Identification of *in vivo* CFTR conformations during biogenesis and upon misfolding by covalent protein painting (CPP)

Sandra Pankow[1,2], Casimir Bamberger[1], Salvador Martínez-Bartolomé[1], Sung-Kyu Park[1], John R. Yates III[1,2]

Address:

[1] Department of Chemical Physiology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

[2] Corresponding authors: jyates@scripps.edu, pankows@scripps.edu

Abstract

In vivo characterization of protein structures or protein structural changes after perturbation is a major challenge. Therefore, experiments to characterize protein structures are typically performed *in vitro* and with highly purified proteins or protein complexes. Using a novel low-resolution method named Covalent Protein Painting (CPP) that allows the characterization of protein conformations *in vivo*, we are the first to report how an ion channel, the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR), is conformationally changed during biogenesis and channel opening in the cell. Our study led to the identification of a novel opening mechanism for CFTR by revealing that the interaction of the intracellular loop 2 (ICL2) with the nucleotide binding domain 2 (NDB2) of CFTR is needed for channel gating, and this interaction occurs concomitantly with changes to the narrow part of the pore and the walker A lysine in NBD1. However, the ICL2:NBD2 interface, which forms a "ball-in-a-socket" motif, is uncoupled during biogenesis, likely to prevent inadvertent channel activation during transport. In particular, solvent accessibility of lysine 273 (K273) in ICL2 changes with the opening and closing of the channel. Mutation of K273 severely impaired CFTR biogenesis and led to accumulation of CFTR in the Golgi and TGN. CPP further revealed that, even upon treatment with current approved drugs or at permissive temperature, the uncoupled state of ICL2 is a prominent feature of the misfolded CFTR mutants ΔF508 and N1303K that cause Cystic Fibrosis (CF), which suggests that stabilization of this interface could produce a more efficient CF drug. CPP is able to characterize a protein in its native environment and measure the effect of complex PTMs and protein interactions on protein structure, making it broadly applicable and invaluable for the development of new therapies.

Keywords

Chemical footprinting, high resolution mass spectrometry, misfolding, conformational defects

Introduction

The tertiary structure of a protein is closely linked to its function and determines how it interacts with other proteins. Mutations generating conformational defects are frequently associated with loss- or gain-of-function of a protein, and result in a variety of protein misfolding diseases such as CF or Alzheimer's [1-3]. Additionally, many of the proteins that cause misfolding diseases contain either long stretches of inherently disordered regions, are membrane proteins, or both, presenting a significant challenge to methods that depend on the determination of high-resolution structures to understand the conformational defects.

Recent cryo-EM studies of the human Cystic Fibrosis Transmembrane Regulator (CFTR), an anion channel important for maintaining the salt balance in stratified epitheliums such as the lung and intestine, have provided insight into channel architecture and structural rearrangements occurring during gating. However, the structures of several CFTR mutants that cause Cystic Fibrosis, the most common genetic childhood disease in Caucasians, have still not been solved, mainly due to technical problems to express and purify enough of these unstable proteins. In fact, even the wt CFTR structure could only be obtained by introducing stabilizing mutations and deleting a small disordered region, the RI element, which is physiologically important for CFTR maturation and is highly phosphorylated *in vivo* [4]. To overcome these constraints, we have developed covalent protein painting (CPP), a novel structural proteomics approach that allows identification and quantification of protein misfolding events *in vivo*.

CPP covalently labels the epsilon amine of a lysine side chain *in vivo*. Lysine side chains contain primary amines that are positively charged at physiological pH. They occur primarily on exposed surfaces of the native tertiary structure of a protein where they are readily accessible for interaction with other proteins as well as for labeling with amine-reactive reagents [5]. If lysines are involved in a protein interaction, they may be masked for labeling as they become part of the protein binding interface and therefore become solvent-excluded [6]. Thus, labeling reagents can be used to map protein structure and interactions by measuring the differences in the availability of the amino acid side chain to react with a label. These techniques have been successfully employed in top down approaches [5] for peptide mass mapping to identify modification sites and provide site-specific structural information on purified proteins or simple purified protein complexes.

Results

To characterize CFTR conformations *in vivo*, surface exposed lysines in human bronchial epithelial cells expressing wt CFTR (HBE16o-) were covalently labeled with isotope encoded dimethyl groups directly in-dish for 10 min on ice (CPP) [7] (Fig 1A). Subsequently, cells were lysed and CFTR was co-immunoprecipitated using the CoPIT protocol [8]. Labeled CFTR and its interactors were then digested into peptides with lysine-insensitive chymotrypsin. Subsequently, we used a second isotope combination to label the lysines that were solvent excluded in the

cell but accessible after digestion of the proteins into peptides. Using quantitative mass spectrometry, labeled lysines present in a complex cell lysate can be directly identified and quantified and their solvent accessibility determined. Twenty-six lysine sites in CFTR were quantified in 3 different biological replicates and mapped onto the CFTR Cryo-EM structure (PDB 5UAK) [9] showing a good general agreement between solvent accessibility determined by Cryo-EM and solvent accessibility determined by CPP (Fig. 1B,1C; Table 1). For example, lysines located on the outside of the CFTR molecule such as Lys 283 were all more than 98 % solvent accessible. In addition, CPP was able to quantify the solvent accessibility of lysines in intrinsically unstructured regions such as the Regulatory Region (R-region), which is not resolved in Cryo-EM structures and the Regulatory Insertion (RI) element, which has to be deleted in recombinant CFTR used for Cryo-EM structural studies for reasons of stabilization and yield [9].
We then performed CPP on isogenic CFBE41o- cells expressing ΔF508 CFTR and compared its solvent accessibilities with wt CFTR. Our experiment revealed large differences in solvent accessibility in the transmembrane helix 1 (TMH1) and the intracellular loop 2 (ICL2) as well as smaller, but significant differences in the RI element, the N-terminus of the nucleotide binding domain 1 (NBD1) and the NBD2 (Fig.1B, left panel). A label swap experiment confirmed the detected differences between wt and ΔF508 CFTR (Fig. 1B, right panel). Particularly interesting changes were observed for Lys 95 (K95) and Lys 464 (K464). Lys 95 (K95) is part of the narrow region of the internal vestibule and is fully accessible in wt CFTR and K464 is the Walker A lysine in NBD1 and involved in ATP binding; both were less accessible in ΔF508 CFTR. However, the largest difference was observed for K273 in the ICL2, which was solvent-excluded in 40% of wt CFTR molecules, but solvent accessible to greater than 95% in ΔF508 CFTR (Fig. 1D). The ICL2 is a short “coupling helix” that reaches into the NBD2 in what has been called a “ball-in-a-socket” motif [10]. K273 is located at the bottom of the loop and is solvent excluded in all published Cryo-EM structures as well as in 40 % of wt CFTR by CPP. Solvent accessibility of K273 thus represents a CFTR conformation in which ICL2 does not reach into NBD2, e.g. an uncoupled state. Interestingly, 60 % of wt CFTR is also in such an uncoupled conformation.

To gain a better understanding of the observed differences between wt and ΔF508 CFTR, we performed CPP experiments on ΔF508 CFTR after treatment with 5 μM of the FDA approved VX-809 CF drug[11]. We also performed CPP experiments on ΔF508 CFTR at permissive temperature of 28˚C, a temperature at which ΔF508 CFTR is efficiently trafficked to the plasma membrane and partially active [12] (Figure 2). Treatment with VX-809 did not significantly change solvent accessibility of lysines in ΔF508 CFTR except for lysine 1218, which is located in NBD2 and is ubiquitinated in ΔF508 CFTR [4,13]. However, permissive temperature of 28˚C reduced solvent accessibility of Lys273 in ICL2 to 89% (e.g. by 6 %), showing that coupling of ICL2 to NBD2 now occurred in about 11% of CFTR molecules. We therefore speculated that ICL2 coupling to NBD2 may be crucial for CFTR activity. Comparing solvent accessibility of K273 in Cryo-EM models of inactive CFTR [9]

(PDB 5UAK) and activated, dimerized CFTR [14] (PDB 5UAR) revealed that K273 was solvent excluded in both models as well as in additional homology models [15].

To investigate the possibility that *in vivo* activation may be different from *in vitro* activation of CFTR, we treated wt CFTR with 500 nM VX-770 for 10 min[16] and then measured K273 accessibility by CPP. VX-770 is a channel potentiator that opens CFTR channels independent of ATP hydrolysis and is used to treat CF caused by the G551D mutation [17-19]. Surprisingly, activation of wt CFTR by VX-770 treatment led to complete solvent exclusion of K273. Only in 0.66 % of CFTR molecules was K273 still solvent accessible, thus showing a clear correlation of K273 accessibility with wt CFTR activity (Fig. 2B). Coupling of ICL2 to NBD2 upon activation with VX-770 was accompanied by additional surface accessibility changes in the vestibule (K95), the chloride pathway (Lys 370) and the Walker A lysine Lys 464, all of which became less accessible (Fig. 2C, Supplemental Figure 1A). The data suggested that the CFTR conformation in which K273 is solvent excluded and the ICL2 reaches into NBD2 is an active open state, while the conformation in which K273 is solvent accessible represents an inactive state in which ICL2 is uncoupled from NBD2.

To probe whether this previously undescribed inactive CFTR conformation is associated with gating of the channel or is rather an inactive confirmation attained by CFTR when it is not present at the plasma membrane, we performed additional CPP experiments with CFTR variants N1303K CFTR and well as G551D CFTR. N1303K CFTR, like ΔF508 CFTR, is misfolded and not present at the plasma membrane, and G551D CFTR is trafficked normally to the plasma membrane, but has a gating defect [20-23]. CPP results on these CFTR variants showed that solvent availability of K273 was high (>95%) in the mainly ER and Golgi located, misfolded CFTR variants N1303K and ΔF508 CFTR, but that it was solvent excluded in G551D, suggesting that this conformation is not associated with CFTR gating per se (Figure 3A, 3B). To investigate when this CFTR conformation occurs, we mutated CFTR K273 to K273A (which is a small, polar, cavity creating amino acid) and analyzed biogenesis of K273A CFTR by confocal microscopy. Confocal images showed that the K273A CFTR mutant exits the ER, but does not reach the plasma membrane; instead it accumulates in the cells (Figure 4A). However, these are not aggregates, as shown by missing colocalization with autophagy markers such as SQSTM and cleaved LC-3 (Supplemental Figure 2). Interactome analysis of K273A CFTR versus a GFP control followed by GO enrichment analysis revealed that proteins localized to membrane bounded organelles and vesicles were highly enriched as interactors of K273A CFTR as well as proteins involved in endosomal transport and vesicle mediated transport (Fig. 4B, 4C, Supplemental table 1). Examples of these interactors include the ESCRT complex, and the retromer cargo-selective complex (CSC), specifically the SNX3-retromer. The ESCRT complex is required for endosomal sorting of protein cargo into multivesicular bodies (MVBs)[24,25], and the SNX3-retromer mediates retrograde transport from the endosome to the trans-Golgi-network (TGN), and transport from endosome to plasma membrane; it was also previously shown to promote cell

surface expression of ENAC sodium channels [26-28] (Fig. 4D). However, interactors that recognize misfolded proteins in the ER and belong to the ER associated degradation (ERAD) pathway and are over-represented in ΔF508 CFTR interactomes, were not enriched. CPP analysis of the K273A CFTR mutant further showed that solvent accessibility of lysines was similar to wt CFTR in non-activated state (Figure 5D).

Discussion

Here, we report a novel CFTR conformation that reflects an immature, closed state of the channel. This conformation is attained by 60 % of wt CFTR molecules in the cell as well as the misfolding mutants ΔF508 CFTR and N1303K[JY1][MOU2]. The data also suggest previously unknown conformational rearrangements that appear upon activation of the channel in the ion permeation pathway and which add important detail to the knowledge gained from Cryo-EM structures.

The most striking feature of the newly discovered immature CFTR conformation is the altered solvent accessibility of K273 in the ICL2 coupling helix. In general, ABC transporter ICLs mediate the coupling of ATPase (catalytic) activity at the nucleotide binding sites to substrate translocation through the membrane spanning domains (MSDs) [29]. Previous data suggested that ICL2 is involved in stabilization of the full conductance state of the CFTR Cl channel, as a 19 amino acid deletion mutant of CFTR ICL2 is non-responsive to forskolin stimulation and does not reach the cell surface [30]. In models of the ABC transporter P-glycoprotein the ICLs may move like a ball in a socket to accommodate conformational change during the gating cycle [10]. In this arrangement, the ICL2 loop reaches into the NBD2 as in all current Cryo-EM structures of CFTR and thus is not solvent accessible. Such solvent exclusion of K273 was observed upon activation of the CFTR channel with VX-770. Interestingly, crosslinking cysteines engineered in CFTR's ICL2 and NBD2 decreased the open probability of the channel (P_o) dramatically [31,32], suggesting that movement of the ICL2 is needed for proper channel gating. However, if K273 is fully solvent accessible as it is in ΔF508, N1303K and 60% of wt and G551D CFTR, it must represent a conformation in which the ICL2 is removed from the socket in NBD2 (i.e., an "uncoupled" state), and thus represents an inactive CFTR conformation.

One could interpret the different conformations simply as open/closed states during the gating cycle. However, the fact that this inactive conformation is also observed in ΔF508 and N1303K CFTR, two CFTR variants that are misfolded and are largely degraded before reaching the plasma membrane, suggests otherwise. Furthermore, based on its interactome and confocal images, K273A mutation led to accumulation of CFTR in the Trans-Golgi network (TGN), ERGIC and endosomes. This indicates that the inactive conformation is an immature CFTR conformation that is attained during transport from the ER to the plasma membrane and may prevent accidental channel activation during biogenesis and transport.

Thus, CPP was used to discover an immature CFTR conformation that was previously unobservable. It is possible that this conformation is stabilized by

interactors that prevent ICL2 from reaching into NBD2 or by internal molecular re-arrangements, such as movement of the R-domain or the RI element upon phosphorylation. While movement of the R-domain from its place between the two NBDs to the peripheral surface has been observed in the PKA phosphorylated, ATP-bound CFTR [14], it cannot be the only factor that prevents the coupling of ICL2 to NBD2, as ICL2 also reaches into NBD2 in the structure of the un-phosphorylated, ATP-free CFTR. We therefore suspect that additional *in vivo* phosphorylation events are needed, or that movement of the RI element (which is deleted in all CFTR structures) is involved since it is highly phosphorylated in wt CFTR, but not in ΔF508 CFTR and the amount of RI element phosphorylation correlates with the amount of fully mature CFTR in the plasma membrane [4].

The CPP results also present an opportunity to further examine changes in the ion permeation pathway that occur upon channel activation. Changes include the walker A lysine K464 becoming less solvent accessible upon VX-770 activation, likely reflecting ATP-binding and dimerization of the NBDs. This is consistent with the observation that NBD binding symmetrically closes off the ATP-binding sites in the activated state [14]. Further changes include K95 and K370, two charged residues in the vestibule and close to the chloride entry pore [9], which become partially inaccessible in activated wt CFTR treated with VX-770 and the G551D mutant. This could reflect a percentage of channels that have closed, or may reflect conformational changes that reflect regional pore constriction/tightening at the mouth of the pore and the vestibule. Regional pore constriction/tightening is consistent with a role of these positively charged residues in anion selection and in creating an electrostatic potential for chloride ions to move along the pore [33]. Comparing CPP results across all analyzed CFTR mutants and upon activation, we also noticed changes in solvent accessibility of K52, which is part of the “lasso motif” and a structure unique to CFTR [34]. K52 was accessible in ΔF508 and N1303K, but only partially accessible in G551D and in VX-770 activated wt CFTR, suggesting a role of the lasso in channel activation or conformational changes occurring upon localization to the plasma membrane. Such a role would be consistent with previous observations that the lasso motif is important for interactions with the membrane traffic machinery and would confirm suggestions that the lasso regulates CFTR channel gating through interactions with the R domain [14,35-37].

Finally, the CPP data revealed a few changes that appear to be unique to N1303K, ΔF508 and G551D CFTR and thus related to the misfolding and the gating defect, respectively. In comparison to ΔF508 CFTR, N1303K CFTR exhibited higher solvent accessibility at K584 in NBD1 and K643 located at the border of the NBD1 and R-region. The G551D mutant exhibited a specific difference in solvent accessibility of K536. A change unique to ΔF508 CFTR was the slightly altered solvent accessibility of K1420 in NBD2. Interestingly, structural differences between wt and ΔF508 CFTR in NBD2 have been previously identified by limited trypsin digestion assays [38]. Finally, CPP data also allowed us to better rationalize a VX-809 corrector mechanism, which remains largely unknown. VX-809 has been

reported to bind to NBD1 using isolated NBD1 and to stabilize domain:domain interfaces, particularly NBD1:ICL4, although others observed that VX-809 alone do not stabilize isolated NBD1 [39-41]. In this study, we observed a reduction in solvent accessibility of K1218 by more than 30%, a very clear change that was not observed in any other condition (Supplemental table 2). Point mutation of K1218 to K1218R has been shown to increase the amount of CFTR, stabilize it at the plasma membrane and reduce lysosomal degradation of CFTR [13], while its deletion improves folding efficiency of ΔF508 CFTR [42]. Binding of VX-809 to the region or K1218 itself could thus diminish ΔF508 CFTR targeting to the lysosome, although it is unlikely to completely prevent it, as multiple mechanism exist for peripheral quality control of CFTR [43,44]. Such a mechanism would be in line with the moderate rescue of ΔF508 CFTR processing observed for VX-809.

Methods

Cell culture, plasmids and point mutations

CFBE and HBE cells were grown in A-MEM supplemented with 10% FBS, 1% Penicillin-Streptomycin (GibCo, Carlsbad, CA) and appropriate antibiotics at 37˚C, 5% CO_2 . FRT cells expressing either wt CFTR or the variants G551D, N1303K or ΔF508 CFTR were cultured in Ham's F12 medium, Coon's modification supplemented with 5 % FBS and 100 µg/ml Hygromycin. K273A CFTR was generated using the Quick change method (Quiagen) in the pCMV6-ac-CFTR-GFP plasmid (Origene). Hek293T cells were transfected with Lipofectamine 3 according to manufacturer's recommendation (Thermo Fisher, Carlsbad, CA) and expression of K273A CFTR was analyzed 48 h to 72 h after transfection.

In vivo dimethyl labeling and CFTR-IP

Cell culture media was removed and the intact cell layer was washed with 1x PBS before labeling of surface-exposed lysines with either formaldehyde (CH3, 0.3 % final concentration) and sodiumcyanoborohydrate (NaBH3CN, final concentration 30 mM) for light labeling or with deuterated formaldehyde containing ^{13}C ($^{13}CD3$) and deuterated sodiumcyanoborohydrate ($NaBD_3CN$) for heavy labeling. The labeling reaction was allowed to continue for 10 min on ice before it was stopped by adding 1% ammonium acetate ($NH_4C_2H_3O$). Subsequently, cells were harvested with a cell scraper and lysed by adding 2 x TNI buffer [8] containing complete Ultra protease inhibitor (Roche) and Halt phosphatase inhibitor (Pierce). The cell lysate was sonicated for 3 min in a water bath sonicator and insoluble material was removed by centrifugation at 18,000 x g, 4˚C, 15 min. CFTR-IP was then performed as described in Pankow et al., 2016 [8], except that the last three washes were performed with HNN buffer (50 mM Hepes, pH 7.5, 150 mM NaCl, 1 mM EDTA). Proteins were precipitated by methanol-chloroform precipitation (Lysate : Methanol : Chloroform, 1:4:1, *v*:*v*:*v*) and the precipitate was washed with 95% Methanol before digestion. For interactome analyses, CFTR was immunoprecipitated as described [45] using GFP-Trap Agarose (Chromotek).

Protein digestion and peptide labeling

Precipitated proteins were resolubilized by sonication in 0.2% Rapigest, 100 mM HEPES, pH 8.0 in a water bath sonicator. The protein lysate was heated to 95˚C for 10 min, reduced by 5 mM TCEP (Pierce), and alkylated with 10 mM iodoacetamide (Sigma-Aldrich, St.Louis, MO) before digestion with chymotrypsin (Protea Biosciences, Morgantown, WV, 1:100 w:w) at 30˚C for 14-16 hours. Peptides were labeled according to Boersema et al., 2009 [46] with 0.16% CH3 or ^{13}CD3 and 30 mM NaBH3CN or $NaBD_3CN$, respectively, for 1 h at room temperature. The labeling reaction was stopped with 1% $NH_4C_2H_3O$, Rapigest was inactivated by incubation with 9% formic acid (1h, 37˚C) and samples were reduced to near dryness *in vacuo*. Tryptic digestion of unlabeled CFTR IPs was carried out as described previously [8].

Mass spectrometry

Peptides were reconstituted in buffer A (95% H_2O, 5% acetonitrile/0.1% formic acid), loaded onto a preparative MudPIT column and analyzed by nano-ESI-LC/LC-MS/MS on an LTQ-Elite or LTQ-VelosPro Orbitrap (Thermo Fisher, San Jose, CA) by placing the triphasic MudPIT column in-line with an Agilent 1200 quaternary HPLC pump (Agilent, Palo Alto, CA) and separating the peptides in multiple dimensions with a 10-step gradient (0%, 10%, 20%, 30%, 40%, 50%, 60%, 70 %, 80%, 90% Buffer C ((500 mM ammonium acetate/5% acetonitrile/0.1% formic acid)) over 20 h as described previously [47]. To avoid cross contaminations between different samples, each sample was loaded onto a fresh column. Each full scan mass spectrum (400-2000 m/z) was acquired at 60,000 R, followed by 20 data-dependent MS/MS scans at 35% normalized collisional energy and an ion count threshold of 1000. Dynamic exclusion was used with an exclusion list of 500, repeat time of 60 s and asymmetric exclusion window of -0.51 and +1.5 Da. Monoisotopic peaks were extracted from Raw-files with RawConverter [48] and MS/MS spectra were searched with ProLuCID [49] against the human Uniprot database (release Nov 2019) using a target-decoy approach in which each protein sequence is reversed and concatenated to the normal database [50]. The search was carried out using a search window of 50 ppm, no enzyme specificity, N-terminal static modification according to protein dimethyl label (e.g. 28.0313 or 36.0757), and lysine dimethylation (28.0313) and carbamidomethylation of cysteine (57.02146) as static modifications. Additionally, oxidation of methionine (15.9949) and serine, threonine or tyrosine phosphorylation (79.9663) were considered as differential modifications. A lysine mass shift of 8.0444 was used for metabolic labeling search. Minimum peptide length was set to 7 amino acids (AA). Search results were filtered to less than 0.2 % FDR on spectrum level, corresponding to less than 0.5 % peptide level FDR using DTASelect 2.1 within the Integrated Proteomics Pipeline (IPA, San Diego, CA). Additionally, a deltaCN value of ≥ 0.1, and precursor delta mass cutoff of less than 8 ppm were applied as filter.

Quantification of solvent-excluded vs solvent-exposed lysines and mapping of solvent accessible areas

To determine the proportion of solvent-exclusion to solvent-exposure of a lysine residue, the area under each peptide peak was quantified using Census and the Integrated Proteomics Pipeline (IPA, San Diego, CA) as described previously [51,52]. Individual isotopes were extracted within a 10 ppm window. A determinant factor of ≥ 0.5, a profile score of ≥ 0.5 and no retention time shift allowed were set as parameters to filter for correct peak assignments. The determined ratio of light over heavy directly reflects solvent-exposure/ solvent-exclusion. Peak shapes were also manually assessed for large ratio measurements. Protein structures available in PDB were visualized with JMOL (version 14.30.1) and the isosurface (isoSurface saSurface) of the epsilon amino group of lysine calculated and visualized in JMOL with a rolling probe radius of 1.2 Å. Area ratios were expressed as percentage of surface accessibility as described [7] and plotted in Prism 7 (Graphpad Inc.) Statistical analysis of area ratios per site was carried out in PCQ as described [53]. Briefly, ratio values for each lysine site were calculated with the SoPaX algorithm, which is part of ProteinClusterQuant (PCQ, https://github.com/proteomicsyates/ProteinClusterQuant), a One-way ANOVA analysis carried out and obtained results subjected to Benjamini-Hochberg correction.

Immunofluorescence staining and western blotting

Cells were grown in Lab Tek II chambers (Thermo Fisher), fixed with 4 % paraformaldehyde/1xPBS for 10 min at RT and solubilized with 0.1% Triton X-100 /1xPBS before staining with antibodies against LC3A/B (Cell Signaling Technologies, cat no. 12741S) and SQSTM1 (Cell Signaling Technologies, cat no. 8025S). The plasma membrane was visualized by wheat germ agglutin CF 594 conjugate (Biotium) and the ER was visualized using the ER Cytopainter kit (Abcam) according to manufacturer's recommendations. Slides were mounted in ProLong Gold Antifade Mountant with DAPI to visualize nuclei (Thermo Fisher). Images were taken with a Zeiss LSM 710 confocal laser scanning microscope. Protein lysates were prepared in TNI buffer as described previously [8]. CFTR was detected using mouse monoclonal antibodies 24.1 (ATCC) and M3A7 (EMD Millipore). Antibodies against b-actin (AC-15, Sigma) and Na+/K+ATPase a (H300, sc28800, Santa Cruz) were used for normalization. Horseradish-peroxidase-conjugated secondary antibodies (Jackson Immunoresearch) were detected with enhanced chemiluminescence (ECL, Pierce).

Author contributions

S.P., C.B. and J.R.Y. designed the research and S.P. and C.B. performed the experiments. S.M.B. implemented the protein residue-specific quantification and SoPaX in PCQ. S.K.P. supported mass spectrometric data analysis and developed quantification software. J.R.Y. provided materials and funding. S.P. wrote the manuscript and prepared the figures with help from all authors. All authors read and approved the manuscript.

Acknowledgements
We would like to thank E. Sorscher (University of Alabama) for the kind gift of the FRT cell lines and Claire Delahunty (TSRI) for reading the manuscript and editing it for clarity. Funding was provided by NIH grants R01HL131697-05 and R33CA212973-01 awarded to John R. Yates III.

Figure legends

Figure 1. CPP analysis of wt and ΔF508 CFTR. A. Experimental and mass spectrometric workflow. **B.** Solvent accessibility of quantified lysines in different CFTR domains upon labeling with CH_3 ("light") followed by $^{13}CD_3$ ("heavy") labeling (left panel) and upon label-swap (right panel). **C**. Quantified lysines were mapped onto the CFTR Cryo-EM structure (PDB:5UAK) and solvent accessible isosurfaces displayed as blue spheres. **D**. Representative MS1 spectra of K273 depicting light and heavy precursors used for quantification in wt CFTR (upper panel) or ΔF508 CFTR (lower panel). Wt CFTR n=4; ΔF508 CFTR n=4 (biological replicates).

Figure 2. **Solvent accessibility of CFTR lysines in ICL-2 and the ion permeation pathway upon correction and activation. A.** Changes in solvent accessibility at permissive temperature of 28 ˚C (24h), upon VX-809 treatment (5 µM, 36h) or in DMSO control. **B**. Channel activation by VX-770 (500 nM, 10 min) changes solvent accessibility of K273 in ICL-2. Mean is indicated for each condition. **C**. Changes in solvent accessibility of lysines located in the ion permeation pathway and ATP binding site upon channel activation with VX-770. Wt (n=4), wt_VX770 (n=3), ΔF508 (n=4), ΔF508_VX809 (n=3), ΔF508_DMSO (n=3), ΔF508_28˚C (n=3). n represents number of biological replicates.

Figure 3. K273 solvent accessibility in G551D and N1303K CFTR. **A**. Solvent accessibility of K273. **B**. Representative MS1 spectrum of the peptide containing K273 in G551D CFTR. Light and heavy precursors and measured m/z for each are indicated. G551D (n=3), N1303K (n=3). n represents biological replicates.

Figure 4. **K273A CFTR interactome and cellular localization. A**. HEK293T cells transiently transfected with GFP tagged wt or K273A CFTR (green) were stained with wheat germ agglutinin (pink) to visualize the plasma membrane (upper panel). ER was stained with the ER Cytopainter kit (lower panel). Nuclei were counterstained with DAPI (blue).n=3 biological replicates. **B**. Interactome analysis of K273A CFTR. True K273A CFTR interactors were distinguished from background using CoPITgenerator. **C**. Network of enriched GO-terms in the K273A CFTR interactome. **D**. ESCRT and retromer complexes were specifically enriched in the K273A interactome.

Table 1. Mapping and comparison of lysine solvent accessibilities to Cryo-EM structures. Quantified lysines and respective peptide sequences were mapped onto the available Cryo-EM structures 5UAK and 5UAB and solvent accessible isosurfaces probed with a rolling probe.

Supplemental Figure 1. A. Solvent accessibility of all quantified CFTR lysines upon activation with VX-770 (n=3). **B**. Solvent accessibility of all quantified N1303K CFTR lysines compared to ∆F508 and wt CFTR. N1303K (n=3).

Supplemental Figure 2. HEK293T cells transiently transfected with GFP-tagged wt or K273A CFTR were stained for autophagy markers SQSTM (pink) or cleaved LC-3 (pink).

Supplemental table 1. Enriched GO-terms in K273A interactome with p-values and adjusted p-values.

Supplemental table 2. Means and standard deviation (SD) of all quantified lysines, the identified peptide sequences and CFTR AA position are displayed for all measured conditions.

References

1 Winklhofer, K. F., Tatzelt, J. & Haass, C. The two faces of protein misfolding: gain- and loss-of-function in neurodegenerative diseases. *Embo J* **27**, 336-349, doi:10.1038/sj.emboj.7601930 7601930 [pii] (2008).

2 Cohen, F. E. & Kelly, J. W. Therapeutic approaches to protein-misfolding diseases. *Nature* **426**, 905-909, doi:10.1038/nature02265 (2003).

3 Gregersen, N., Bross, P., Vang, S. & Christensen, J. H. Protein misfolding and human disease. *Annu Rev Genomics Hum Genet* **7**, 103-124, doi:10.1146/annurev.genom.7.080505.115737 (2006).

4 Pankow, S., Bamberger, C. & Yates, J. R., 3rd. A posttranslational modification code for CFTR maturation is altered in cystic fibrosis. *Sci Signal* **12**, doi:10.1126/scisignal.aan7984 (2019).

5 Mendoza, V. L. & Vachet, R. W. Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass spectrometry reviews* **28**, 785-815, doi:10.1002/mas.20203 (2009).

6 Luchini, A., Espina, V. & Liotta, L. A. Protein painting reveals solvent-excluded drug targets hidden within native protein-protein interfaces. *Nature communications* **5**, 4413, doi:10.1038/ncomms5413 (2014).

7 Bamberger, C., Pankow S., *et al.*.Covalent Protein Painting Reveals Structural Changes in the Proteome in Alzheimer Disease doi:10.1101/2020.01.31.929117 (2020).

8 Pankow, S., Bamberger, C., Calzolari, D., Bamberger, A. & Yates, J. R., 3rd. Deep interactome profiling of membrane proteins by co-interacting protein identification technology. *Nat Protoc* **11**, 2515-2528, doi:10.1038/nprot.2016.140 (2016).

9 Liu, F., Zhang, Z., Csanady, L., Gadsby, D. C. & Chen, J. Molecular Structure of the Human CFTR Ion Channel. *Cell* **169**, 85-95 e88, doi:10.1016/j.cell.2017.02.024 (2017).

10 Oldham, M. L., Davidson, A. L. & Chen, J. Structural insights into ABC transporter mechanism. *Curr Opin Struct Biol* **18**, 726-733, doi:10.1016/j.sbi.2008.09.007 (2008).
11 Van Goor, F. *et al.* Correction of the F508del-CFTR protein processing defect in vitro by the investigational drug VX-809. *Proc Natl Acad Sci U S A* **108**, 18843-18848, doi:10.1073/pnas.1105787108 (2011).
12 Denning, G. M. *et al.* Processing of mutant cystic fibrosis transmembrane conductance regulator is temperature-sensitive. *Nature* **358**, 761-764, doi:10.1038/358761a0 (1992).
13 Lee, S. *et al.* Interference with ubiquitination in CFTR modifies stability of core glycosylated and cell surface pools. *Mol Cell Biol* **34**, 2554-2565, doi:10.1128/MCB.01042-13 (2014).
14 Zhang, Z., Liu, F. & Chen, J. Conformational Changes of CFTR upon Phosphorylation and ATP Binding. *Cell* **170**, 483-491 e488, doi:10.1016/j.cell.2017.06.041 (2017).
15 Kleizen, B. *et al.* CFTR: New insights into structure and function and implications for modulation by small molecules. *J Cyst Fibros* **19 Suppl 1**, S19-S24, doi:10.1016/j.jcf.2019.10.021 (2020).
16 Csanady, L. & Torocsik, B. Cystic fibrosis drug ivacaftor stimulates CFTR channels at picomolar concentrations. *Elife* **8**, doi:10.7554/eLife.46450 (2019).
17 Accurso, F. J. *et al.* Effect of VX-770 in persons with cystic fibrosis and the G551D-CFTR mutation. *N Engl J Med* **363**, 1991-2003, doi:10.1056/NEJMoa0909825 (2010).
18 Van Goor, F. *et al.* Rescue of CF airway epithelial cell function in vitro by a CFTR potentiator, VX-770. *Proc Natl Acad Sci U S A* **106**, 18825-18830, doi:10.1073/pnas.0904709106 (2009).
19 Eckford, P. D., Li, C., Ramjeesingh, M. & Bear, C. E. Cystic fibrosis transmembrane conductance regulator (CFTR) potentiator VX-770 (ivacaftor) opens the defective channel gate of mutant CFTR in a phosphorylation-dependent but ATP-independent manner. *J Biol Chem* **287**, 36639-36649, doi:10.1074/jbc.M112.393637 (2012).
20 Bompadre, S. G., Sohma, Y., Li, M. & Hwang, T. C. G551D and G1349D, two CF-associated mutations in the signature sequences of CFTR, exhibit distinct gating defects. *J Gen Physiol* **129**, 285-298, doi:10.1085/jgp.200609667 (2007).
21 Cutting, G. R. *et al.* A cluster of cystic fibrosis mutations in the first nucleotide-binding fold of the cystic fibrosis conductance regulator protein. *Nature* **346**, 366-369, doi:10.1038/346366a0 (1990).
22 Kerem, B. S. *et al.* Identification of mutations in regions corresponding to the two putative nucleotide (ATP)-binding folds of the cystic fibrosis gene. *Proc Natl Acad Sci U S A* **87**, 8447-8451 (1990).
23 Drumm, M. L. *et al.* Chloride conductance expressed by delta F508 and other mutant CFTRs in Xenopus oocytes. *Science* **254**, 1797-1799 (1991).

24 Henne, W. M., Buchkovich, N. J. & Emr, S. D. The ESCRT pathway. *Dev Cell* **21**, 77-91, doi:10.1016/j.devcel.2011.05.015 (2011).
25 Vietri, M., Radulovic, M. & Stenmark, H. The many functions of ESCRTs. *Nat Rev Mol Cell Biol* **21**, 25-42, doi:10.1038/s41580-019-0177-4 (2020).
26 Boulkroun, S. *et al.* Vasopressin-inducible ubiquitin-specific protease 10 increases ENaC cell surface expression by deubiquitylating and stabilizing sorting nexin 3. *Am J Physiol Renal Physiol* **295**, F889-900, doi:10.1152/ajprenal.00001.2008 (2008).
27 Harterink, M. *et al.* A SNX3-dependent retromer pathway mediates retrograde transport of the Wnt sorting receptor Wntless and is required for Wnt secretion. *Nat Cell Biol* **13**, 914-923, doi:10.1038/ncb2281 (2011).
28 Harrison, M. S. *et al.* A mechanism for retromer endosomal coat complex assembly with cargo. *Proc Natl Acad Sci U S A* **111**, 267-272, doi:10.1073/pnas.1316482111 (2014).
29 Locher, K. P. Review. Structure and mechanism of ATP-binding cassette transporters. *Philos Trans R Soc Lond B Biol Sci* **364**, 239-245, doi:10.1098/rstb.2008.0125 (2009).
30 Xie, J., Drumm, M. L., Ma, J. & Davis, P. B. Intracellular loop between transmembrane segments IV and V of cystic fibrosis transmembrane conductance regulator is involved in regulation of chloride channel conductance state. *J Biol Chem* **270**, 28084-28091, doi:10.1074/jbc.270.47.28084 (1995).
31 He, L. *et al.* Multiple membrane-cytoplasmic domain contacts in the cystic fibrosis transmembrane conductance regulator (CFTR) mediate regulation of channel gating. *J Biol Chem* **283**, 26383-26390, doi:10.1074/jbc.M803894200 (2008).
32 Sorum, B., Czege, D. & Csanady, L. Timing of CFTR pore opening and structure of its transition state. *Cell* **163**, 724-733, doi:10.1016/j.cell.2015.09.052 (2015).
33 Li, M. S., Cowley, E. A., El Hiani, Y. & Linsdell, P. Functional organization of cytoplasmic portals controlling access to the cystic fibrosis transmembrane conductance regulator (CFTR) chloride channel pore. *J Biol Chem* **293**, 5649-5658, doi:10.1074/jbc.RA117.001373 (2018).
34 Zhang, Z. & Chen, J. Atomic Structure of the Cystic Fibrosis Transmembrane Conductance Regulator. *Cell* **167**, 1586-1597 e1589, doi:10.1016/j.cell.2016.11.014 (2016).
35 Naren, A. P. *et al.* CFTR chloride channel regulation by an interdomain interaction. *Science* **286**, 544-548, doi:10.1126/science.286.5439.544 (1999).
36 Fu, J., Ji, H. L., Naren, A. P. & Kirk, K. L. A cluster of negative charges at the amino terminal tail of CFTR regulates ATP-dependent channel gating. *J Physiol* **536**, 459-470, doi:10.1111/j.1469-7793.2001.0459c.xd (2001).

37 Peters, K. W., Qi, J., Johnson, J. P., Watkins, S. C. & Frizzell, R. A. Role of snare proteins in CFTR and ENaC trafficking. *Pflugers Arch* **443 Suppl 1**, S65-69, doi:10.1007/s004240100647 (2001).
38 Lukacs, G. L. *et al.* Conformational maturation of CFTR but not its mutant counterpart (delta F508) occurs in the endoplasmic reticulum and requires ATP. *EMBO J* **13**, 6076-6086 (1994).
39 Hudson, R. P. *et al.* Direct Binding of the Corrector VX-809 to Human CFTR NBD1: Evidence of an Allosteric Coupling between the Binding Site and the NBD1:CL4 Interface. *Mol Pharmacol* **92**, 124-135, doi:10.1124/mol.117.108373 (2017).
40 Okiyoneda, T. *et al.* Mechanism-based corrector combination restores DeltaF508-CFTR folding and function. *Nat Chem Biol* **9**, 444-454, doi:10.1038/nchembio.1253 (2013).
41 Loo, T. W., Bartlett, M. C. & Clarke, D. M. Corrector VX-809 stabilizes the first transmembrane domain of CFTR. *Biochem Pharmacol* **86**, 612-619, doi:10.1016/j.bcp.2013.06.028 (2013).
42 Du, K. & Lukacs, G. L. Cooperative assembly and misfolding of CFTR domains in vivo. *Mol Biol Cell* **20**, 1903-1915, doi:10.1091/mbc.E08-09-0950 (2009).
43 Okiyoneda, T. *et al.* Chaperone-Independent Peripheral Quality Control of CFTR by RFFL E3 Ligase. *Dev Cell* **44**, 694-708 e697, doi:10.1016/j.devcel.2018.02.001 (2018).
44 Okiyoneda, T. *et al.* Peripheral protein quality control removes unfolded CFTR from the plasma membrane. *Science* **329**, 805-810, doi:10.1126/science.1191542 (2010).
45 Pankow, S. *et al.* F508 CFTR interactome remodelling promotes rescue of cystic fibrosis. *Nature* **528**, 510-516, doi:10.1038/nature15729 (2015).
46 Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc* **4**, 484-494, doi:10.1038/nprot.2009.21 (2009).
47 Washburn, M. P., Wolters, D. & Yates, J. R., 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* **19**, 242-247 (2001).
48 He, L., Diedrich, J., Chu, Y. Y. & Yates, J. R., 3rd. Extracting Accurate Precursor Information for Tandem Mass Spectra by RawConverter. *Anal Chem* **87**, 11361-11367, doi:10.1021/acs.analchem.5b02721 (2015).
49 Xu, T. *et al.* ProLuCID: An improved SEQUEST-like algorithm with enhanced sensitivity and specificity. *J Proteomics* **129**, 16-24, doi:10.1016/j.jprot.2015.07.001 (2015).
50 Elias, J. E., Haas, W., Faherty, B. K. & Gygi, S. P. Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nat Methods* **2**, 667-675, doi:10.1038/nmeth785 (2005).
51 Park, S. K. *et al.* Census 2: isobaric labeling data analysis. *Bioinformatics* **30**, 2208-2209, doi:10.1093/bioinformatics/btu151 (2014).

52 Park, S. K. & Yates, J. R., 3rd. Census for proteome quantification. *Curr Protoc Bioinformatics* **Chapter 13**, Unit 13 12 11-11, doi:10.1002/0471250953.bi1312s29 (2010).
53 Bamberger, C. *et al.* Deducing the presence of proteins and proteoforms in quantitative proteomics. *Nature communications* **9**, 2320, doi:10.1038/s41467-018-04411-5 (2018).

Figure 1

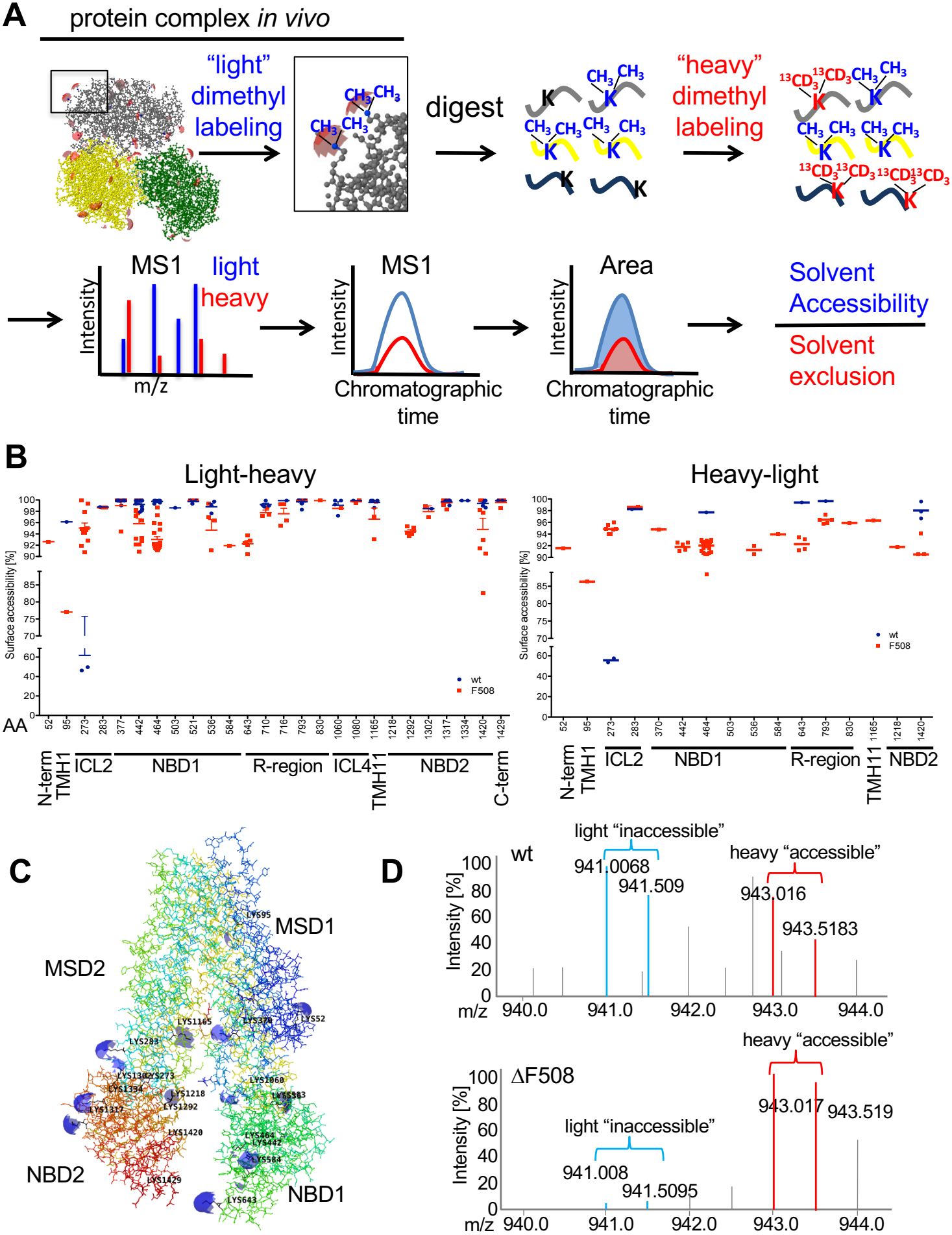
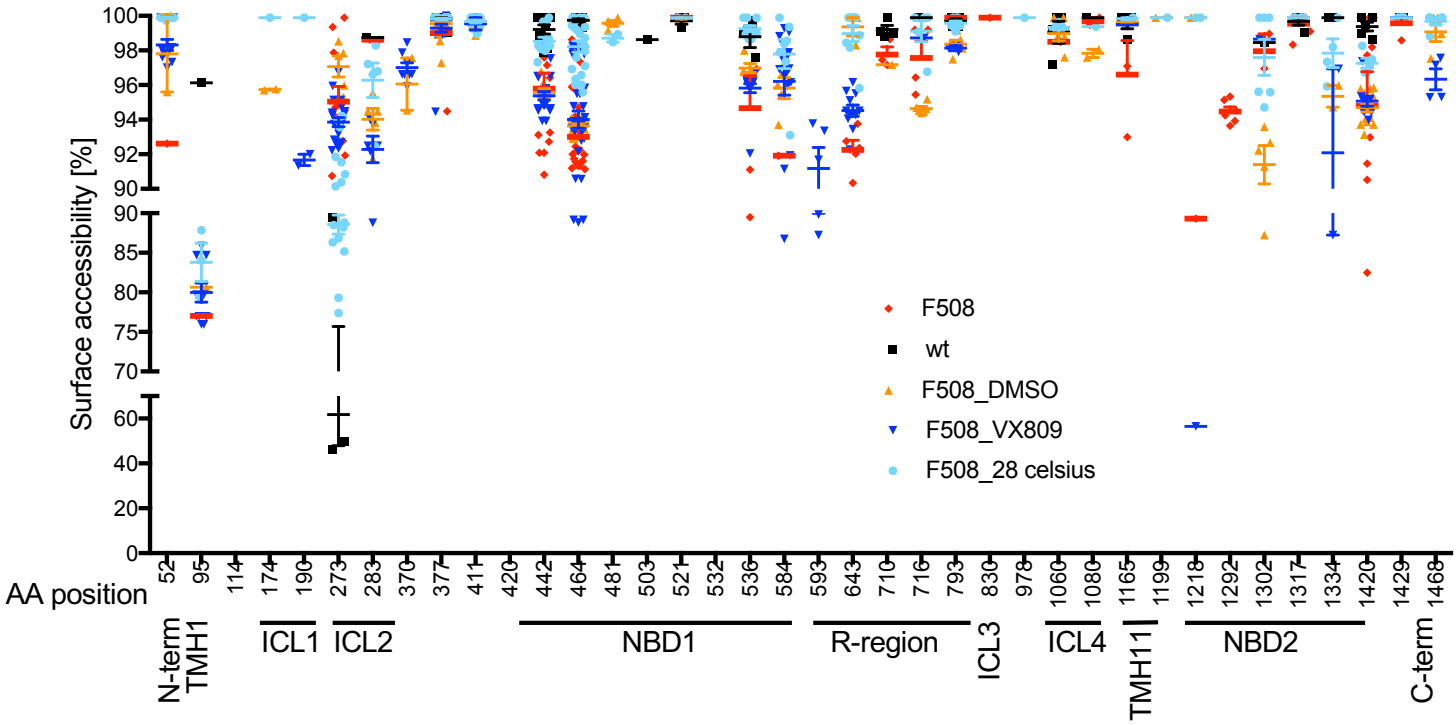
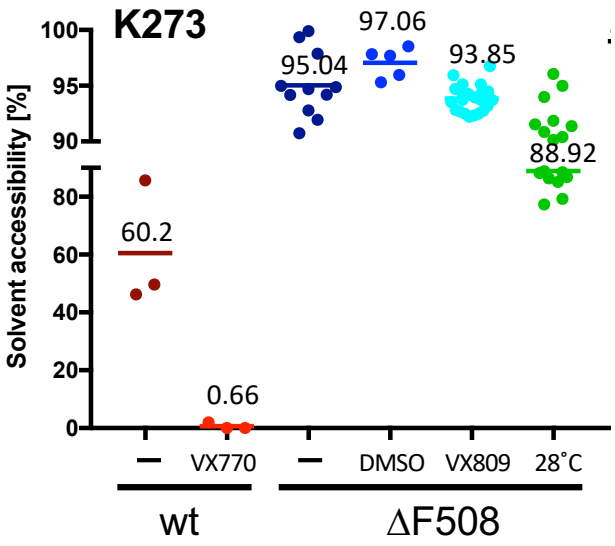

A
protein complex in vivo
"light" dimethyl labeling
digest
"heavy" dimethyl labeling
MS1
light
heavy
Intensity
m/z
MS1
Intensity
Chromatographic time
Area
Intensity
Chromatographic time
Solvent Accessibility
Solvent exclusion
B
Light-heavy
Heavy-light
Surface accessibility [%]
wt
F508
AA
N-term
TMH1
ICL2
NBD1
R-region
ICL4
TMH11
NBD2
C-term
C
MSD1
MSD2
NBD2
NBD1
D
wt
light "inaccessible"
heavy "accessible"
941.0068
941.509
943.016
943.5183
Intensity [%]
m/z
ΔF508
heavy "accessible"
light "inaccessible"
943.017
943.519
941.008
941.5095

Figure 2

A


Surface accessibility [%]
AA position
F508
wt
F508_DMSO
F508_VX809
F508_28 celsius
N-term
TMH1
ICL1
ICL2
NBD1
R-region
ICL3
ICL4
TMH11
NBD2
C-term

B


K273
Solvent accessibility [%]
60.2
0.66
95.04
97.06
93.85
88.92
VX770
DMSO
VX809
28˚C
wt
ΔF508

C

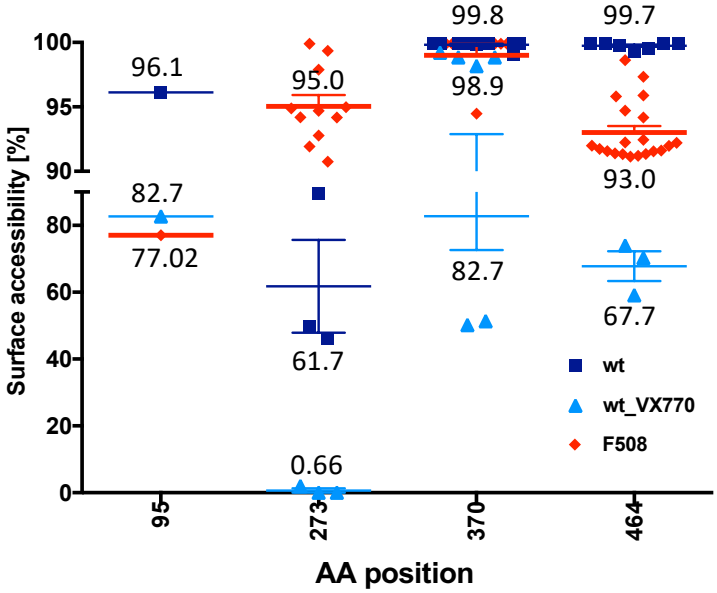
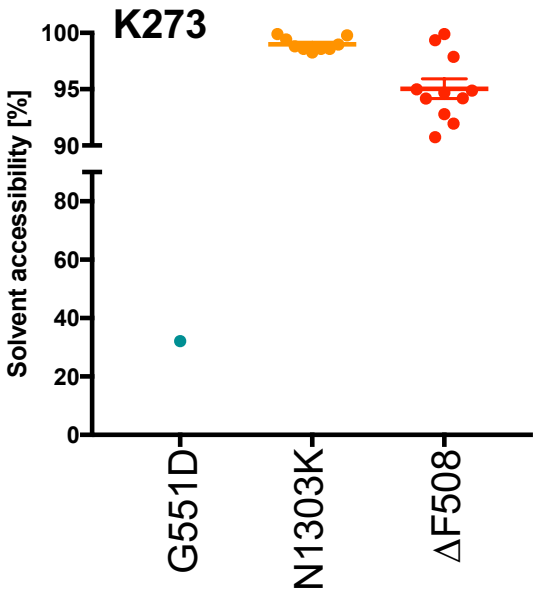

Surface accessibility [%]
96.1
82.7
77.02
95.0
61.7
0.66
99.8
98.9
82.7
99.7
93.0
67.7
wt
wt_VX770
F508
AA position

Figure 3

A


K273
Solvent accessibility [%]
100
95
90
80
60
40
20
0
G551D
N1303K
ΔF508

B

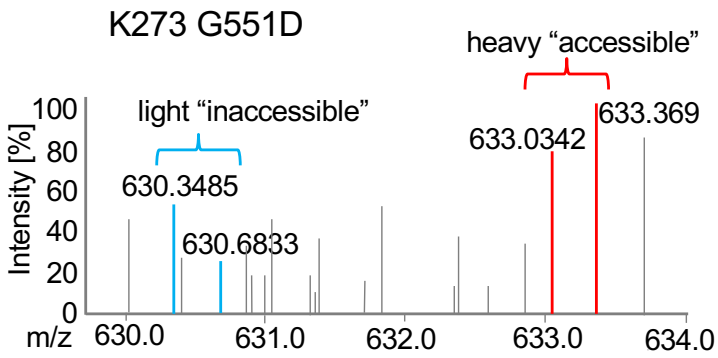

K273 G551D
heavy "accessible"
light "inaccessible"
633.369
633.0342
630.3485
630.6833
Intensity [%]
100
80
60
40
20
0
m/z
630.0
631.0
632.0
633.0
634.0

Figure 4

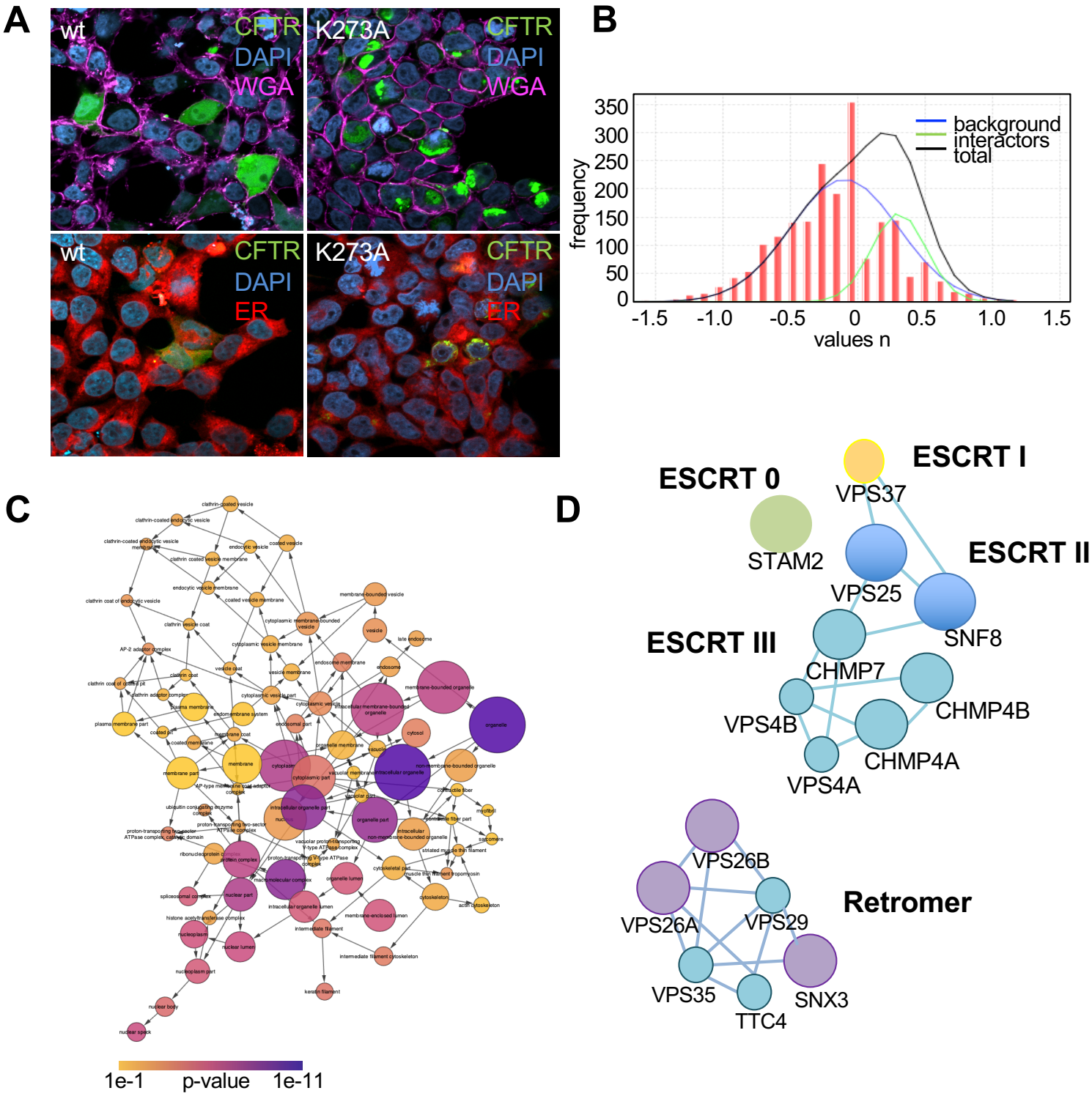
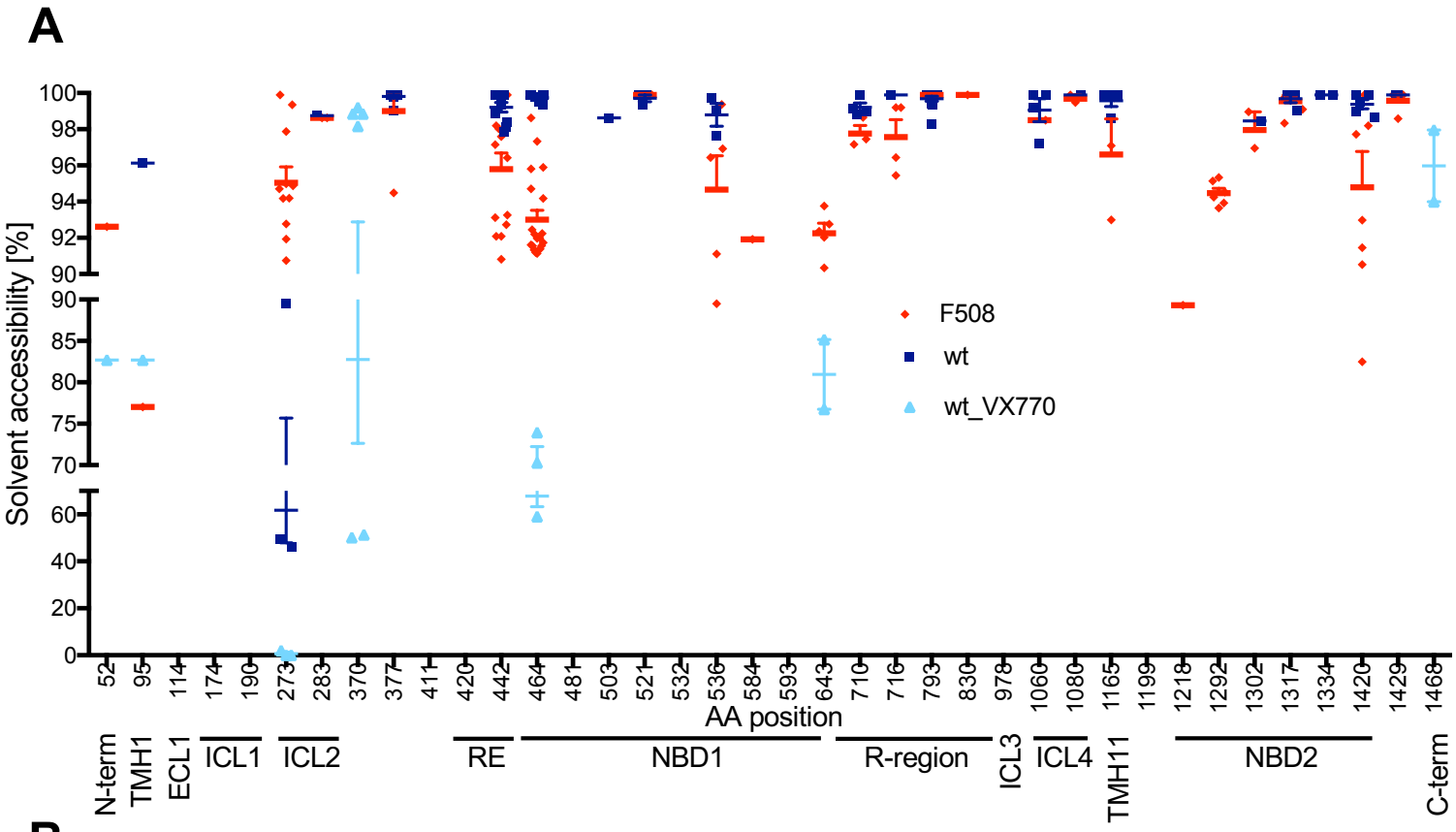
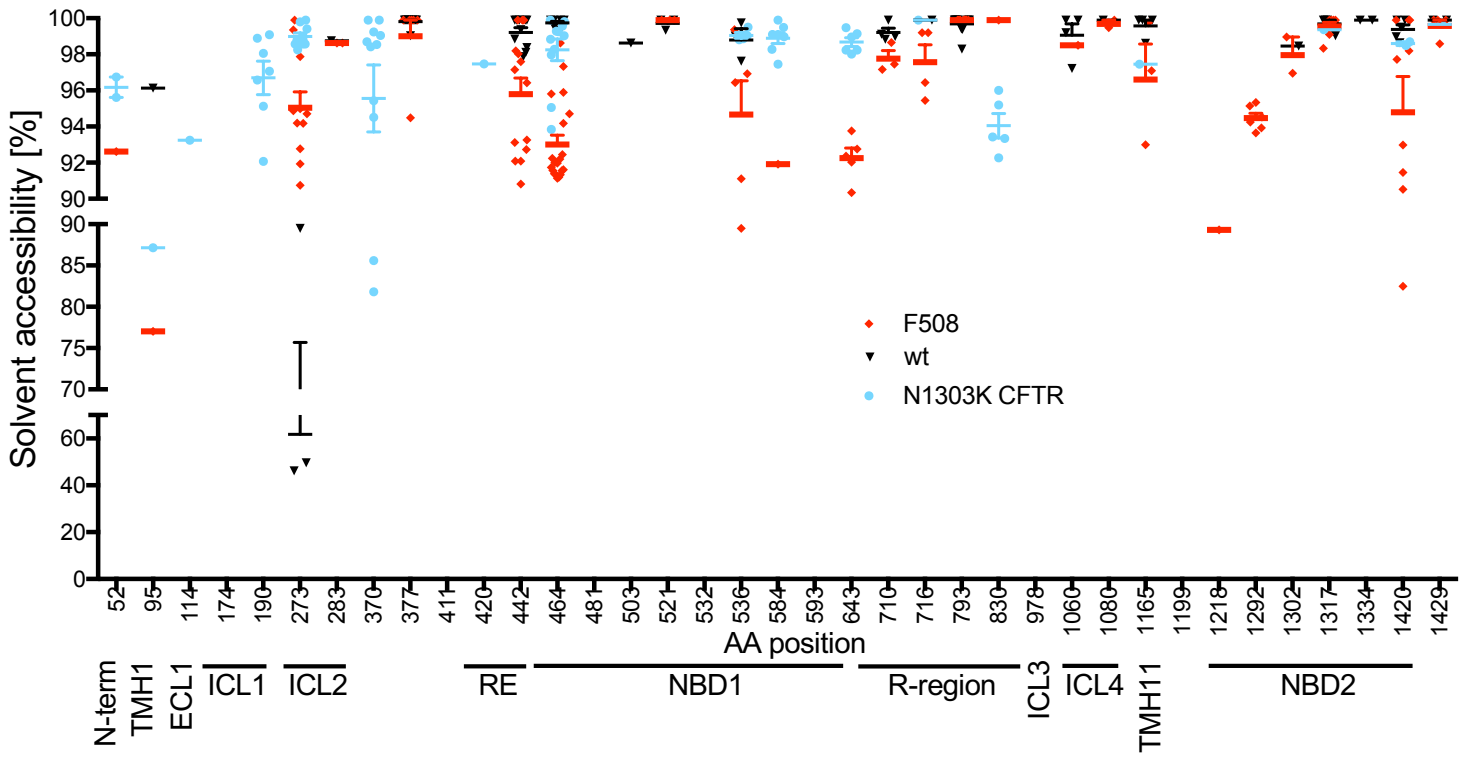
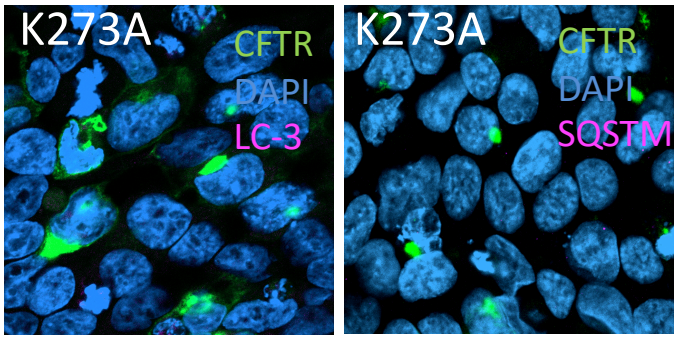

A
wt
CFTR
DAPI
WGA
K273A
CFTR
DAPI
WGA
wt
CFTR
DAPI
ER
K273A
CFTR
DAPI
ER
B
frequency
values n
background
interactors
total
C
1e-1
p-value
1e-11
D
ESCRT 0
STAM2
ESCRT I
VPS37
ESCRT II
VPS25
SNF8
ESCRT III
CHMP7
CHMP4B
VPS4B
CHMP4A
VPS4A
Retromer
VPS26B
VPS26A
VPS29
VPS35
SNX3
TTC4

Table 1

Position	sequence	Mean surface accessibility (%)		Wt Cryo-EM structure	Wt cryo-EM structure, ATP-bound, phosphorylated
		wt	ΔF508		
K52	QIPSVDSADNLSEKL	nd	92.6	accessible (93.75 Å^2)	accessible (13.39 Å^2)
K95	LGEVTKAVQPL	96.1	77.0	partially accessible (3.51 Å^2)	accessible (35.78 Å^2)
K273	VITSEMIENIQSVKAY	61.8	95.0	inaccessible (0.48 Å^2)	partially accessible (6.92 Å^2)
K283	CWEEAMEKMIENLRQTEL	98.7	98.9	accessible (53.98 Å^2)	accessible (39.15 Å^2)
K370/K377	DSLGAINKIQDFLQKQEY	99.8	98.9	accessible (41.45 Å^2, 39.11 Å^2)	accessible (29.92 Å^2, 36.11 Å^2)
K411	not present in CryoEM				
K420	not present in CryoEM				
K442	SLLGTPVLKDINF	99.2	95.7	accessible (39.53 Å^2)	accessible (47.84 Å^2)
K464	AVAGSTGAGKTSLL	99.7	93.0	accessible (9.59 Å^2)	partially accessible (4.36 Å^2)
K536	AEKDNIVLGEGGITL	98.8	94.7	accessible (51.94 Å^2)	accessible (26.99 Å^2)
K584	GYLDVLTEKEIF	nd	91.9	accessible (8.91 Å^2)	accessible (18.03 Å^2)
K1060	ESEGRSPIFTHLVTSLKGLW	99.0	98.5	accessible (49.05 Å^2)	partially accessible (3.15 Å^2)
K1080	FETLFHKAL	99.9	99.7	accessible (59.33 Å^2)	accessible (35.59 Å^2)
K1165	KFIDMPTEGKPT	99.5	96.6	accessible (19.66 Å^2)	accessible (54.57 Å^2)
K1218	TAKYTEGGNAILENISF	nd	89.3	accessible (41.39 Å^2)	accessible (30.56 Å^2)
K1292	GVIPQKVF	nd	94.4	accessible (48.20 Å^2)	partially accessible (3.01 Å^2)
K1302	RKNLDPYEQWSDQEIW	98.4	97.9	accessible (26.64 Å^2)	accessible (13.86 Å^2)
K1317	KVADEVGLRSVIEQFPGKLDF	99.7	99.5	accessible (37.29 Å^2)	accessible (53.09 Å^2)
K1334	RSVIEQFPGKLDF	99.9	nd	accessible (56.92 Å^2)	accessible (49.60 Å^2)
K1420	LVIEENKVRQY	99.3	94.8	accessible (51.71 Å^2)	accessible (51.65 Å^2)
K1429	DSIQKLL	99.9	99.5	accessible (57.27 Å^2)	accessible (33.64 Å^2)

Supplemental Figure 1





Supplemental Figure 2