD1's association with the R region is revealed in NBD1 crystal structures that include the regulatory extension (R-region residues 655–673)^{6,32,33}. The regulatory extension forms a helix that packs against NBD1 along the putative NBD1-NBD2 dimerization interface in many of the structures^{6,32,33}, demonstrating that the R region could inhibit CFTR function by blocking NBD1-NBD2 association, channel opening and nucleotide hydrolysis. Inhibition of NBD1-NBD2 dimerization by the R region is also supported by previously observed effects of phosphorylation. NBD1 initially bound to a glutathione S-transferase (GST)-tagged R region dimerizes more readily with maltose-binding protein (MBP)-fused NBD2 upon phosphorylation *in vitro*³⁹, and phosphorylation also increases NBD1-NBD2 cross-linking⁴⁰ in the context of full-length CFTR. Our data suggest that the association of the regulatory extension helix with NBD1 along the dimerization interface is transient. This is consistent with an NBD1-regulatory extension crystal structure in which the regulatory extension is rotated away from the dimerization interface³², having a different conformation from that of the first structure reported. Studies of CFTR composite channels missing the regulatory extension have found largely wild-type function⁴¹, supporting a model of transient association of multiple R-region segments with NBD1 in full-length CFTR. The regulatory extension helix observed in crystal structures determined in various nucleotide-bound states⁶ is an amphipathic helix, packing against a primarily hydrophobic surface of NBD1 and burying 880 Å² of surface area. According to helical-wheel analysis, residues 733–739 and 761–777 (depending on the side chain conformation of Arg766) may also form amphipathic helices, and these might bind the same NBD1 surface as the regulatory extension, or other NBD1 surface(s), in a dynamic equilibrium. Additional residues of the R region, including 767–780, which interact with ATP-bound NBD1 independently of phosphorylation, may bind this same NBD1 surface or other NBD1 surface(s)⁴². Further NMR studies are currently in progress that focus on NBD1 to clarify the sites of R-region interactions.

Both ‘switch’ and ‘rheostat’ mechanisms are known modes of regulation via phosphorylation of disordered proteins. Multiple phosphorylated sites of the intrinsically disordered cyclin-dependent kinase inhibitor Sic1 bind a single site on the Cdc4 component of SCF ubiquitin ligase, with the dynamic equilibrium of this interaction allowing high-affinity switch-like binding⁴³. In contrast, the intrinsically disordered N-terminal segment of the Ets-1 transcription factor has multiple phosphorylation sites that act as a rheostat to control autoinhibition of the ETS domain⁴⁴. The R region may be a similar rheostat, except that primarily its nonphosphorylated state binds. CFTR function is regulated by phosphorylation-dependent modulation of the structural properties of a large number of potential interacting segments, a mechanism that allows a graded CFTR response to PKA binding and phosphorylation^{8,10,11}. Moderate channel activity is possible with low levels of phosphorylation, but additional phosphorylation increases CFTR activity⁸. Phosphorylation of various sites in the R region has generally additive results, with no single phosphorylation site required for channel activity^{8,9} and various impacts on channel activity from mutations at different sites^{10,11}.

Transient interaction of R-region segments with different affinities for NBD1 may act together to shift the equilibrium between NBD1 monomer and NBD1-NBD2 heterodimer, as well as equilibria with other binding partners (**Fig. 5**). In this model, the nonphosphorylated

R region inhibits NBD dimerization by binding NBD1. The dynamic interconversion, as various R-region segments bind NBD1 and dissociate from it, allows PKA to access the R region's PKA consensus motifs, enabling phosphorylation, which leads to an equilibrium shift away from binding NBD1. In addition to facilitating NBD1-NBD2 dimerization, this enhances binding of the R region to the N terminus, lengthening the duration of channel opening⁴⁵, and to the SLC26A3/6 STAS domains, decreasing channel interburst duration⁴⁶. There are certainly other inter- and intramolecular interaction partners of the R region still to be identified that control this equilibrium shift. In the context of full-length CFTR binding to its physiological partners, the large, approximately 200-residue R region and its binding sites may interact with more than one binding partner at a time, with these sites coming on and off of their targets and exchanging binding partners. Further NMR experiments to probe the specific residues of the R region involved in interactions with the N terminus and the NBD2 and SLC26A3/6 STAS domains will shed more light on the detailed mechanisms by which the disordered, phosphoregulated R region controls CFTR function.

METHODS

Expression and purification of R region

The human CFTR (NCBI Protein P13569) R region was expressed in BL21(DE3) CodonPlus RIL cells (Strata-gene) from a plasmid encoding a 185-residue fragment from position 654 to 838 with an N-terminal His₆ tag in the pPROEX HTb vector (Invitrogen). The Ser654 N-terminal boundary was chosen to include the PKA phosphorylation consensus residues on the N-terminal side of the important Ser660 phosphorylation site, as well as Ser654, which is homologous to the Thr654 residue found to act as an N-terminal helix cap in the murine NBD1 crystal structure⁶. The C-terminal boundary was chosen to include a group of negatively charged residues with α -helical propensity suggested to have a functional role in CFTR channel inhibition¹³. The R-region sequence carried the polymorphism Leu833 from the original CFTR cloning paper¹, as the wild-type Phe833 was deleterious for protein solubility (data not shown). Comparison of HSQC spectra of R-region proteins bearing Phe833 and Leu833 showed similar structural properties for the two proteins (data not shown).

Uniformly isotopically enriched R region with either ¹⁵N labeling alone or combined ¹⁵N,¹³C labeling was expressed in BL21 CodonPlus cells grown to A₆₀₀ = 0.6–0.8 and induced with 1 mM IPTG for 12–16 h at 16 °C. The R region was purified from the insoluble lysate fraction using a 6 M guanidinium-HCl purification including Ni²⁺ affinity chromatography and denaturing size-exclusion chromatography on a Superdex 200 column (Pharmacia), followed by HPLC with a Jupiter 10u C4 300A reverse phase column (Phenomenex). Samples were lyophilized, rehydrated, cleaved overnight with His₆-tagged tobacco etch virus protease, Ni²⁺ affinity purified to remove the His tag and protease, and exchanged into the buffer of interest. R region from the soluble fraction was also purified as above, but without HPLC and using nondenaturing buffers, for comparison with R region from the insoluble fraction. Superimposed HSQCs of the two samples (**Supplementary Fig. 2** online) are virtually identical.

PKA phosphorylation

Reactions were done in 50 mM Tris (pH 7.4), 50 mM MgCl₂, 50 mM ATP and 2 mM DTT with 10 μ M R-region protein. R region was incubated for 1 h at 37 °C after addition of 100 units of PKA (Promega) per 300 μ g of protein, then another 1 h with an additional 100 units of PKA per 300 μ g of protein. Reactions were stopped with the addition of 6 M guanidinium-HCl and 0.1% (v/v) trifluoroacetic acid. PKA was removed by reverse-phase HPLC chromatography.

NMR experiments

All NMR data were collected on a Varian Inova 800-MHz spectrometer at 10 °C with a room-temperature, triple-resonance probe with actively shielded gradients. Nonphosphorylated and highly phosphorylated samples of R region (residues 654–838 with Leu833, as above) at 0.25 mM (in 125 mM potassium phosphate (pH 6.8), 125 mM KCl, 2 mM EDTA, 2 mM benzamidine, 2 mM DTT) were used for assignment experiments. Triple-resonance assignment experiments⁴⁷ included CBCA(CO)NNH and HNCACB to examine C α /C β chemical shifts. HNCO and HN(CA)CO experiments were important in the assignment process, because CO chemical shifts have greater dispersions than C α and C β chemical shifts for disordered proteins. Additionally, CCC-TOCSY experiments were recorded to examine correlations between side chain ¹³C nuclei with backbone ¹H and ¹⁵N nuclei⁴⁸. All spectra were referenced using the internal reference sodium 2,2-dimethyl-2-silpentate-5-sulfonate (DSS). Data was processed using NMRPipe⁴⁹ and analyzed using NMRView⁵⁰.

¹⁵N R_1 , $R_{1\rho}$ and heteronuclear NOE values were measured using previously published pulse schemes²⁸. ¹⁵N R_1 values were measured from 2D spectra recorded with the relaxation delays 10.1, 70.6, 151.2, 252, 362.9, 493.9 and 655.2 ms. ¹⁵N $R_{1\rho}$ values were measured from 2D spectra recorded with delays of 2, 13, 28, 45, 66, 90 and 120 ms. ¹⁵N R_2 values for each residue were calculated by correction of the observed relaxation rate $R_{1\rho}$ for the offset of the applied spin-lock rf field to the resonance, using the equation $R_{1\rho} = R_1^2 \cos^2 \theta + R_2 \sin^2 \theta$, where $\theta = \arctan(\omega_{SL}/\Delta\omega)$. ω_{SL} was 1824.8 Hz and 1811.6 Hz for nonphosphorylated and phosphorylated R region, respectively. For both nonphosphorylated and phosphorylated R region, steady-state NOE values with saturation of ¹H were obtained by scanning for 5 s with a 10 s delay between scans, and values without ¹H saturation by using a 15 s delay between scans. All data sets were processed using Fuda (S.M. Kristensen, University of Copenhagen, and D.F. Hansen, University of Toronto, personal communication).

SSP calculations

SSP uses reference chemical shift values for each amino acid type in α -helix, β -strand and random coil conformations and determines a percentage helix or percentage strand value by calculating the experimental chemical shift difference from random coil relative to that expected for stable helix or strand for each amino acid type²⁶. The values are weighted for each amino acid depending on the magnitude of overall chemical shift expected for secondary structure formation. The default averaging of a 5-residue sliding window was used. Phosphorylated serines were excluded from the calculations because of a lack of reference chemical shift values for fully stabilized helix and strand conformations.

Expression and purification of mouse NBD1

NBD1 from mouse CFTR (NCBI protein P26361) was expressed in BL21(DE3) CodonPlus cells from a plasmid encoding residues 389–653 as an N-terminal, His₆-smt3 (SUMO) fusion protein and purified as described⁶. Mouse protein was used because wild-type human CFTR NBD1 is highly insoluble⁶. The final protein was concentrated to approximately 0.2 mg ml⁻¹ and buffer-exchanged into 50 mM NaPO₄, 200 mM NaCl, 5 mM MgCl₂, 5 mM ATP, 5 mM DTT and 2% (v/v) glycerol (pH 7).

R-region interaction experiments with mouse NBD1

HNCO spectra were recorded with 16 transients at 10 °C (as above) on a 50- μ M ¹⁵N, ¹³C R-region sample in the absence and presence of 100 μ M unlabeled mouse NBD1. The solution contained either 50 mM NaPO₄ (pH 6.8), 200 mM NaCl, 5 mM ATP, 5 mM MgCl₂, 2% (v/v) glycerol and 2 mM DTT (for ATP-bound NBD1), or 50 mM NaPO₄ (pH 6.8), 200 mM

NaCl, 2% (v/v) glucose, 2% (v/v) glycerol and 2 mM DTT (for apo-NBD1). Buffer conditions were optimized for NBD1 solubility and resulted in small chemical shift changes, compared with spectra of the free ^{15}N , ^{13}C R region from assignment experiments. Data were analyzed by calculating the ratio of bound to free peak intensity measured in NMRView⁵⁰. A uniform scaling factor averaged from several noninteracting residues was applied to the peak intensities of the R region bound to apo-NBD1, to compensate for a small amount of NBD1 precipitation during the experiment. This precipitation also reduced the amount of R region in solution, as evidenced by the overall decrease in intensity of R region resonances across the full length of the protein.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank R. Muhandiram and L.E. Kay for technical assistance with NMR experiments, D.F. Hansen for assistance with NMR data analysis using Fuda, C.E. Bear and T. Mittag for many helpful discussions, and J.W. Hanrahan for critically reading the manuscript. This work was funded by grants from the Canadian Cystic Fibrosis Foundation and the Canadian Institutes of Health Research to J.D.F.-K. and from the US National Institutes of Health (DK49835) and the Robert Welch Foundation to P.J.T. J.M.R.B. was supported by scholarships from the Natural Sciences and Engineering Research Council of Canada.

References

1. Riordan JR, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science*. 1989; 245:1066–1072. [PubMed: 2475911]
2. Dawson RJ, Locher KP. Structure of a bacterial multidrug ABC transporter. *Nature*. 2006; 443:180–185. [PubMed: 16943773]
3. Ostedgaard LS, Baldursson O, Vermeer DW, Welsh MJ, Robertson AD. A functional R domain from cystic fibrosis transmembrane conductance regulator is predominantly unstructured in solution. *Proc. Natl. Acad. Sci. USA*. 2000; 97:5657–5662. [PubMed: 10792060]
4. Dulhanty AM, Riordan JR. Phosphorylation by cAMP-dependent protein kinase causes a conformational change in the R domain of the cystic fibrosis transmembrane conductance regulator. *Biochemistry*. 1994; 33:4072–4079. [PubMed: 7511414]
5. Smith PC, et al. ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. *Mol. Cell*. 2002; 10:139–149. [PubMed: 12150914]
6. Lewis HA, et al. Structure of nucleotide-binding domain 1 of the cystic fibrosis transmembrane conductance regulator. *EMBO J*. 2004; 23:282–293. [PubMed: 14685259]
7. Rich DP, et al. Regulation of the cystic fibrosis transmembrane conductance regulator Cl⁻ channel by negative charge in the R domain. *J. Biol. Chem*. 1993; 268:20259–20267. [PubMed: 7690753]
8. Chang XB, et al. Protein kinase A (PKA) still activates CFTR chloride channel after mutagenesis of all 10 PKA consensus phosphorylation sites. *J. Biol. Chem*. 1993; 268:11304–11311. [PubMed: 7684377]
9. Cheng SH, et al. Phosphorylation of the R domain by cAMP-dependent protein kinase regulates the CFTR chloride channel. *Cell*. 1991; 66:1027–1036. [PubMed: 1716180]
10. Wilkinson DJ, et al. CFTR activation: additive effects of stimulatory and inhibitory phosphorylation sites in the R domain. *Am. J. Physiol*. 1997; 273:L127–L133. [PubMed: 9252549]
11. Vais H, Zhang R, Reenstra WW. Dibasic phosphorylation sites in the R domain of CFTR have stimulatory and inhibitory effects on channel activation. *Am. J. Physiol. Cell Physiol*. 2004; 287:C737–C745. [PubMed: 15140750]
12. Baldursson O, Ostedgaard LS, Rokhlina T, Cotten JF, Welsh MJ. Cystic fibrosis transmembrane conductance regulator Cl⁻ channels with R domain deletions and translocations show

- phosphorylation-dependent and -independent activity. *J. Biol. Chem.* 2001; 276:1904–1910. [PubMed: 11038358]
13. Xie J, et al. A short segment of the R domain of cystic fibrosis transmembrane conductance regulator contains channel stimulatory and inhibitory activities that are separable by sequence modification. *J. Biol. Chem.* 2002; 277:23019–23027. [PubMed: 11950844]
 14. Csanady L, et al. Severed channels probe regulation of gating of cystic fibrosis transmembrane conductance regulator by its cytoplasmic domains. *J. Gen. Physiol.* 2000; 116:477–500. [PubMed: 10962022]
 15. Winter MC, Welsh MJ. Stimulation of CFTR activity by its phosphorylated R domain. *Nature.* 1997; 389:294–296. [PubMed: 9305845]
 16. Wright PE, Dyson HJ. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* 1999; 293:321–331. [PubMed: 10550212]
 17. Romero P, et al. Sequence complexity of disordered protein. *Proteins.* 2001; 42:38–48. [PubMed: 11093259]
 18. Dunker AK, Obradovic Z, Romero P, Garner EC, Brown CJ. Intrinsic protein disorder in complete genomes. *Genome Inform. Ser. Workshop Genome Inform.* 2000; 11:161–171.
 19. Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z, Dunker AK. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* 2002; 323:573–584. [PubMed: 12381310]
 20. Haynes C, et al. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput. Biol.* 2006; 2:e100. [PubMed: 16884331]
 21. Gadsby DC, Nairn AC. Control of CFTR channel gating by phosphorylation and nucleotide hydrolysis. *Physiol. Rev.* 1999; 79:S77–S107. [PubMed: 9922377]
 22. Dyson HJ, Wright PE. Nuclear magnetic resonance methods for elucidation of structure and dynamics in disordered states. *Methods Enzymol.* 2001; 339:258–270. [PubMed: 11462815]
 23. Bienkiewicz EA, Lumb KJ. Random-coil chemical shifts of phosphorylated amino acids. *J. Biomol. NMR.* 1999; 15:203–206. [PubMed: 10677823]
 24. Neville DC, et al. Evidence for phosphorylation of serine 753 in CFTR using a novel metal-ion affinity resin and matrix-assisted laser desorption mass spectrometry. *Protein Sci.* 1997; 6:2436–2445. [PubMed: 9385646]
 25. Townsend RR, Lipniunas PH, Tulk BM, Verkman AS. Identification of protein kinase A phosphorylation sites on NBD1 and R domains of CFTR using electrospray mass spectrometry with selective phosphate ion monitoring. *Protein Sci.* 1996; 5:1865–1873. [PubMed: 8880910]
 26. Marsh JA, Singh VK, Jia Z, Forman-Kay JD. Sensitivity of secondary structure propensities to sequence differences between alpha and gamma synuclein: implications for fibrillation. *Protein Sci.* 2006; 15:2795–2804. [PubMed: 17088319]
 27. Andrew CD, Warwicker J, Jones GR, Doig AJ. Effect of phosphorylation on alpha-helix stability as a function of position. *Biochemistry.* 2002; 41:1897–1905. [PubMed: 11827536]
 28. Farrow NA, et al. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ¹⁵N NMR relaxation. *Biochemistry.* 1994; 33:5984–6003. [PubMed: 7514039]
 29. Schwalbe H, et al. Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry.* 1997; 36:8977–8991. [PubMed: 9220986]
 30. Cohn JA, Nairn AC, Marino CR, Melhus O, Kole J. Characterization of the cystic fibrosis transmembrane conductance regulator in a colonocyte cell line. *Proc. Natl. Acad. Sci. USA.* 1992; 89:2340–2344. [PubMed: 1372442]
 31. Picciotto MR, Cohn JA, Bertuzzi G, Greengard P, Nairn AC. Phosphorylation of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* 1992; 267:12742–12752. [PubMed: 1377674]
 32. Lewis HA, et al. Impact of the deltaF508 mutation in first nucleotide-binding domain of human cystic fibrosis transmembrane conductance regulator on domain folding and structure. *J. Biol. Chem.* 2005; 280:1346–1353. [PubMed: 15528182]
 33. Thibodeau PH, Brautigam CA, Machius M, Thomas PJ. Side chain and backbone contributions of Phe508 to CFTR folding. *Nat. Struct. Mol. Biol.* 2005; 12:10–16. [PubMed: 15619636]

34. Csanady L, et al. Preferential phosphorylation of R-domain Serine 768 dampens activation of CFTR channels by PKA. *J. Gen. Physiol.* 2005; 125:171–186. [PubMed: 15657296]
35. Dyson HJ, Wright PE. Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* 2002; 12:54–60. [PubMed: 11839490]
36. Pawson T. Protein modules and signalling networks. *Nature.* 1995; 373:573–580. [PubMed: 7531822]
37. Fuxreiter M, Simon I, Friedrich P, Tompa P. Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J. Mol. Biol.* 2004; 338:1015–1026. [PubMed: 15111064]
38. Oldfield CJ, et al. Coupled folding and binding with alpha-helix-forming molecular recognition elements. *Biochemistry.* 2005; 44:12454–12470. [PubMed: 16156658]
39. Howell LD, et al. Protein kinase A regulates ATP hydrolysis and dimerization by a cystic fibrosis transmembrane conductance regulator (CFTR) domain. *Biochem. J.* 2004; 378:151–159. [PubMed: 14602047]
40. Mense M, et al. In vivo phosphorylation of CFTR promotes formation of a nucleotide-binding domain heterodimer. *EMBO J.* 2006; 25:4728–4739. [PubMed: 17036051]
41. Csanady L, Chan KW, Nairn AC, Gadsby DC. Functional roles of nonconserved structural segments in CFTR's NH2-terminal nucleotide binding domain. *J. Gen. Physiol.* 2005; 125:43–55. [PubMed: 15596536]
42. Ma J, Zhao J, Drumm ML, Xie J, Davis PB. Function of the R domain in the cystic fibrosis transmembrane conductance regulator chloride channel. *J. Biol. Chem.* 1997; 272:28133–28141. [PubMed: 9346969]
43. Nash P, et al. Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication. *Nature.* 2001; 414:514–521. [PubMed: 11734846]
44. Pufall MA, et al. Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region. *Science.* 2005; 309:142–145. [PubMed: 15994560]
45. Naren AP. CFTR chloride channel regulation by an interdomain interaction. *Science.* 1999; 286:544–548. [PubMed: 10521352]
46. Ko SB, et al. Gating of CFTR by the STAS domain of SLC26 transporters. *Nat. Cell Biol.* 2004; 6:343–350. [PubMed: 15048129]
47. Sattler M, Schleucher J, Griesinger C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* 1999; 34:93–158.
48. Kanelis V, Forman-Kay JD, Kay LE. Multidimensional NMR methods for protein structure determination. *IUBMB Life.* 2001; 52:291–302. [PubMed: 11895078]
49. Delaglio F, et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* 1995; 6:277–293. [PubMed: 8520220]
50. Johnson BA, Blevins RA. NMRView—a computer program for the visualization and analysis of NMR data. *J. Biol. NMR.* 1994; 4:603–614.

Figure 1.

R-region phosphorylation. (a) Schematic showing sites of R-region phosphorylation by PKA. (b) ^1H - ^{15}N HSQC NMR spectrum of the isolated human R region, comprising CFTR residues 654–838. Nonphosphorylated R region (black) and phosphorylated R region (red) are superimposed. Owing to the spectral width, the apparent chemical shift of the single tryptophan side chain peak is moved (aliased) to the top left corner (marked with blue asterisk). The true ^{15}N chemical shift is 129.7 p.p.m. Apparent chemical shifts of arginine side chain resonances are also aliased and appear at lower right (dashed blue ellipse). More arginine side chain resonances were observed for the phosphorylated R region, probably because interactions between phosphates and arginine guanidinium groups led to reduced solvent exchange. A dashed blue rectangle marks glutamine and asparagine side chain resonances at upper right. (c) Calculated ^1H chemical shift difference between nonphosphorylated and phosphorylated R region resonances (nonphosphorylated minus phosphorylated). Red circles indicate phosphorylation sites.

Figure 2.

Structural properties of the free R region. **(a)** SSP values calculated for the free phosphorylated or nonphosphorylated R region and averaged over a sliding window of 5 residues. Positive SSP values reflect the fractional helical structure for each residue in a weighted average of R-region conformations. Negative SSP values reflect the fractional β -strand structure. Red circles indicate phosphorylation sites. **(b,c)** R_2 relaxation rates for nonphosphorylated **(b)** and phosphorylated **(c)** R region. Error bars indicate propagation of errors calculated from the covariance matrix of the least-squares fit of R_1 and $R_{1\rho}$ values. Black arrows indicate residues with decreased R_2 relaxation rates upon phosphorylation.

Figure 3.

Interaction of the R region with NBD1. **(a)** Superposition of a selected portion of the ^1H - ^{15}N HSQC spectra of nonphosphorylated ^{15}N , ^{13}C -labeled R region in the free state (black) and in the presence of ATP-bound NBD1 (red). Resolved peaks are labeled with their assignments. Minor chemical shift changes with NBD1 binding were readily tracked in HNCO spectra. **(b)** Slices at 115.36 p.p.m. (^{15}N frequency) of the HNCO spectra of free and bound R region, showing the Ser790 peak, the only one that maintains intensity in this experiment upon addition of NBD1. **(c)** Slices at 115.07 p.p.m. (^{15}N frequency) of the HNCO spectra, showing the Thr760 peak, which is weaker upon NBD1 binding, indicating interactions with NBD1. In **b** and **c**, a trace through the approximate center of the peak (dashed line represents slice through 3D data) is shown at the bottom of each spectrum, illustrating the lineshape.

Figure 4.

Analysis of R-region interactions with NBD1. **(a–d)** Graphs plot ratios of ^{15}N , ^{13}C -labeled R-region peak intensities with and without NBD1, from HNCOC experiments. Shown are interactions of nonphosphorylated or phosphorylated R region with apo-NBD1 or ATP-bound NBD1. Red circles indicate phosphorylation sites. Bar colors reflect secondary structural properties of the R region in the absence of NBD1 (**Fig. 2a**). Stretches of at least three consecutive residues with SSP values above 5% (indicating fractional α -helical structure) or below –5% (indicating fractional β -strand structure) are colored blue or yellow, respectively, with the remaining residues colored gray. Error bars are derived from error propagation of the noise level of the spectrum. Bound peak intensities were calculated at the chemical shifts of the bound peak, as some peaks had small chemical shift changes upon interaction with NBD1. Absence of bar indicates that data could not be analyzed for that residue, except for some residues that had ratios of bound to free peak intensity equal to 0, including residues 765, 771 and 774 in **a**, and 763, 766, 771, 772 and 775 in **c**.

Figure 5.

Schematic illustrating how phosphorylation-induced structural changes in the R region lead to a redistribution of binding equilibria with various regulatory interaction partners. R region is shown as red curve, with multiple helices reflecting the fractional helical structure in portions of the sequence. Gray ellipses, putative binding surfaces of interaction partners; green arrows, potential tertiary interactions within the R region; gold arrows, interactions with binding partners (multiple arrows represent dynamic exchange of both multiple R-region binding sites and multiple interaction interfaces, without implying involvement of specific R-region segments). Other components of CFTR are also shown, including membrane-spanning domains (MSDs), intracellular domains (ICDs), nucleotide-binding domains (NBDs) and cytoplasmic, helical N terminus (N tail). **(a,b)** Interactions favored by nonphosphorylated **(a)** and phosphorylated **(b)** R region. Upon phosphorylation, R region has less helical structure and reduced interactions with NBD1 and possibly NBD2, as well as within the R region; NBD1-NBD2 dimerization leads to channel opening.