Cryo-EM structure of the tetra-phosphorylated R-domain in Ycf1 reveals key interactions for transport regulation.

Rodolpho S. A. de Carvalho[1], Md Shamiul I. Rasel[1], Nitesh K. Khandelwal[1,2], Thomas M. Tomasiak[1]

[1]Department of Chemistry and Biochemistry, University of Arizona, Tucson, AZ 85721

[2]Current Address: Department of Biochemistry and Biophysics, University of California – San Francisco, San Francisco, CA 94

Correspondence should be addressed to Tomasiak, Thomas M.

Email: tomasiak@arizona.edu

Author Contributions: R.S.A.C and T.M.T designed research; R.S.A.C conducted protein expression and purification, Cryo-EM studies, viability assay, ATPase activity assay, mass spectrometry, data analysis and manuscript preparation. M.S.I.R performed protein purification viability assay and ATPase activity assay. N.K.K benchmarked protein purification and supported cryo-EM studies. R.S.A.C, M.S.I.R and T.M.T prepared the manuscript.

Competing Interest Statement: The authors declare no competing interest.

Classification: Structural Biology, Membrane Proteins

Keywords: ABC-transporter, r-domain, phosphorylation, PKC, heavy metal

Abstract

Many ATP-binding cassette (ABC) transporters are regulated by phosphorylation on long and disordered loops which present a challenge to visualize with structural methods. We have trapped an activated state of the regulatory domain (R-domain) of Yeast Cadmium Factor 1 (Ycf1) by enzymatically enriching the phosphorylated state. A 3.2 Å cryo-EM structure reveals an R-domain structure with four phosphorylated residues and a position for the entire R-domain. The structure reveals key R-domain interactions including a bridging interaction between NBD1 and NBD2 as well as an interaction with the R-insertion, another regulatory region. We systematically probe these interactions with a linker substitution strategy along the R-domain and find a close match with these interactions and survival under Ycf1-dependent growth conditions. We propose a model where four overlapping phosphorylation sites bridge several regions of Ycf1 to engage in a transport-competent state.

Main Text

Introduction

ATP-binding cassette (ABC) transporters catalyze the translocation of a wide range of substrates across membranes and support a variety of physiological processes including ion homeostasis, adaptative immune responses, lipid distribution, and detoxification. They consist of two cytosolic nucleotide-binding domains (NBDs) that can bind and hydrolyze ATP which promotes conformational changes in their transmembrane domains (TMDs). Consequently, substrate binding and release is allowed in the transmembrane core driving transport (1–4). Mutations in ABC transporters cause numerous clinically relevant diseases, like cystic fibrosis, cancer multidrug resistance, diabetes, atherosclerosis, coronary artery disease, as well as Tangier disease (5–9).

ABC proteins are classified into seven subfamilies (A to G) based on their topology and domain packing. The C-subfamily is the largest ABC subfamily in the human genome and is characterized by a single polypeptide chain, containing two NBDs and three TMDs in most members (accessory TMD0 and common TMD1 and TMD2). The ABCC family also expands the canonical ABC architecture with the addition of a long disordered link between NBD1 and TMD2 called the regulatory domain (R-domain) in some ABCC transporters (4, 10, 11). Some of the relevant transporters included in this family are Cystic fibrosis conductance receptor (CFTR), involved in chloride ion homeostasis; Multidrug resistance protein 1 (MRP1) involved in the secretion of leukotrienes and drugs; the Sulfonylurea receptor 1 (SUR1) a potassium channel that regulates insulin secretion; and the Yeast cadmium factor 1, which is responsible for heavy metal and drug detoxification in yeast.

Several ABCC members present similarities in their regulatory core, especially the R-domain (12, 13). Much of the foundational research of the R-domain extends from CFTR (14). From the original cloning of CFTR, several canonical phosphorylation sites, mostly cAMP-dependent protein kinase (PKA), were discovered (15). These were later shown in patch-clamp experiments to drastically stimulate CFTR transport (15, 16). Further research identified additional kinases, including Casein Kinase II (CKII), Protein Kinase C (PKC), and cGMP-dependent protein kinase (cGK), as modulators of CFTR activity. CKII was explored through *in vivo* cellular assays, while PKC and cGK were examined using patch clamp studies, illustrating the varied and significant roles these kinases play in regulating CFTR function (17–19). In purified samples, mutagenesis on CFTR and Ycf1 can diminish ATPase activity and can impact cellular activity (20, 21). Indeed, R-domain phospho-regulation seems to involve a complex kinase recruitment mechanism in which a myriad of phospho-sites that suggest PKA, PKC, CKII kinases can potentiate or suppress each other in the context of ABCC R-domain phosphorylation (22, 23). Recent work points to a specific PKA site (24), S813, as the predominant rate-limiting site, similar to the Ycf1 site S908 (20, 25).

The interactions of the R-domain with the rest of the transporter architecture have remained a mystery owing to its intrinsically disordered nature. Isolated fragments of the functional R-domain was identified as a random coil by circular dichroism (26). Early attempts to map specific R-domain regions important for activity initially relied on deletion studies (14, 26, 27). A patch-clamp study using CFTR R-domain deletions has shown that R-domain phosphorylation is required for the transport (26). Surprisingly, deletion of the entire R-domain results in only a minimal loss of activity in CFTR (14), with the C-terminal half being more important for PKA-dependent activity. A study with *Xenopus oocytes* expressing CFTR deletion mutants negatively impacted CFTR trafficking to the plasma membrane (27). Recent structures of the CFTR and more recently, Ycf1, have shown the phosphorylated R-domain binds to parts of NBD1 and intracellular loops (28, 29). Our previous triple-phosphorylated Ycf1 structure (ref) showed this arrangement but only showed fragments of the R-domain and was missing the fourth phosphorylation site at S903. Notably, these R-domain positions lie directly over the CFTR F508 site (F713 in Ycf1). This site causes approximately 70% of cystic fibrosis disease-causing mutations and is found in close to the R-insertion and R-domain intersection, which further highlights the importance of the R-domain regulatory environment (26).

More information is now being uncovered about the dephosphorylated state. Newer structures of the Ycf1 dephosphorylated state and the ABCC2 substrate-free state that the R-domain also adopts an autoinhibitory structure in the substrate cavity, and CFTR has shown additional electron density between

both NBDs (30–32). The dephosphorylated Ycf1 structure showed this effect to be dependent on phosphorylation levels. Computational studies of low-energy conformations in the R domain suggest that its phosphorylation-regulated intra- and inter-domain interactions might be mediated by Molecular Recognition Elements (MoREs) (33, 34). These preformed elements promote interaction with disordered binding partners and enable rapid association and dissociation rates (35), which is in agreement with the later structural studies with Ycf1 and its dephosphorylated state (29, 32). Furthermore, these studies have shown that the R domain itself adopts a more compact state, consistent with these observations (36). However, there have been no investigations comparing the stepwise deactivation from the substrate cavity or being able to image a complete interaction of the R-domain. Strikingly, there are no structures of any ABCC member available with the entire intact R-domain in any form, which is thought to be too mobile, to elucidate this mechanism. Although numerous CFTR structures are available in different states, once again they miss scanning the total R-domain sequence, which leaves a major gap in ABC transporter studies.

Here, we enriched the phosphorylated state of Ycf1 PKA sites using an enzymatic strategy used for CFTR to uncover a complete map of the R-domain. We determined the structure of this state to 3.2Å resolution using cryo-EM and observed a new state phosphorylated in four positions along the R-domain with four overlapping phosphorylated motifs, PKC, PKA, and two CKII sites. We developed a new scrambling strategy to make scrambled chimeras in place of deletion constructs to maintain the length of the R-domain while ablating specific sites. We show that scrambling of those regions in context with our cryo-EM model impacts function to some extent with the most powerful effects near the phosphorylation sites. We also show those regions not in contact do not impact cell survival. Taken together, this study provides a detailed description of the R-domain structure, and critical phospho-regulatory interactions as well as revealing cellular effects that match closely to our novel structural observation.

Results

Cryo-EM structure determination of inward-facing wide PKA-phosphorylated Ycf1

We expressed and purified the NBD2 Walker B mutant E1435Q Ycf1 in the *S. cerevisiae* DSY5 strain as described previously (29) and in Materials and Methods. This protein is expressed with a phosphorylated R-domain (**Fig. 1**) as confirmed by phospho-stained SDS-PAGE (**Fig. 1D**) and mass spectrometry analysis (**Table 1**). Since our previous structures of Ycf1 resulted in many cryo-EM classes (29) and a large amount of heterogeneity, we hypothesized that higher phosphorylation levels could drive Ycf1 into a uniform population. We further enriched the phosphorylated state by subjecting purified Ycf1 sample to *in vitro* phosphorylation with PKA as described in the methods in a similar experimental setup to phosphorylated CFTR (28). *In-vitro* phosphorylation yielded a relative increase in the overall phosphorylation level compared to wild-type as judged by the phospho-Q stain (**Fig. 1D**), and the phosphorylation on S903, S908, T911, and S914 was confirmed with the mass spectrometry data (**Supplemental Fig. 4**), consistent with the previous wild-type purified structure (29).

Consequently, while measuring the thermostability of different Ycf1 phospho-states we observe a significant increase, about 2°C, in the melting temperature of our PKA-treated Ycf1 compared to wild-type, and 4°C compared to the dephosphorylated Ycf1 (**Fig. 1D**). The ATPase activity assay with different phospho-states of Ycf1 has shown that upon PKA treatment the ATP hydrolysis rate increases while the dephosphorylation is reduced compared to the wild-type (**Fig. 2A**). This suggests that phosphorylation modulates the regulation of this transporter which is consistent with previous studies with CFTR (31).

Cryo-EM grids of phosphorylated Ycf1 were prepared as described in the methods section, yielding homogenous particles from which we identified a single inward-facing conformation in a 3.23Å map (**Fig. 1A, Table 2**). The reconstruction yielded high-quality maps for most regions of the transporter and filled gaps from our previous structures (PDB: 7M69 and 7M68), which allowed modeling of the amino gaps in TMD0 (125-132), TMD1 (329-332, 335-340) and NBD2 (1259-1267, 1485-1515). The overall conformation has previously been unobserved with a 40.0Å separation between NBDs, ~5Å wider separation than the previous widest model (G668 on NBD1 and S1411 on NBD2) (**Fig. 1B**).

The structure reveals an R-domain with four phosphorylation sites

The new structure reveals four total phosphorylation sites in the span of 11 residues, S903, S908, T911, and T914. Although we predicted S903 to be a PKC phosphorylation site by a computational prediction tool (Netphos 3.1) (37, 38) with a signature motif of (S/T) x (R/K), the phosphate occurrence at this site had not been structurally confirmed. The new position of phosphorylated S903 shows a rotamer shifted 90 degrees towards NBD1 compared to the previous unphosphorylated models (**Fig. 2B**). The S903 residue and its phosphate group make extensive interactions in a highly basic pocket defined by the surface of NBD1 (L619, K655, H803, K805 and K807) and R-insertion (615-643), especially at the K617 site, the latest being consistent with the R-insertion of CFTR in a proposed alternate conformation (39). Direct contacts amenable to ionic interaction are observed among three out of the six residues in this cluster: the positively charged residues K617, K655 and K805 at 3.0, 3.5 and 3.3 Å respective distances to the phosphate oxygen group. The other three residues K619, H803 and K807, although further apart from the phosphate (~7-8 Å) might be weakly contributing through long-range electrostatic interactions. The ATPase activity assay conducted on the S903A and S903E mutants further underscores the regulatory significance of this interaction site, demonstrating a significant decrease in the ATP hydrolysis rate of the mutants (**Fig. 2A**).

This new arrangement sheds light on the peculiar arrangement of phosphorylation sites. From the confirmation of S903 phosphorylation in the Ycf1 sequence, we observe that the PKC-PKA-CKII-CKII sites (phosphorylated on S903-S908-T911-S914) share their kinase consensus sequences in an overlapping manner (**Fig. 2C-D**). The individual PKC, PKA and CKII recognition motifs are shared by at least one residue to the next kinase motif in the R-domain sequence: R905 shared by PKC and PKA motif; S908 shared by PKA and CKII and T911 shared between two CKII motifs (**Fig. 2C**). This stabilized tetra-phosphorylated motif (PKC-PKA-CKII-CKII) makes a linear segment completely involved or bound to basic charges (**Fig. 2D**). In contrast, in the previous dephosphorylated structure (32), they instead make a compact assembly in the 899-914 alpha helix that must be straightened to achieve this new conformation (**Fig. 2D)**.

The phosphorylated R-domain completely wraps around NBD1

The cryo-EM maps of the enriched tetra-phosphorylated Ycf1 also reveal a continuous R-domain density (**Fig. 1A**) that encloses entirely around NBD1 in path resembling previously published Ycf1 structures (29, 40) . The segments 882-891 and 900-935 were of enough quality to model and refine a polypeptide chain, but the rest of the R-domain (856-881, 892-899) density was too unfeatured to assign the correct sequence register to. Nevertheless, the rest of the density is continuous and reveals a general topology that is in agreement with the predicted AlphaFold model. We generated a second map using the local refinement procedure and particle subtraction in Cryosparc masking out NBD2 to better isolate density around NBD1. In this locally refined map (3.4 Å), we could place a modified Alphafold model with slight adjustments from our structure into this position to map where the remaining R-domain interactions are likely to be (**Supplemental Fig. 2**).

The structure expands on our previous R-domain models (29) and traces its path along the periphery of NBD1. At its N-terminus, the R-domain coils to make a previously unobserved bridging contact with NBD2. The R-domain then returns towards NBD1 to make a helical structure (882-891) followed by a long strand with important bridging contacts between negative charges along the R-domain (E885, D893, D895 and D900) to a series of histidines (H786, H790, H794, H803) (**Supplemental Fig. 2**). Following this segment, the R-domain's four phosphorylated residues extend the electronegative segment with extensive contacts to a positively charged groove along the TMD2-NBD1 surface (**Fig. 3**). Finally, the R-domain ends at the elbow helix of TMD2 with a short helical segment as described before.

Finally, the patterning of charged and hydrophobic interactions along the entirety of NBD1 reveals a common feature in the bound state of intrinsically disordered domains (41, 42). The charged regions of the R-domain are intercalated with hydrophobic contacts. The pairs of R-domain S908-T911-L912-I915 with NBD1 T1145-F1151-I1154, as S914-F917 and I1123 were observed in the previous Ycf1 structure. We now infer the R-domain positions of S903-A901-I902-L904 contacting NBD1 I621-F721-L801-L802.

Cellular evaluation of R-domain contacts through the linker replacement strategy

To assess the importance of the R-domain, we systematically altered segments of the R-domain in a linker replacement strategy by mutation of 6-10 amino acid segments to poly glycine-asparagine-alanine. The rationale was twofold. First, we reasoned that Ycf1 R-domain is shorter than the CFTR R-domain and that deletion constructs previously used in CFTR (14, 26, 27) would interfere with proper folding in Ycf1. Second, we reasoned that polar and flexible constructs would be the least intrusive replacement while still measuring the loss of specific contacts. The linker insertion mutants created were 855-864, 865-874, 875-884, 885-894, 895-904, 905-912, 913-919, 920-926 and 927-935, numbered as the residues of the segment scrambled (**Fig. 4**).

The effect of each insertion was tested on a cadmium survival assay (100 μM $CdCl_2$) in a Ycf1 knockout strain dependent on mutant Ycf1 function for survival. We also assessed the full-R-domain-scrambled construct and as controls, we evaluated the E1435Q catalytic dead mutant and the empty vector that is missing the Ycf1 gene. The R-domain scrambled mutants generally showed a significant reduction in survival compared to the wild-type, with the strongest effects localized to the regions near the phosphorylation sites (segment 895-904, 905-912) and the C-terminus (927-935) (**Fig. 4A-C**). All other regions in contact with NBD1 in the cryo-EM structure showed defects in growth though to a lesser degree. Unexpectedly and in contrast to CFTR deletion experiments (14), the full R-domain scrambled construct had a severe effect. The general pattern exactly matches the cryo-EM structure, with regions in contact between the R-domain and NBD1 being necessary for survival and those that we observe as not in contact are not important (**Fig. 4B**). This trend overall matches that of the Alphafold model pLDDTT confidence scores, with higher confidence scores being associated with higher requirements for survival.

Strikingly, the segments 884-894 and 927-935, which do not contain phosphorylation sites, had a drastic impact on survival similar to the key phosphorylated sites mentioned earlier (**Fig. 4A**). Thus, the 884-894 region matches closely with the novel R-domain alpha-helix (882-891) modeled in our structure, so we hypothesize that this hydrophobic and charged helix might be generating an important, novel, R-domain interaction with implications in Ycf1 activity. In terms of the 927-935 scrambled importance, our structure also reveals potential charge-charge interaction between the helix and a positively charged region (R1112, R1115, R1119).

Comparatively, the scrambled mutants presented a similar survival trend to what has been observed with single-point mutants (S903A, S908A, T911A or S914A) in previous studies (21, 25, 29, 32). The scrambled 895-904 and S903A had a similar drastic effect on cellular survival. Interestingly, the S903E mutant did not present phospho-mimetic activity as observed in previous studies with S908D/E and T911D and instead led to a loss of function (21). The glutamic acid substitution in S903 may not be sufficient to recall the ionic interaction from the phosphate group, which is extensive in the phosphorylated S903 structure. As mentioned earlier, the S903 binding pocket recruits three immediate ionic partners instead of one, as observed in S908 and T911, that way the single charge provided by the glutamic acid cannot replace the contacts of phosphorylated S903.

The R-domain contains co-evolving kinase motifs in the phosphoregulatory region

In our previous work, we performed an evolutionary coupling analysis using the EVcouplings webserver (43) to understand how the R-domain regions coupled to each other and the rest of the Ycf1 (29). Here we reperform the analysis and focus on how segments of the R-domain co-evolve with each other in light of the interactions that are observed. The R-domain generally reveals few contacts because of the high variability in this region. We find that the entire segment of continuous phosphorylation sites from residue 905 to 914 co-evolve as one linear segment (**Supplemental Fig. 5**). This segment then co-evolves with another segment, 872-875 on the R-domain N-terminus. Interestingly, this site also contains numerous presumptive phosphorylation motifs of PKA, PKC, and CKII arranged in a similar order to the functionally important site in S903-S914. These sites are phosphorylated in our data here and previously published data Strikingly, phosphorylation of S878 seems not to impact cellular transport, as observed in our scrambled mutants and previous studies have also shown no impact on ATPase activity (32).

Discussion

The Ycf1 structure in this study shows an enriched PKA phosphorylated state that unveils new features within the ABCC family transporters which explain the interaction of the phosphorylated R-domain. It describes a new 4th phosphorylated site, S903, that coevolves as a critical unit with S908, T911, and S914 and for the first time provides cryo-EM density that shows the entire path of the R-domain. A new scrambled linker insertion functional assay shows in an unbiased way that this segment is the predominately important region for survival on cadmium, with supporting regions that flank it. Importantly, the cellular assay mirrors the regions of the cryo-EM maps that can be visualized with R-domain density in direct contact with NBD1. This supports a model where R-domain phosphorylation predominantly stabilizes NBD1 through contacts around the periphery of NBD1 (895-919) while a major role and most severe defects are localized to 905-912. This exactly coincides with the pLDDT scores in the Alphafold model of Ycf1.

This structure enables the first modeling of an intact regulatory domain presenting four phosphates occupying their respective binding sites. Density is visible for the entirety of the R-domain (855-935) and matches closely with the Alphafold prediction of this region, while detailed atomic modeling was possible in residues 882-891 and 900-935. Taken together, the secondary structure assignment also closely matches the proposed one for CFTR in an analogous R-domain region of the phosphorylated state found by NMR (44). Previously, two independent phosphorylated Ycf1 structures (7M69 and 7MPE) (29, 40) and the structure of CFTR (44) showed that the NBD1 supports binding of the R-domain, here the extended R-domain regions present a similar morphology further enveloping the NBD1. Comparatively, these phosphorylated structures resemble the overall global conformation of the dephosphorylated auto-inhibited Ycf1 structure (32), except for the R-domain placement. This phosphorylated Ycf1 structure and our previous models suggest that disordered regions can acquire an array of R-domain conformations with distinct structural characteristics displaying regulatory function.

The phosphorylated residue (S903) shown here directly interacts with a highly basic pocket in the lower NBD1 portion and drives Ycf1 function in our cellular assay. The crucial role of serine phosphorylation in the R-domain in CFTR activation has been extensively described (15, 45, 46). Similar studies have suggested S908 phosphorylation as the rate-limiting step for Ycf1 activity (32). Here, we demonstrate how the requirement for S903 phosphorylation is also linked to transport activation. Nonetheless, CFTR studies have shown that combined PKA and PKC stimulation can drastically increase R-domain complex formation and activation (23, 47, 48). The phosphorylated S903 site phosphate engages a distinct triple ionic interaction with the basic groove in the N-terminal region (K617, K655, and K805), which stands out when compared to the other phospho-sites (S908, T911 and S914). Critically, the interaction between the K617 site and S903 involves a region that is homologous to another vital regulatory area in CFTR, known as the R-insertion (28), suggesting an R-domain and R-insertion interaction mechanism. In CFTR, this interaction is likely to be a proposed alternate state of the R-insertion that brings it into contact with the R-domain but is present natively in Ycf1 (39).

Finally, this structure also reveals broad physical-chemical properties important for disordered domains (48). The R-domain is dominated by alternating charged regions from the phosphorylation sites and then aspartates/glutamates interspersed by hydrophobic regions. This feature is suggested to favor binding towards hydrophobic grooves and promote stabilization less stringently (41, 42, 48). This type of dynamic interface is proposed to be advantageous to conformational fluctuations and facilitate posttranslational modification in a signaling network (42, 49–52). Similarly, the hydrophobic lower NBD1 interactions with the R-domain (Supplemental Fig. 2) supported by our 884-894 scrambled mutant indicates that non-phosphorylated networks can also impact activity. This N-terminus segment, although not conserved, still provides information regarding the low enthalpy state that the R-domain needs to achieve in this inward-facing conformation.

In conclusion, our tetra-phosphorylated Ycf1 structure contributes to a detailed understanding of the post-translational modification required for regulation in a relevant group of transporters, specifically the ABCC family. The structural elucidation of a novel phosphorylated residue with a well-defined binding site confirms the existence of additional regulatory sites with different mechanisms in the same regulatory core. This discovery may have implications for numerous other transporter classes featuring similarly unstructured

domains or loops with multiple phosphorylation modes. Furthermore, this precisely identified regulatory region has the potential to serve as a foundation for a complex activation mechanism as well as allosteric modulation in various cellular contexts.

Materials and Methods

Cloning, expression and purification

The *S. cerevisiae* YCF1 (Yeast Cadmium Factor 1) gene was codon-optimized and cloned into the p423_GAL1 yeast expression vector as an N-terminal Flag (DYKDDDDK) and C-terminal deca-histidine (10X His) tagged fusion protein (GenScript) (**Supplementary Fig. 1A**). The E1435Q, R-domain phosphorylation sites and interacting residues mutants were generated by site-directed mutagenesis using a Q5-New England Biolabs PCR mutagenesis kit together with primers synthesized from Millipore sigma and verified by sequencing (Elim Biopharmaceuticals, Inc). For protein expression, the *S. cerevisiae* strain DSY538 (Genotype MATa leu2 trp1 ura3-52 his3 pep4 prb1) was transformed with the Ycf1 expression plasmid construct and grown in YNB-HIS selection media plates. Cellular growth was inoculated into a 50 mL primary culture grown for at least 24 h at 30 °C with shaking at 200 rpm in YNB-His media (0.67% w/v yeast nitrogen base without amino acids, 2% w/v glucose, and 0.08% w/v amino acid dropout mix without histidine). A secondary 750 mL culture of YNB-His media was inoculated with 2% of the primary culture (15 mL) and grown under the same growth conditions for an additional 24 h before induction by adding YPG media (1% w/v yeast extract, 1.5% w/v peptone, and 2% w/v galactose final concentration) from a 4X YPG media stock. The culture was grown for an additional 16-18 h at 30 °C before harvesting by centrifugation at 4000 × g for 45 min at 4 °C.

For protein purification, harvested cells were resuspended with ice-cold lysis buffer (50 mM Tris-Cl pH 8.0, 300 mM NaCl, and complete, EDTA-free protease inhibitor cocktail tablets (Roche)) at a ratio of 3.1 mL/g of cell pellet. Resuspended cells were lysed on ice by bead beating with 0.5 mm glass beads for 8 cycles consisting of 45 seconds of beating, with 5 min between cycles. Lysates were collected by vacuum filtration through a coffee filter and membranes were harvested by ultracentrifugation at 112,967 × g for 1.5 h prior to storage at −80 °C. Membranes were solubilized in resuspension buffer (50 mM Tris-Cl pH 7.0, 300 mM NaCl, 0.5% 2,2-didecylpropane-1,3-bis-β-D-maltopyranoside (LMNG)/0.05% cholesteryl hemisuccinate (CHS) supplemented with protease inhibitor as described above) at a ratio of 15 mL/g of membrane at 4 °C for 4 h. Solubilized membranes were clarified by centrifugation at 34,155 × g for 30 min at 4 °C. The clarified supernatant was filtered through a 0.4 µM filter to remove the insoluble fraction and supplemented with 30 mM Imidazole pH 7.0 immediately before loading at a flow rate of 2 mL/min onto a 5 mL Ni-NTA immobilized metal affinity chromatography (IMAC) column (Bio-Rad) equilibrated in Buffer A (50 mM Tris-Cl, 300 mM NaCl, 0.01% LMNG/0.001% CHS, pH 7.0). Following loading, the column was washed with 10 column volumes (CV) of Buffer A to remove nonspecifically bound proteins then followed by a gradient of Buffer B (50 mM Tris-Cl, 300 mM NaCl, 500 mM Imidazole 0.01% LMNG/0.001% CHS, pH 7.0) consisting of the following step sizes: 6% (10CV), 10% (2CV), 16% (2CV), and 24% (2CV). Protein was eluted with 4CV of 60% buffer B and immediately diluted 10-fold with Buffer A before concentration and 3 rounds of buffer exchange to remove excess imidazole by centrifugation at 3095xg at 4 °C in 100 kDa cutoff concentrators (Amicon). Concentrated, buffer exchanged sample was lastly purified by size exclusion chromatography (SEC) at 4 °C by injecting sample onto a Superose 6 Increase 10/ 300 GL column (GE Healthcare) equilibrated in SEC buffer (50 mM Tris, 300 mM NaCl, pH 7.0) supplemented with either 0.01% LMNG/0.001%CHS or 0.06% digitonin and immediately used for biochemical assay or cryo-EM grid preparation following quantification by BCA Assay (Pierce)

Ycf1 phosphorylation and dephosphorylation reactions

Following size exclusion purification, both wild-type and E1435Q YCF1 were concentrated in a 100 kDa cutoff concentrator (Amicon) and BSA quantification was performed. For phosphorylation reactions, the concentrated protein at 3 mg/mL was incubated with c-AMP Protein Kinase A PKA (NEB) (molar ratio of 40:1) at room temperature for 1 hour. A total reaction mixture of 500 µL was used for downstream size exclusion purification utilizing a Superose 6 10/300 column (GE Healthcare) to remove excess PKA and

phosphorylation reagents. Dephosphorylation reactions were performed with Lambda Protein Phosphatase (Lambda PP) (NEB). 1,000 units of Lambda PP were utilized per µmol of Ycf1 protein in a 500 µL reaction. Following a 1-hour treatment at 30 °C, further size exclusion chromatography was performed in a Superose 6 10/300 column (GE Healthcare) for the removal of dephosphorylation reagents.

Cryo-EM sample preparation and data acquisition

Cryo-EM grids for E1435Q PKA-treated Ycf1 were prepared immediately following a second SEC purification for E1435Q Ycf1 protein after treatment with protein kinase A (PKA) (Supplemental **Fig. 1C**). 5 µL of concentrated E1435Q (5.94 mg/mL) PKA-treated Ycf1 sample was applied to a QF-1.2/1.3-4Au 400 mesh grid (E1435Q Ycf1) purchased from Electron Microscopy Sciences. Grids were placed inside of a Leica EM GP2 equilibrated to 10 °C and 80% humidity. Following a 10 s incubation, the side of the grid to which the sample was applied was blotted on Whatman 1 paper (8 s for wild-type; 3.5 s for E1435Q), then immediately plunged frozen in liquid ethane equilibrated to −185 °C. A total of 8,587 movies were captured for E1435Q, on a Titan Krios at 300 kV equipped with a K3 Summit detector (Gatan) at the Pacific Northwest Center for Cryo-EM. Movies were collected at 22,500X magnification with automated super-resolution mode and defocus ranges of −0.9 to −2.1 µm. Movie frames contained 60 frames with a per frame exposure of 0.9 electrons /Å2 dose rate (~54 electrons /Å2 total dose).

Cryo-EM data processing

The phosphorylated Ycf1 E1435Q dataset was processed in Cryosparc (version 4.2.1)(53). 8,587 micrographs were motion corrected by Cryosparc Patch Motion Correction and drift correction to generate an image stack with a pixel size of 0.822 Å/pixel. The contrast transfer function (CTF) was estimated for dose-weighted micrographs using Cryosparc Patch CTF before particle picking using the automated Blob picker. Interactive selection of particles was performed with Inspect Particle Picks on the total automated picks and subject to reference-free 2D Classification to generate references. Selected 2D classes representing the Ycf1 expected morphology were utilized to a reference-based particle picking with the Template picker function. 2D Classification was once again performed and the best classes were utilized on *ab initio* reconstruction. Several rounds of 3D heterogeneous and non-uniform refinements in CRYOSPARC (53, 54) were performed leading to a 3.89 Å resolution map of phosphorylated Ycf1 obtained (**Supplemental Fig. 6**). This map was then used as a reference for artificial intelligence-assisted TOPAZ (ref) particle picking in Cryosparc in the initial micrograph dataset (**Supplemental Fig. 7**). A total 2,159,582 particles were automatically picked, extracted with 4X binning resulting in a box size of 440 pixels with 6.576 Å/pixel. Multiple rounds of 2D classification were performed to remove bad particles resulting in 1,626,297 particles subject to 3D analysis in ChimeraX following extraction with 2X binning and a box size of 220 pixels with 2.062 Å/ pixel.

Iterative rounds of local and global CTF refinements in association with heterogenous and non-uniform refinement were utilized to achieve a final 3.23 Å map, utilizing 68,169 particles. Local refinement was performed in the final map utilizing particle subtraction, in which NBD2 was masked out to decrease particle heterogeneity from this dynamic domain. Local refinement generated a 3.4Å map with the expanded region of the N-termini R-domain portion in association with NBD1. Cryosparc 3D Flexible Refinement was also performed utilizing the 3.23Å map and particle stack to generate a 3D volume series that represents conformational motion in the phosphorylated sample. Local resolution and FSC validation were performed in Cryosparc. The final and locally refined map was then used for manual model building as specified below. Data processing workflow is shown in Supplementary Figure 3-4, as well as the final EM maps and quality report.

Model building and refinement

An initial model of Ycf1 was built utilizing the ISOLDE suite (version 1.3) in UCSF ChimeraX (version 1.3) (55) and the AlphaFold2 Ycf1 model was used as the initial reference template. Manual model building was performed in both COOT (version 0.9) and ISOLDE ChimeraX (56, 57). Multiple Iterative cycles of real-space refinement and analysis were performed in Phenix (version 1.20.1) and CCP-EM modules (56,57). Secondary structure restraints were extensively used as an additional Phenix restrain. For the structure-building process map blurring and sharpening features of the COOT package were utilized for structure

analysis, especially on locally refined maps for R-domain modeling. Molprobity dedicated web service was also utilized in association with the Phenix platform to optimize geometry. To maintain proper geometry, starting model restraints and harmonic restraints were used extensively in Phenix. Model visualization, analysis and figure preparation were done using UCSF Chimera and UCSF ChimeraX.

ATPase activity assay

For evaluating ATPase activity, wild-type and mutants were expressed and purified as described above in a buffer containing 0.01% LMNG and 0.001% CHS. Different phosphorylation states were achieved by enzymatic treatments (c-AMP PKA and Lambda-PP phosphatase) as described above and a second size exclusion was performed to exclude the enzymatic treatment components, following BCA quantification. The ATPase rates were measured as a colorimetric endpoint assay for inorganic phosphate detection described by (58), which allowed the indirect quantification of protein-driven ATP hydrolysis. The Michaelis-Menten kinetics for the transport of four replicates (n = 4) at an ATP range (0.05 to 4 mM) are reported as the mean ± standard quantified in GraphPad Prism. Reaction was performed at near the physiological growth temperature (30°C) of its host organism, *S.cerevisae,* for 30 min. The ATPase reaction was started by the addition of an Mg-ATP solution prepared in Tris buffer, bringing the final concentration of Ycf1 to 0.135 mM in a 50 mL reaction containing 10 mM MgCl2 and varying final concentrations of ATP (0–4 mM). Samples in the absence of Mg-ATP solution or Ycf1 were prepared as negative controls. The reaction was quenched for a single time point by adding 40 mL of 5% SDS. Following quenching, 200 mL of detection reagent (8.75 mM Ammonium Molybdate, 3.75 mM Zinc Acetate, pH 5.0, and 7.5% Ascorbic Acid pH 5.0 prepared fresh prior to use) were added to each sample and incubated for 25 minutes at 37°C for revelation. Data were fit using nonlinear regression in GraphPad Prism 9 to derive and Vmax values.

LC-MS/MS analysis of phosphorylated Ycf1

LC-MS/MS analysis was used to identify the phosphorylated residues of Ycf1. LC-MS/MS was performed at the University of Arizona mass spectrometry core using a Thermo Fisher Q Exactive Plus. Separations were performed on an Acclaim PepMap RSLC column (75 micron x 25 cm) using a gradient of solvent A (water and 0.1% formic acid) and solvent B (acetonitrile, 0.1% formic acid). The gradient uses 3-20% solvent B over 90 min, 20-50% solvent B over 10 min, 50-95% solvent B for 10 min, and finally a 95% solvent B for 10 minutes with a final 3% solvent B for 10 minutes. The Xcalibur v 4.0.27.19 software (59) was used for data-dependent acquisition with a 70,000 resolution. A range of 350-1500 mass/charge (m/z) with automatic gain control set to 1 x 10^6 and a maximum injection time (IT) of 65msec was used and the 10 most intense ions are subjected to higher-energy collisional dissociation (HCD) at 27 normalized collision energy with 1.5m/z, AGC of 5e4, and maximum IT of 65msec and dynamic exclusion used to exclude single MS/MS ions for 30 seconds following an acquisition. Ions with charge states of +1, >7, unassigned, or isotopes are excluded.

MS and MS/MS data were analyzed against the *Saccharomyces cerevisae* UniProt database containing 9124 sequence entries, using Thermo Proteome Discoverer v 2.2.0388 with additional common contaminant proteins added. Trypic peptides were identified and considered matches with up to 2 missed cleavage sites and variable modifications considered include phosphorylation (98.00 Da), methionine oxidization (15.995 Da) and cysteine carbamidomethylation (57.021 Da). Protein identification was performed at 99% confidence with XCorr (60) in a reversed database search and label-free quantification was also performed. Protein and peptide data were further analyzed using Scaffold Q+S v4.8.7. (61). Protein identifications will be accepted that pass a minimum of two peptides identified at 0.1% peptide False Discovery Rate and 90-99.9% protein confidence by the Protein Profit algorithm within Scaffold.

Yeast Spotting Assay

Yeast survival assays were performed with BY4741 - YDR135C Ycf1 knock-out *S. cerevisae* strain as reported in (62). Cells were chemically transformed (Zymogen Inc) with wild-type Ycf1, E1455Q Ycf1 and R-domain scrambled plasmids. Transformants were plated onto YNB-His plates and incubated for 2-3 days at 30°C. Once colonies had developed, individual colonies were inoculated into 5 mL YNB-His liquid media and grown overnight at 30°C while shaking at 200 rpm. Yeast cells were pelleted by slow centrifugation at 800g for 5 minutes, then washed with 0.9% saline solution before pelleting them again. Final resuspension

was done in 1-3 mL sterile water, absorbance at OD600 was read for all cultures. All cultures were then adjusted to 0.5 OD600 (± 5% error). Serial dilutions of each culture was performed in a 96-well plate, 200µL of normalized yeast solution was added to the first row (row A) while the following 3-4 columns were used as technical replicas while 120µL of sterile water was added to rows B to F. Serial dilution on a 1:5 ratio was performed using 30µL of row A culture to the following rows, creating a OD600 serial range of 0.5, 0.1, 0.02, 0.004, 0.0008, 0.00016. YRG agar plates (yeast nitrogen base with ammonium sulfate 0.67% w/v, raffinose 1% w/v, galactose 2% w/v, CSM-His 0.077% w/v and 2% w/v agar) with and without 100 µM $CdCl_2$ were shortly incubated at 30°C at until stamping. Cells were then stamped onto 30°C warm YRG plates containing or not 150 µM $CdCl_2$. Cellularly stamped YRG plates were kept at 30°C for 4 days before imaging on a Biorad ChemiDoc MP Imaging System. 3 technical replicates were performed per plate, and 3 biological replicates were performed in total for each sample.

Evolutionary coupling analysis of Ycf1

Evolutionary coupling analysis was performed using the EV-couplings web server (evcoupling.org) (43) to identify co-evolving residues within the Ycf1 protein (UniProt ID:P39109). The residues K268 to S965 were selected as the search range for the coupling analysis so coverage of the full R-domain region (855-935) could occur. The multiple sequence alignment of the homologous sequence was selected by using the server's default parameters, selecting sequences with a maximum of 80% identity to cover a broad evolutionary spectrum while minimizing redundancy. The co-evolving pairs were then mapped onto the Ycf1 protein structure presented here to identify potential interaction networks and functional sites critical to the R-domain and its phosphorylation sites. Scores were used with a cutoff of ≥ 1 and probability of ≥70% to account for the low conservation of the R-domain sequence.

Acknowledgments

This research received generous support through grants from the National Institute of Allergy and Infectious Diseases (NIAID) under grant number NIH R01 AI156270 (awarded to TMT), and the NIH S10 OD011981, which supported the Life Sciences North Imaging Facility at the University of Arizona. Additional funding was provided by the University of Arizona through a BIO5 Postdoctoral Fellowship Award granted to NKK. We extend our gratitude to the team at the Pacific Northwest Center for Cryo-EM (PNCC), with special thanks to Nancy Meyer and Claudia Lopez, for their invaluable assistance with data collection.

References

1. K. P. Locher, Mechanistic diversity in ATP-binding cassette (ABC) transporters. *Nat Struct Mol Biol* **23**, 487–493, (2016).

2. M. J. Fath, R. Kolter, ABC transporters: bacterial exporters. *Microbiological Reviews* **57**, 995–1017, (1993).

3. P. M. Jones, A. M. George, The ABC transporter structure and mechanism: perspectives on recent research. *CMLS, Cell. Mol. Life Sci.* **61**, 682–699, (2004).

4. M. Dean, Y. Hamon, G. Chimini, The human ATP-binding cassette (ABC) transporter superfamily. *Journal of Lipid Research* **42**, 1007–1017, (2001).

5. J. ter Beek, A. Guskov, D. J. Slotboom, Structural diversity of ABC transporters. *J Gen Physiol* **143**, 419–435, (2014).

6. Z.-S. Chen, A. K. Tiwari, Multidrug resistance proteins (MRPs/ABCCs) in cancer chemotherapy and genetic diseases. *The FEBS Journal* **278**, 3226–3245, (2011).

7. Z. Ye, Y. Lu, T. Wu, The impact of ATP-binding cassette transporters on metabolic diseases. *Nutr Metab (Lond)* **17**, 61, (2020).

8. L. I. U. Xiang, L. I. U. Xiao-dong, Diabetes and ATP-binding cassette transporters. *Chinese Journal of Clinical Pharmacology and Therapeutics* **14**, 1428, (2020).

9. C.-H. Choi, ABC transporters as multidrug resistance mechanisms and the development of chemosensitizers for their reversal. *Cancer Cell International* **5**, 30, (2005).

10. K. Diederichs, *et al.*, Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon Thermococcus litoralis. *EMBO J* **19**, 5951–5961, (2000).

11. E. I. Stolarczyk, C. J. Reiling, C. M. Paumi, Regulation of ABC transporter function via phosphorylation by protein kinases. *Curr Pharm Biotechnol* **12**, 621–635, (2011).

12. J. A. Wemmie, W. S. Moye-Rowley, Mutational analysis of the Saccharomyces cerevisiae ATP-binding cassette transporter protein Ycf1p. *Mol Microbiol* **25**, 683–694, (1997).

13. K. A. Pickin, *et al.*, Suppression of Ycf1p function by Cka1p-dependent phosphorylation is attenuated in response to salt stress. *FEMS Yeast Res* **10**, 839–857, (2010).

14. D. P. Rich, *et al.*, Effect of deleting the R domain on CFTR-generated chloride channels. *Science* **253**, 205–207, (1991).

15. M. C. Winter, M. J. Welsh, Stimulation of CFTR activity by its phosphorylated R domain. *Nature* **389**, 294–296, (1997).

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

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

18. S. Roosbeek, *et al.*, Phosphorylation by protein kinase CK2 modulates the activity of the ATP binding cassette A1 transporter. *J Biol Chem* **279**, 37779–37788, (2004).

19. D. S. Miller, Regulation of P-glycoprotein and other ABC drug transporters at the blood-brain barrier. *Trends Pharmacol Sci* **31**, 246–254, (2010).

20. S. H. Cheng, *et al.*, Phosphorylation of the R domain by cAMP-dependent protein kinase regulates the CFTR chloride channel. *Cell* **66**, 1027–1036, (1991).

21. P. Eraso, M. Martínez-Burgos, J. M. Falcón-Pérez, F. Portillo, M. J. Mazón, Ycf1-dependent cadmium detoxification by yeast requires phosphorylation of residues Ser908 and Thr911. *FEBS Letters* **577**, 322–326, (2004).

22. G. Seavilleklein, *et al.*, PKC phosphorylation modulates PKA-dependent binding of the R domain to other domains of CFTR. *Am J Physiol Cell Physiol* **295**, C1366-1375, (2008).

23. D. R. Poroca, N. Amer, A. Li, J. W. Hanrahan, V. M. Chappe, Changes in the R-region interactions depend on phosphorylation and contribute to PKA and PKC regulation of the cystic fibrosis transmembrane conductance regulator chloride channel. *FASEB Bioadv* **2**, 33–48, (2020).

24. C. Basso, P. Vergani, A. C. Nairn, D. C. Gadsby, 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, (2003).

25. M. S. Szczypka, J. A. Wemmie, W. S. Moye-Rowley, D. J. Thiele, A yeast metal resistance protein similar to human cystic fibrosis transmembrane conductance regulator (CFTR) and multidrug resistance-associated protein. *J Biol Chem* **269**, 22853–22857, (1994).

26. L. S. Ostedgaard, O. Baldursson, D. W. Vermeer, M. J. Welsh, A. D. Robertson, A functional R domain from cystic fibrosis transmembrane conductance regulator is predominantly unstructured in solution. *Proceedings of the National Academy of Sciences* **97**, 5657–5662, (2000).

27. C. M. Lewarchik, K. W. Peters, J. Qi, R. A. Frizzell, Regulation of CFTR trafficking by its R domain. *J Biol Chem* **283**, 28401–28412, (2008).

28. Z. Zhang, F. Liu, J. Chen, Conformational Changes of CFTR upon Phosphorylation and ATP Binding. *Cell* **170**, 483-491.e8, (2017).

29. N. K. Khandelwal, *et al.*, The structural basis for regulation of the glutathione transporter Ycf1 by regulatory domain phosphorylation. *Nat Commun* **13**, 1278, (2022).

30. Y.-X. Mao, *et al.*, Transport mechanism of human bilirubin transporter ABCC2 tuned by the inter-module regulatory domain. *Nat Commun* **15**, 1061, (2024).

31. Z. Zhang, J. Chen, Atomic Structure of the Cystic Fibrosis Transmembrane Conductance Regulator. *Cell* **167**, 1586-1597.e9, (2016).

32. N. K. Khandelwal, T. M. Tomasiak, Structural basis for autoinhibition by the dephosphorylated regulatory domain of Ycf1. *Nat Commun* **15**, 2389, (2024).

33. C. J. Oldfield, *et al.*, Coupled folding and binding with alpha-helix-forming molecular recognition elements. *Biochemistry* **44**, 12454–12470, (2005).

34. M. Fuxreiter, I. Simon, P. Friedrich, P. Tompa, Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J Mol Biol* **338**, 1015–1026, (2004).

35. P. Radivojac, *et al.*, Intrinsic disorder and functional proteomics. *Biophys J* **92**, 1439–1456, (2007).

36. T. Hegedűs, A. W. R. Serohijos, N. V. Dokholyan, L. He, J. R. Riordan, Computational Studies Reveal Phosphorylation-dependent Changes in the Unstructured R Domain of CFTR. *Journal of Molecular Biology* **378**, 1052–1063, (2008).

37. N. Blom, S. Gammeltoft, S. Brunak, Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* **294**, 1351–1362, (1999).

38. N. Blom, T. Sicheritz-Pontén, R. Gupta, S. Gammeltoft, S. Brunak, Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* **4**, 1633–1649, (2004).

39. D. Scholl, *et al.*, A topological switch in CFTR modulates channel activity and sensitivity to unfolding. *Nat Chem Biol* **17**, 989–997, (2021).

40. S. C. Bickers, S. Benlekbir, J. L. Rubinstein, V. Kanelis, Structure of Ycf1p reveals the transmembrane domain TMD0 and the regulatory region of ABCC transporters. *Proc Natl Acad Sci U S A* **118**, e2025853118, (2021).

41. B. Mészáros, P. Tompa, I. Simon, Z. Dosztányi, Molecular Principles of the Interactions of Disordered Proteins. *Journal of Molecular Biology* **372**, 549–561, (2007).

42. V. J. Hilser, E. B. Thompson, Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc Natl Acad Sci U S A* **104**, 8311–8315, (2007).

43. T. A. Hopf, *et al.*, The EVcouplings Python framework for coevolutionary sequence analysis. *Bioinformatics* **35**, 1582–1584, (2019).

44. J. M. R. Baker, *et al.*, CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. *Nat Struct Mol Biol* **14**, 738–745, (2007).

45. B. Aryal, C. Laurent, M. Geisler, Learning from each other: ABC transporter regulation by protein phosphorylation in plant and mammalian systems. *Biochem Soc Trans* **43**, 966–974, (2015).

46. L. Csanády, *et al.*, Preferential Phosphorylation of R-domain Serine 768 Dampens Activation of CFTR Channels by PKA. *Journal of General Physiology* **125**, 171–186, (2005).

47. D. C. Gadsby, A. C. Nairn, Control of CFTR channel gating by phosphorylation and nucleotide hydrolysis. *Physiol Rev* **79**, S77–S107, (1999).

48. V. Chappe, *et al.*, Stimulatory and inhibitory protein kinase C consensus sequences regulate the cystic fibrosis transmembrane conductance regulator. *Proc Natl Acad Sci U S A* **101**, 390–395, (2004).

49. P. E. Wright, H. J. Dyson, Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* **16**, 18–29, (2015).

50. N. Ishiyama, *et al.*, Dynamic and Static Interactions between p120 Catenin and E-Cadherin Regulate the Stability of Cell-Cell Adhesion. *Cell* **141**, 117–128, (2010).

51. H. N. Motlagh, J. O. Wrabl, J. Li, V. J. Hilser, The ensemble nature of allostery. *Nature* **508**, 331–339, (2014).

52. A. Garcia-Pino, *et al.*, Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. *Cell* **142**, 101–111, (2010).

53. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290–296, (2017).

54. A. Punjani, H. Zhang, D. J. Fleet, Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat Methods* **17**, 1214–1221, (2020).

55. E. F. Pettersen, *et al.*, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **30**, 70–82, (2021).

56. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132, (2004).

57. T. I. Croll, ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, 519–530, (2018).

58. B. Sarkadi, E. M. Price, R. C. Boucher, U. A. Germann, G. A. Scarborough, Expression of the human multidrug resistance cDNA in insect cells generates a high activity drug-stimulated membrane ATPase. *J Biol Chem* **267**, 4854–4858, (1992).

59. N. L. Andon, *et al.*, Proteomic characterization of wheat amyloplasts using identification of proteins by tandem mass spectrometry. *PROTEOMICS* **2**, 1156–1168, (2002).

60. A. A. Klammer, C. Y. Park, W. S. Noble, Statistical Calibration of the SEQUEST XCorr Function. *J. Proteome Res.* **8**, 2106–2113, (2009).

61. A. Keller, A. I. Nesvizhskii, E. Kolker, R. Aebersold, Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* **74**, 5383–5392, (2002).

62. J. M. Falcón-Pérez, M. J. Mazón, J. Molano, P. Eraso, Functional Domain Analysis of the Yeast ABC Transporter Ycf1p by Site-directed Mutagenesis*. *Journal of Biological Chemistry* **274**, 23584–23590, (1999).

Figures and Tables

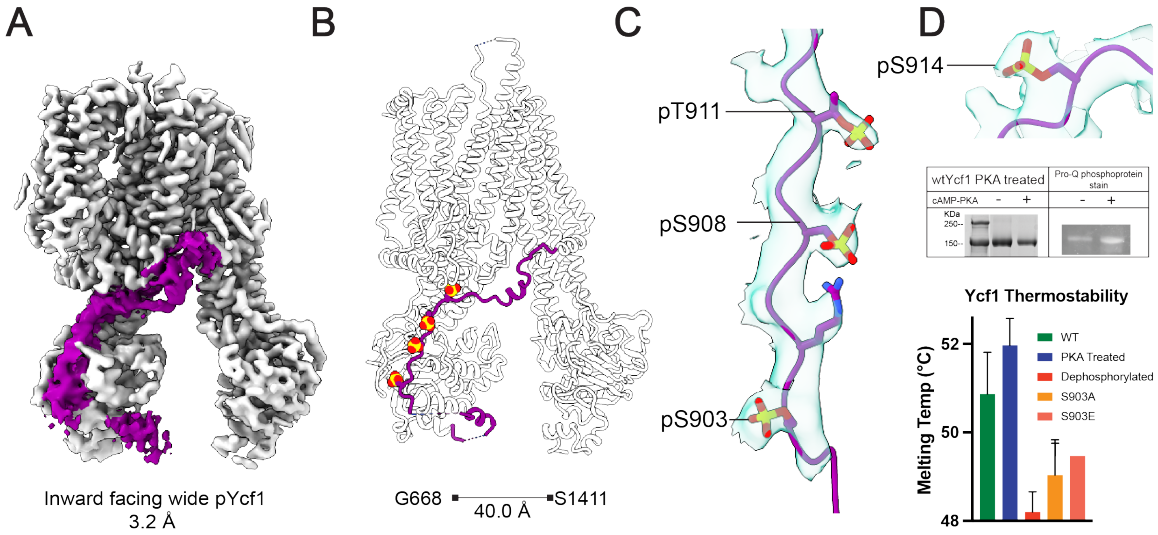

Figure 1. Structure determination of the tetra-phosphorylated Ycf1 by cryo-EM. (A) The refined cryo-EM map of phosphorylated Ycf1 at 3.2 Å resolution. Ycf1 is presented as grey with the R-domain density shown in dark purple. (B) Model of phosphorylated Ycf1. The R-domain is shown in dark purple, with 4 phosphorylation sites visible in the overall R-domain placement. (C) A close-up of the cryo-EM density maps (0.08 contour level, 2.5 selected radius) shows prominent phosphate groups associated with the residues S903, S908, T911 and S914. The phospho-stained SDS-PAGE gel confirms the higher phosphorylation levels achieved with the PKA-treated Ycf1 sample compared to the wild-type protein.

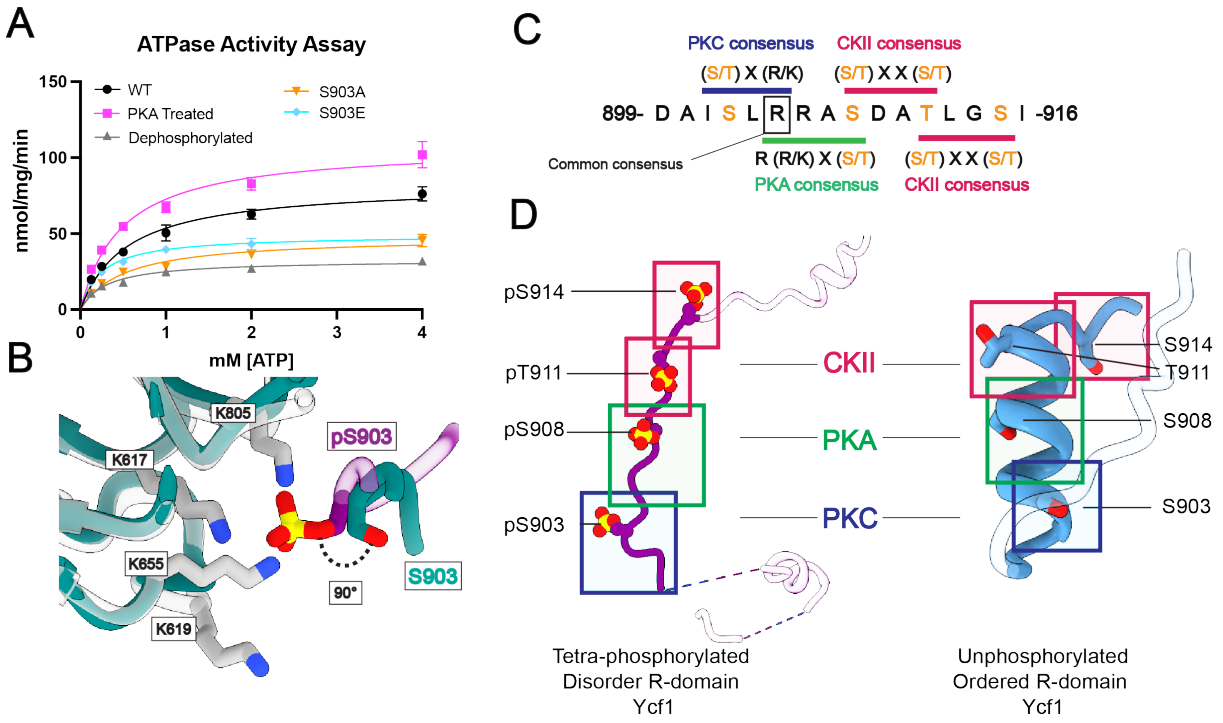

Figure 2. Effects of phosphorylation in Ycf1 R-domain and overlapping kinase consensus motifs. (A) The different phospho-states of Ycf1 and the single site mutant of S903 show relevant changes in the ATPase activity assay. (B) Comparison of the S903 site presented here (R-domain -purple) with the previous phosphorylated Ycf1 model missing the phosphorylation on S903 (colored green, PDBID: 7M69, (1)), the serine residue is shifted at a 90° angle towards its binding pocket due to the presence of the phosphate. (C) The arrangement of phosphorylation sites in kinase motifs within the R-domain presenting a unique pattern. Overlapping kinase motifs suggest a complex phospho-regulation mechanism. (D) A comparison of the phosphorylated R-domain conformation to the unphosphorylated counterpart (2).

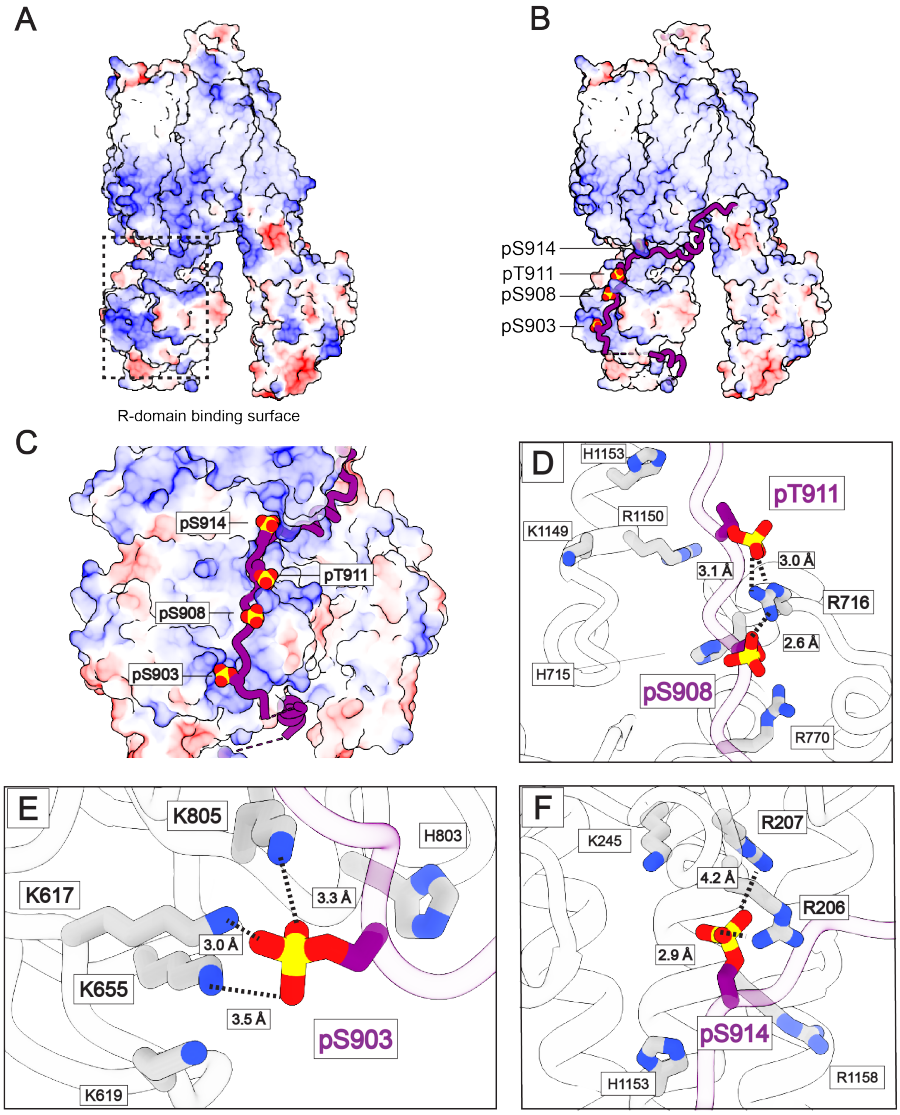

Figure 3. R-domain engagement with NBD1. (A) The electrostatic potential surface of Ycf1 with the R-domain density removed to illustrate the strong basic potential of the NBD1 surface (dashed box). (B) The tetra-phosphorylated R-domain closely represented in cartoon form on the rest of the Ycf1 surface colored as in (A) (C) The enclosed view of the interacting basic surface zone (D) The S908 and T911 phosphate groups are highlighted with their direct ionic interaction pair, the R716 in NBD1 with similar contacts as in (1). Also highlighted are other basic residues in the proximal regions adding to the overall charge in this micro-environment, as R770, R1149 and R1150. (E) The phosphorylated S903 site with the respective contacts to K617, K619, K655, R803 and H805. (F) S914 residue is encircled by the positively charged residues R206 as shown in (ref). In all subfigures, the R-domain is colored purple, the rest of Ycf1 colored grey, nitrogens are blue, oxygens are red, and phosphorous is colored yellow.

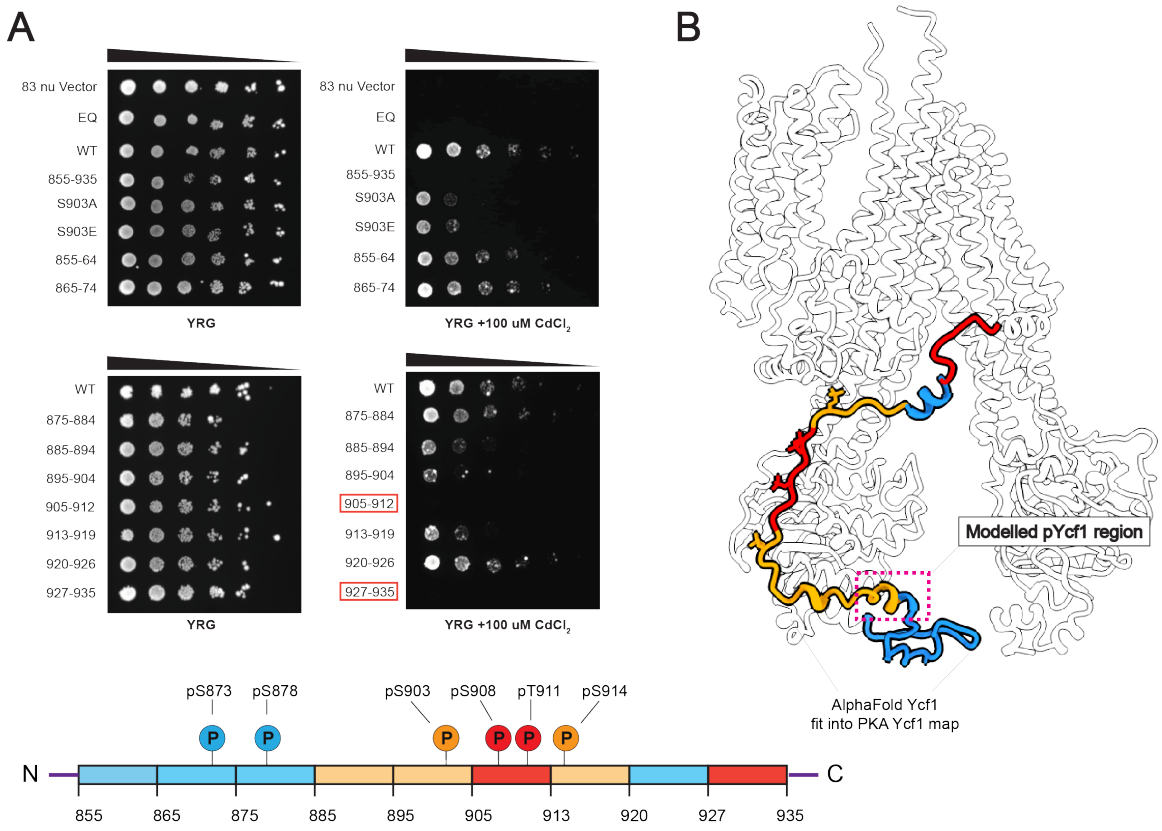

Figure 4. Yeast survival assay with scrambled R-domain constructs. (A) *S. cerevisiae* strain BY4742 with endogenous Ycf1 knocked out expressing different Ycf1 scrambled constructs for the highlighted R-domain segments. Cellular growth is serial diluted (1:5) and plated in a control YRG and CdCl2 toxic YRG plates (B). A representation of B on a linear sequence of Ycf1 with respective phosphorylation sites (C). The R-domain scrambled mutants colored by impact to survival. Red regions showed the most severe defects, blue showed no defect, and yellow are intermediate. Data shown are representative of n=6 (six technical replicates).

Table 1. Cryo-EM data collection and refinement statistics

Data Collection	
Microscope	ThermoFisher Titan Krios
Detector	Gatan K3
Image pixel size	0.822 Å
Defocus range	-0.9 to -2.1 µm
Electron exposure	~ 55 electron/Å
Number of frames	64
Number of micrographs	8,587
Image Processing	
	PKA IF-wide
No. of particles in final reconstruction	68,169
Symmetry	C1
Final box size (pixels)	440
Global resolution (CRYOSPARC map)	3.23 Å
FSC threshold	0.143
Refinement	
Atoms	23,753
Residues	1,481
Water	0
Supplied Resolution	3.21 Å
B-factor ($Å^2$)	
Iso/Aniso (#)	11777/0
Protein (min/max/mean)	45.60/223.33/117.60
Bonds (RMSD)	
Length (Å)	0.005
Angles (°)	0.881
Validation	
MolProbity score	0.66
Clash score	0.46
Ramachandran plot (%)	
Outliners	0
Allowed	1.78
Favored	98.22
Rotamer outliners (%)	0
Model vs Data	
CC (mask)	0.84
CC (box)	0.65
CC (peaks)	0.50
CC (volume)	0.83
PDB ID	9YAC
EMDB ID	XXXXX

Table 2. Mass spectrometry spectrum counts for PKA phosphorylated Ycf1

P39109 – Metal resistance protein YCF1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=YCF1 PE=1 SV=2					
Site	Modification	Best A-score	Localization Probability	Sample 1 100 ng_60 min	Sample 2 100 ng_60 min
S251	Phosphorylation	58.42	100 %	7	2
S846	Phosphorylation	27.76	100 %	1	0
S869	Phosphorylation	30.97	100 %	4	1
S873	Phosphorylation	26.20	100 %	1	0
S878	Phosphorylation	26.20	100 %	8	4
S903	**Phosphorylation**	**104.74**	**100 %**	**56**	**52**
S908	Phosphorylation	1,000.00	100 %	13	56
T911	Phosphorylation	1,000.00	100 %	10	48
S914	Phosphorylation	1,000.00	100 %	10	52
S1313	Phosphorylation	26.20	100 %	0	1
T1437	Phosphorylation	236.24	100 %	1	1