Figure S1

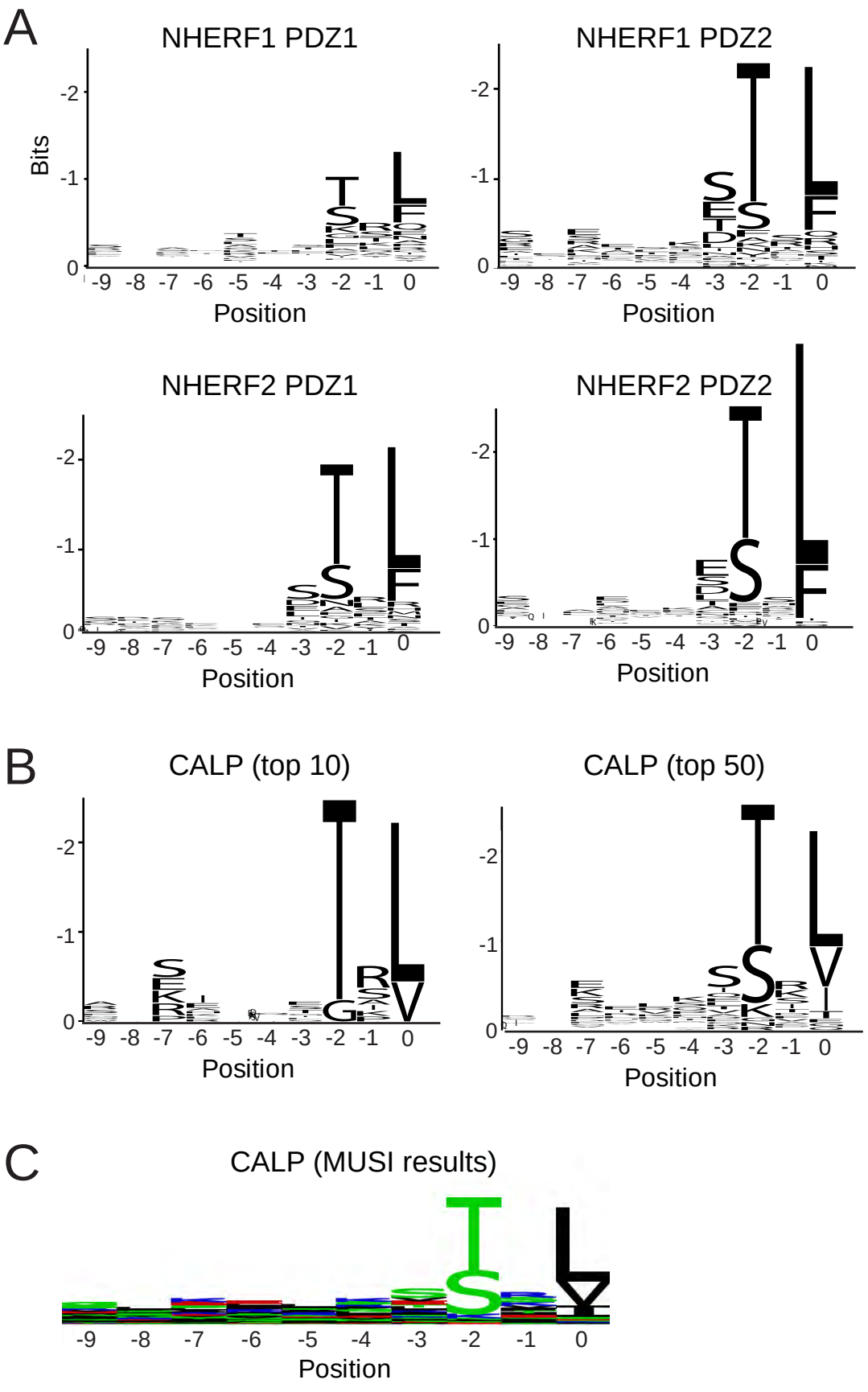

A
NHERF1 PDZ1
NHERF1 PDZ2
NHERF2 PDZ1
NHERF2 PDZ2
Bits
Position
B
CALP (top 10)
CALP (top 50)
C
CALP (MUSI results)

Figure S2

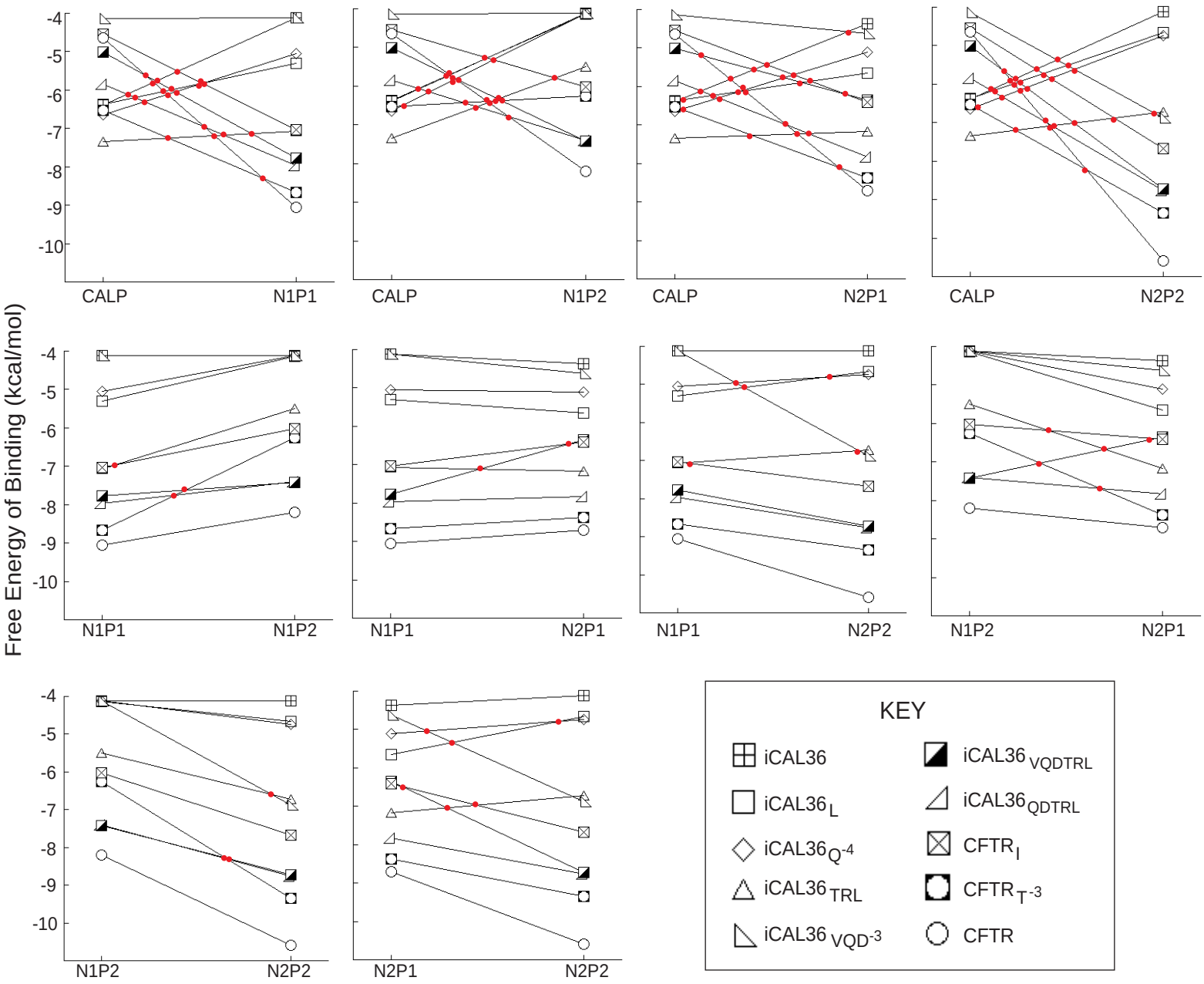
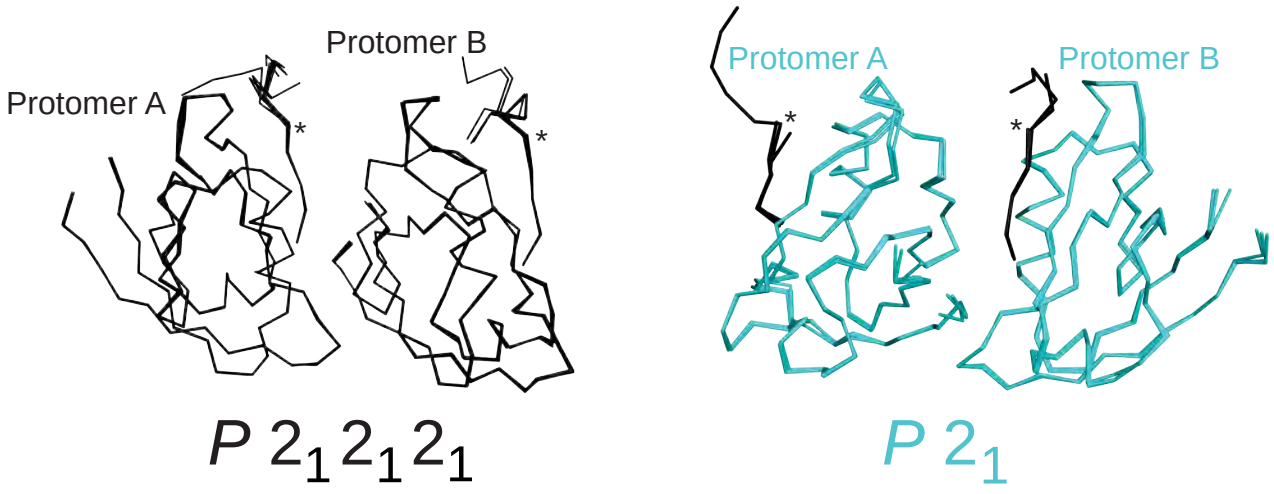

Free Energy of Binding (kcal/mol)
-4
-5
-6
-7
-8
-9
-10
CALP
N1P1
N1P2
N2P1
N2P2
KEY
iCAL36
iCAL36 L
iCAL36 Q-4
iCAL36 TRL
iCAL36 VQD-3
iCAL36 VQDTRL
iCAL36 QDTRL
CFTR I
CFTR T-3
CFTR

Figure S3


Protomer A
Protomer B
*
*
P 2₁ 2₁ 2₁
Protomer A
Protomer B
*
*
P 2₁

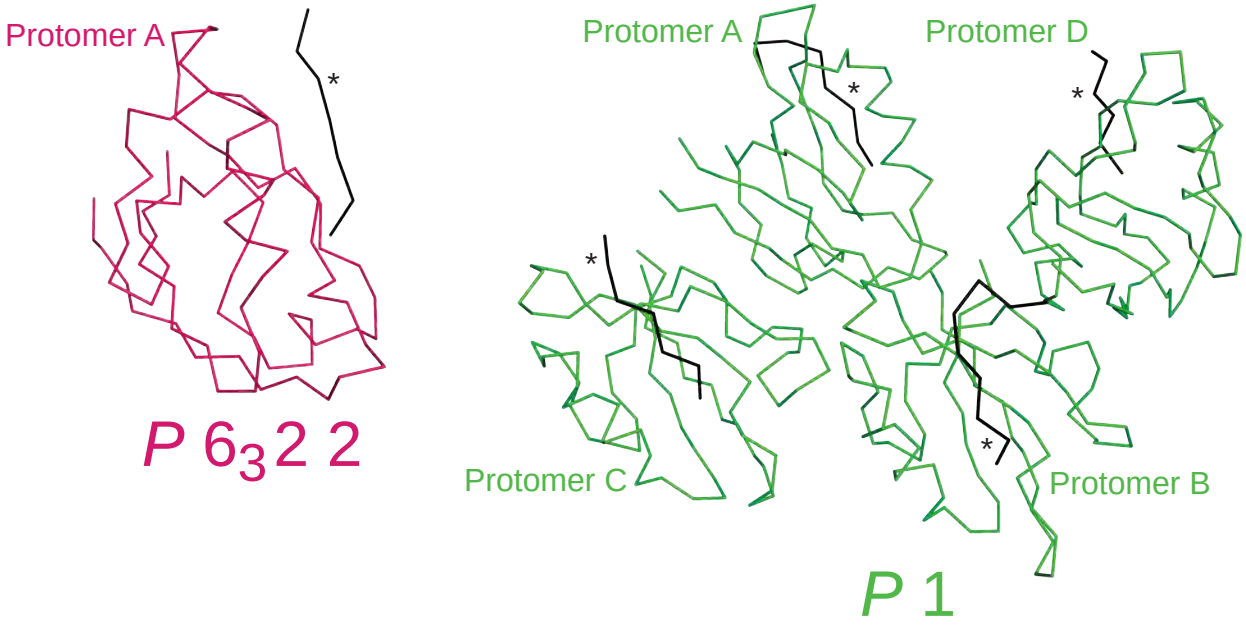

Protomer A
*
P 6₃ 2 2
Protomer A
Protomer D
*
*
*
*
Protomer C
Protomer B
P 1

Figure S4

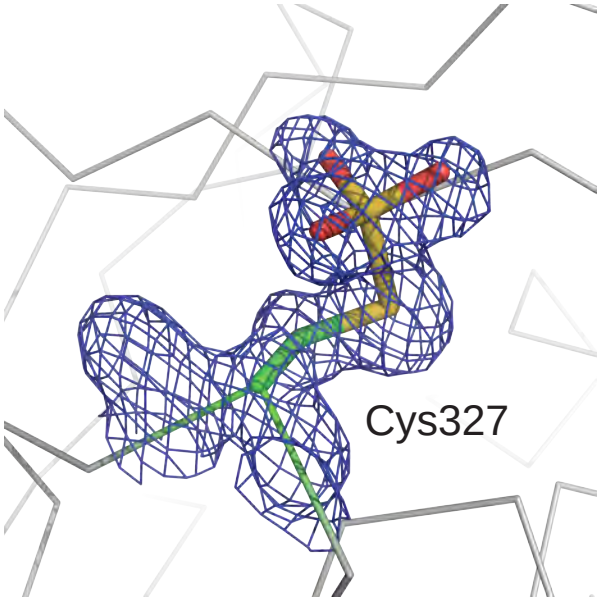
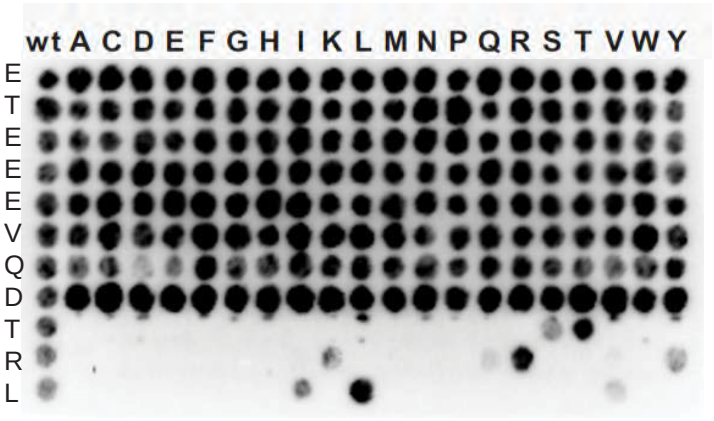
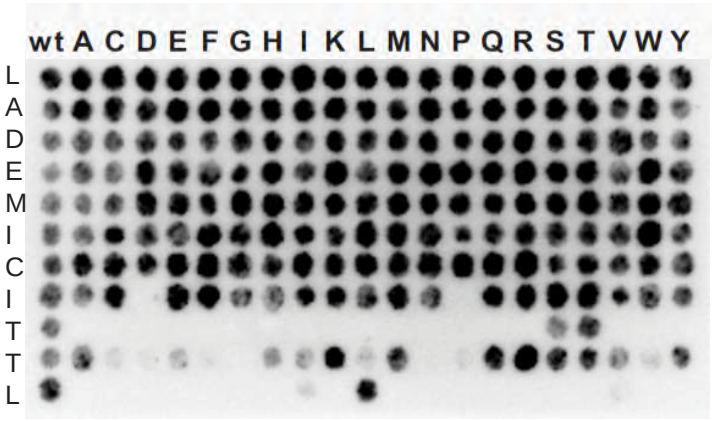


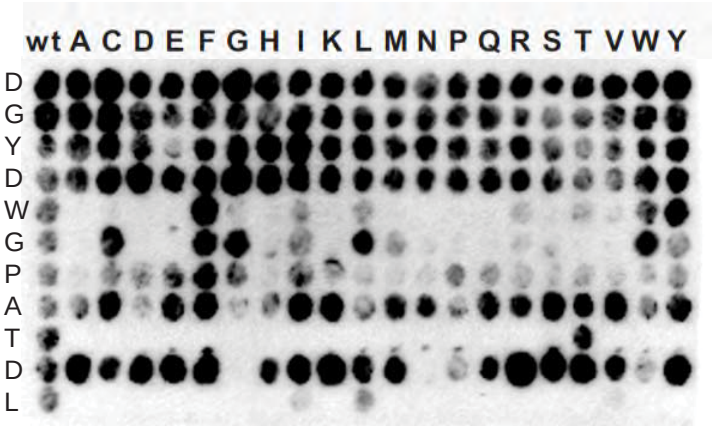
Figure S5



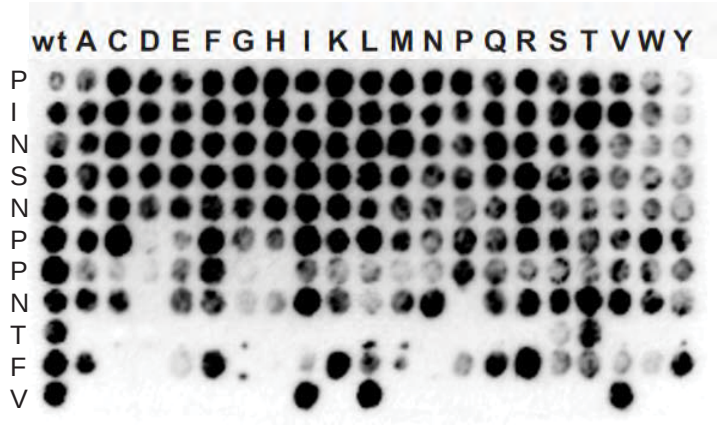
CFTR



CAC1D

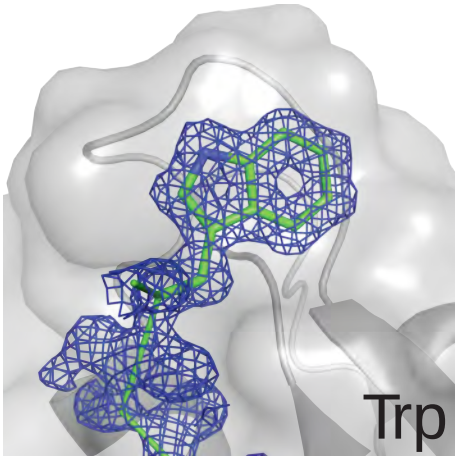


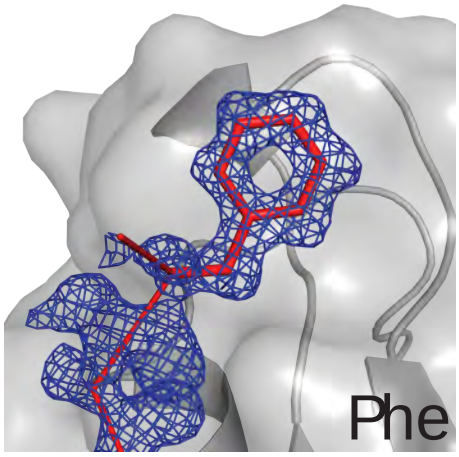
ZO-3

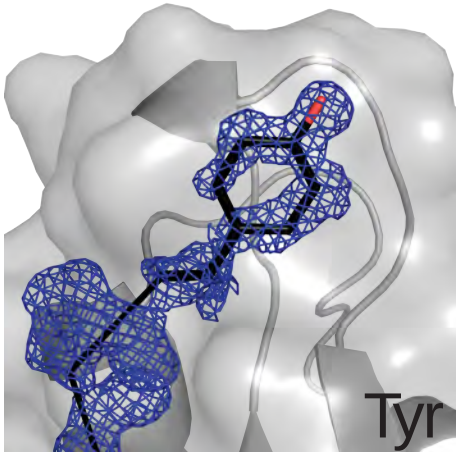


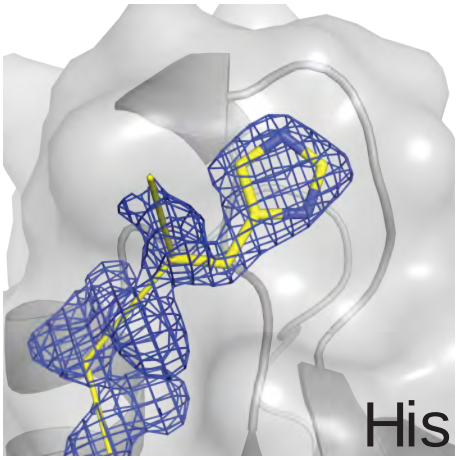
GPR19

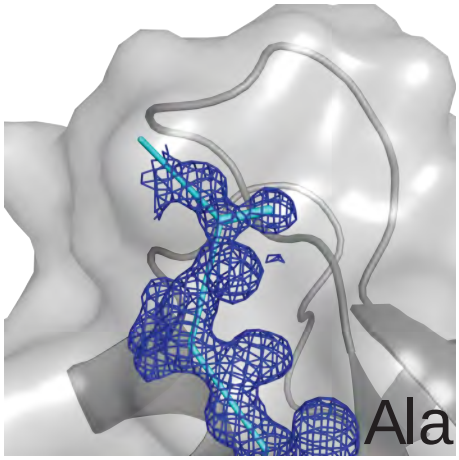
Figure S6

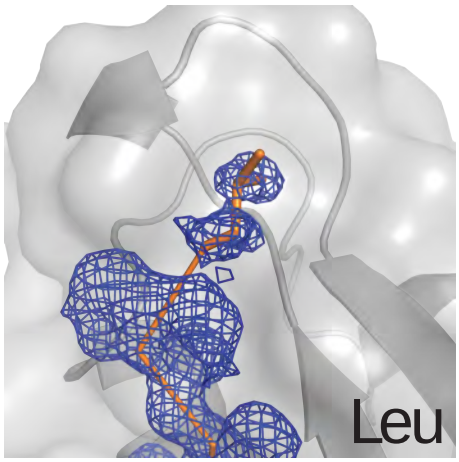












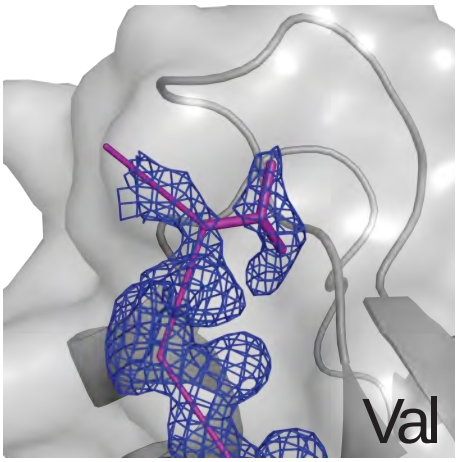


Figure S7

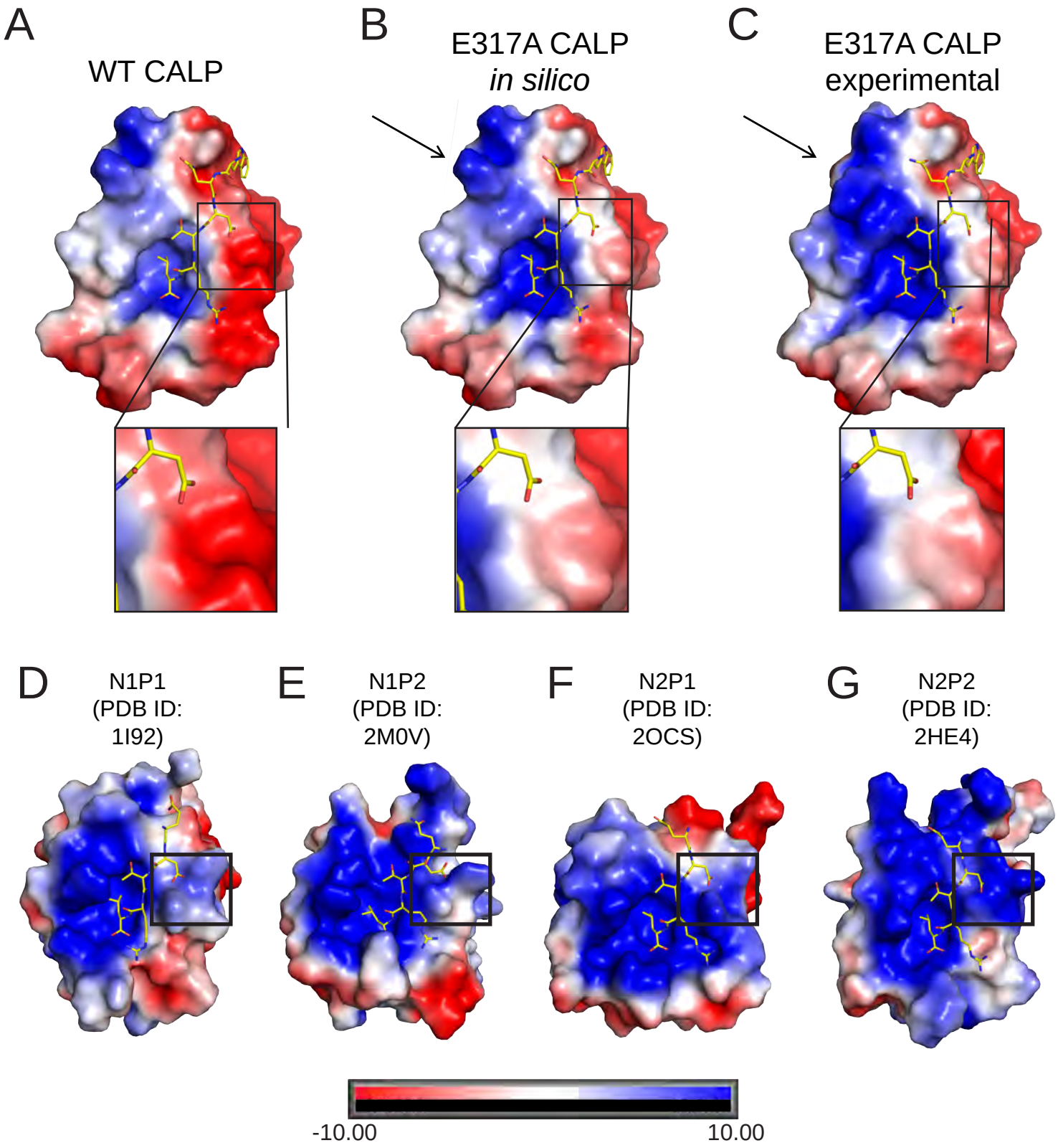

A
WT CALP
B
E317A CALP
in silico
C
E317A CALP
experimental
D
N1P1
(PDB ID:
1I92)
E
N1P2
(PDB ID:
2M0V)
F
N2P1
(PDB ID:
2OCS)
G
N2P2
(PDB ID:
2HE4)
-10.00
10.00

Figure S8

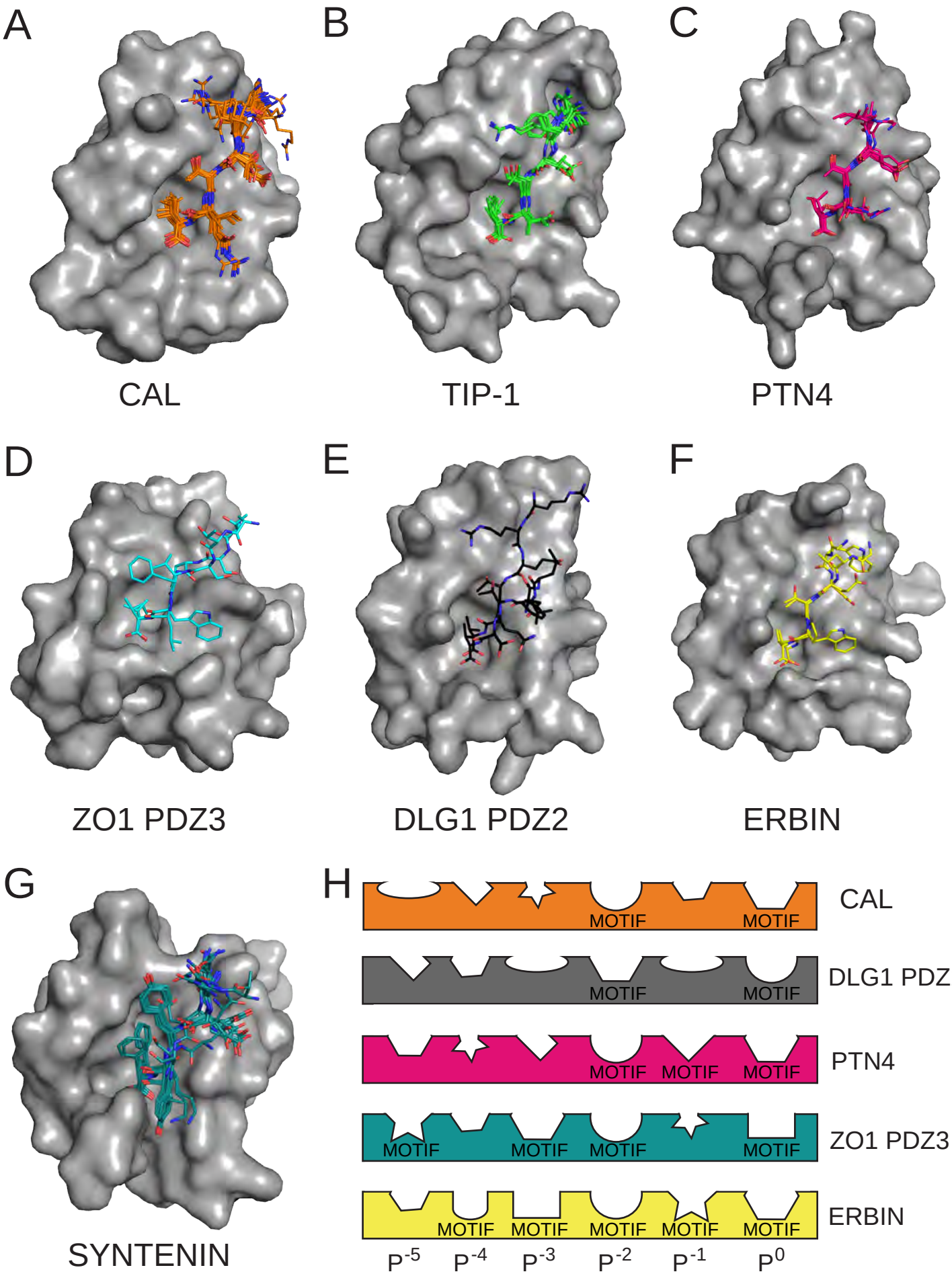

A
CAL
B
TIP-1
C
PTN4
D
ZO1 PDZ3
E
DLG1 PDZ2
F
ERBIN
G
SYNTENIN
H
MOTIF
MOTIF
CAL
MOTIF
MOTIF
DLG1 PDZ2
MOTIF
MOTIF
MOTIF
PTN4
MOTIF
MOTIF
MOTIF
MOTIF
ZO1 PDZ3
MOTIF
MOTIF
MOTIF
MOTIF
MOTIF
ERBIN
P-5
P-4
P-3
P-2
P-1
P0

Figure S1, related to Figure 1A. WebLogo analysis of NHERF1/2 PDZ domains and CALP.
(A) WebLogo analyses are shown for the top 100 sequences that bound N1P1, N1P2, N2P1, or N2P2 in experiments with 6223HumLib peptide arrays. The C-terminal residue position is labeled 0, adjacent residues are -1, -2, etc. The WebLogo reveals clear preferences at the P^0 and P^{-2} positions, and no significant selectivity at other positions.
(B) WebLogo analysis of the top 10 or top 50 sequences that bound CALP in a 6223HumLib array experiment. Although a P^{-1} Arg preference is apparent in the WebLogo from the top 10 sequences, changing preferences at all positions make analysis difficult.
(C) Logo output from the MUltiple Specificity Identifier (MUSI) program (Kim et al., 2011), using the top 100 sequences that bound CALP in a 6223HumLib peptide array reveals an identical, degenerate, binding motif as WebLogo analysis. Positions are numbered according to the MUSI program (10 = P^0, 9 = P^{-1}, and so on).

Figure S2, related to Figure 1B. Non-motif preferences in CAL and NHERF1/2 PDZ domains.
Free energies of binding (kcal/mol) were calculated from the binding affinities of the CALP NHERF1/2 PDZ domains for 10 different peptides, and are plotted in pairwise comparisons. Selectivity determinants are revealed by intersection points, indicating differences in free energy of binding that alter the rank order of binding preferences.

Figure S3, related to Figure 3A. Asymmetric units of CALP:peptide co-crystals formed in four space-groups. CALP:peptide structures (C_α traces) are shown following CALP main-chain alignment and labeled by protomers of the asymmetric unit. Peptides are colored black, and indicated with an asterisk (*) for reference. Structures in the $P2_12_12_1$ space-group (black) include A-iCAL36, V-iCAL36, L-iCAL36, H-iCAL36, F-iCAL36, Y-iCAL36, and $iCAL36_{Q\text{-}4}$. All structures align by main-chain atoms to the A-protomer of CALP:iCAL36 with an RMSD value < 1.7 Å (the smallest difference is 0.11 Å ($iCAL36_Q$) and the largest is 1.7 Å (H-iCAL36). Structures in the $P2_1$ space-group (cyan) include $iCAL36_{VQD\text{-}3}$, and the HPV E6 peptides. When aligned by main-chain atoms, the HPV E6 structures are very similar to each other, RMSD = 0.19 Å. The $iCAL36_{VQD\text{-}3}$ structure aligns to the HPV16 E6 structure with an RMSD value of 0.32 Å. The CALP:$iCAL36_{QDTRL}$ complex crystallized in the $P6_322$ space-group (pink). CALP:$iCAL36_{TRL}$ crystallized in $P1$ (green), with four molecules in the asymmetric unit.

Figure S4. Sulfocysteine residue modeled in $iCAL36_{QDTRL}$ complex structures. CALP is shown as C_α trace (gray), except for Cys^{327}, which is modeled as a sulfocysteine residue (stick model, green= C, red = O, yellow = S). The final refined $2F_o$-F_c electron density map (blue mesh, 1σ) showed clear density for the modification.

Figure S5, related to Figure 4. CALP SubAna peptide arrays for the CFTR, CAC1D, ZO-3, and GPER sequences. The SubAna peptide arrays are shown for the CFTR, CAC1D, ZO-3, and GPER sequences, as labeled. Starting sequences (starting sequence labeled along left side of array), were substituted at each position by all naturally occurring amino acids (substitutions indicated by single letter codes along top of peptide array), and probed for CALP binding using an α-His antibody. Dark spots indicate that binding occurred, whereas the absence of a spot indicates that binding was reduced by the amino acid substitution.

Figure S6, related to Figure 5B. Electron density for P^{-5}-substituted iCAL36 peptides. The B-protomers of CALP (grey cartoon and transparent surface) are shown in the vicinity of the P^{-5} residue (stick figure) together with P^{-5}-substituted peptides (C_α trace and P^{-5} side chain atoms), colored and labeled by P^{-5} residue identity as follows: Trp (green carbons), Phe (red), Tyr (black),

His (yellow), Ala (cyan), Leu (orange), Val (purple). The final refined $2F_o$-F_c density maps (blue mesh) corresponding to each peptide are contoured at 1σ. Non-carbon atoms are colored by element (red = O, blue = N).

Figure S7, related to Figure 5C-D. Electrostatic potential surface maps of CALP and NHERF1 PDZ1.

(A-C) Electrostatic potential surface maps were calculated as in Figure 4, at 10 k_BT/e, and shown for WT-CALP (A), as well as *in silico* (B) and experimental (C) models of the CALP-E317A mutant.

(D-G) Electrostatic potential surface of N1P1:CFTR (PDB entry 1I92), N1P2:CFTR (PDB entry 2M0V), N2P1 (PDB entry 2OCS; peptide from CALP:iCAL36$_{QDTRL}$ structure), and N2P2 (PDB entry 2HE4; peptide from CALP:iCAL36$_{QDTRL}$ structure), for comparison. All surfaces are rendered at 10 k_BT/e. The black boxes highlight the favorable, positive surfaces of all NHERF1/2 PDZ domains that interact with the P^{-3} Asp residue.

Figure S8, related to (A-G) Figure 3 and (H) Figure 6. Pharmacophore models of PDZ domain binding.

(A-G) PDZ domains are shown (gray surface) for which PDB structures are available for complexes with multiple peptides (excluding C-terminal lattice contacts). All structures for a given domain were superimposed by PDZ main-chain atoms. Peptide ligands (stick figures) are shown in the binding pocket, with carbon atoms colored by PDZ target. The domains (labeled) and associated PDB entry codes are: CAL (A; see text); TIP-1 (B; 3DIW, 3GJ9, 3SFJ, 4E3B); PTN4 (C; 3NFK, 3NFL); ZO1 PDZ3 (D; 3SHW, 3TSZ); DLG1 PDZ2 (E; 2OQS, 3RL8); Erbin (F; 1MFG, 1N7T); and syntenin (G; 1OBX, 1OBY, 1OBZ, 1V1T , 1W9E, 1W9O, 1W9Q, 1YBO). (H) Pharmacophore models of the series of side chain-binding sites along the PDZ cleft. Sites were identified as 'motif' positions if high-throughput analysis (Figure 1 for CALP motif; Tonikian et al., 2008) revealed $\leq$ 4 residue preferences at that position ($\leq$ 20% of all possibilities). As seen for CAL (Figure 3), non-motif residues interact with defined local environments, suggesting that modulator residues may influence binding in each of these PDZ domain, as well.

Table S1: Crystallization and data collection conditions for CALP:peptide co-complexes.

Peptide	PDB ID	Sequence	Space Group Number	Crystallization Buffer Conditions	Cryoprotectant Buffer Conditions	Data Collection parameters
A-iCAL36	4JOE	ANSRAPTSII	19	31% (*w/v*) PEG 3350, 50 mM NaCl, 100 mM Tris pH 7.5	Crystallization buffer + 20% (*v/v*) glycerol	$\Delta\varphi^a = 0.3°$ 180° collection $\lambda^b = 1.000$ Å
L-iCAL36	4JOF	ANSRLPTSII	19	36% (*w/v*) PEG 8000, 200 mM NaCl, 10 mM Tris HCl pH 7.95	Crystallization buffer + 20% (*v/v*) glycerol	$\Delta\varphi = 0.5°$ 180° collection $\lambda = 1.000$ Å
V-iCAL36	4JOG	ANSRVPTSII	19	40% (*w/v*) PEG 3350, 200 mM NaCl, 100 mM Tris pH 7.5	35% (*w/v*) PEG 3350, 200 mM NaCl, 100 mM Tris pH 7.5 20% (*v/v*) glycerol	$\Delta\varphi = 0.3°$ 180° collection $\lambda = 1.000$ Å
H-iCAL36	4JOH	ANSRHPTSII	19	35% (*w/v*) PEG 8000, 150 mM NaCl, 100 mM Tris pH 7.5	Crystallization buffer + 20% (*v/v*) glycerol	$\Delta\varphi = 0.3°$ 180° collection $\lambda = 0.9795$ Å
F-iCAL36	4JOJ	ANSRFPTSII	19	34% (*w/v*) PEG 3350, 150 mM NaCl, 100 mM Tris pH 7.5	Crystallization buffer + 20% (*v/v*) glycerol	$\Delta\varphi = 0.3°$ 180° collection $\lambda = 1.000$ Å
Y-iCAL36	4JOK	ANSRYPTSII	19	34% (*w/v*) PEG 3350, 50 mM NaCl, 100 mM Tris pH 7.5	Crystallization buffer + 20% (*v/v*) glycerol	$\Delta\varphi = 0.3°$ 180° collection $\lambda = 1.000$ Å
$iCAL36_{Q-4}$	4K6Y	ANSRWQTSII	19	35% (*w/v*) PEG 3350, 100 mM NaCl, 100 mM Tris pH 7.5	Crystallization buffer + 16% (*v/v*) glycerol	$\Delta\varphi = 0.3°$ 360° collection $\lambda = 1.000$ Å
$iCAL36_{TRL}$	4K76	ANSRWPTTRL	1	26% (*w/v*) PEG 3350, 50 mM NaCl, 100 mM Tris pH 7.5	35% (*w/v*) PEG 3350, 100 mM NaCl, 100 mM Tris pH 7.5, 20% (*v/v*) glycerol	$\Delta\varphi = 0.3°$ 210° collection $\lambda = 0.8856$ Å
$iCAL36_{VQD-3}$	4K72	ANSRVQDSII	4	40% (*w/v*) PEG 1000, 100 mM $N_2S_2O_3{\cdot}5H_2O$, 100 mM Tris pH 8.0	30% (*w/v*) PEG 8000, 100 mM $Na_2S_2O_3{\cdot}5H_2O$ 100 mM Tris pH 8.0 20% (*v/v*) glycerol	$\Delta\varphi = 0.3°$ 180° collection $\lambda = 0.8856$ Å
$iCAL36_{QDTRL}$	4K75	ANSRWQDTRL	182	40% (*w/v*) PEG 1000, 90 mM $N_2S_2O_3{\cdot}5H_2O$, 100 mM Tris pH 8.0	60% (*w/v*) PEG 400, 90 mM $Na_2S_2O_3{\cdot}5H_2O$ 100 mM Tris pH 8.0	$\Delta\varphi = 0.2°$ 200° collection $\lambda = 1.000$ Å
CALP E317A + $iCAL36_{QDTRL}$	4K78	ANSRWQDTRL	182	40% (*w/v*) PEG 1000, 90 mM $N_2S_2O_3{\cdot}5H_2O$, 100 mM Tris pH 8.0	60% (*w/v*) PEG 400, 90 mM $Na_2S_2O_3{\cdot}5H_2O$ 100 mM Tris pH 8.0	$\Delta\varphi = 0.2°$ 180° collection $\lambda = 0.9795$ Å
HPV16 E6	4JOP	SSRTRRETQL	4	32% (*w/v*) PEG 3350, 125 mM NaCl, 100 mM Tris pH 7.5	35% (*w/v*) PEG 3350, 125 mM NaCl, 100 mM Tris pH 7.5, 16% (*v/v*) glycerol	$\Delta\varphi = 1.0°$ 220° collection $\lambda = 1.000$ Å
HPV18 E6	4JOR	RLQRRRETQV	4	29% (*w/v*) PEG 3350, 75 mM NaCl, 100 mM Tris pH 7.5	30% (*w/v*) PEG 3350, 100 mM NaCl, 100 mM Tris pH 7.5, 16% (*v/v*) glycerol	$\Delta\varphi = 0.3°$ 360° collection $\lambda = 1.000$ Å

[a] $\Delta\varphi$ = oscillation range per image.
[b] λ = X-ray wavelength.

Supplemental Experimental Procedures

Structure determination

Co-crystallization was performed by vapor diffusion in hanging-drop format at 291K. Drops contained 5.5 mg ml^{-1} CALP in crystallization buffer (10 mM HEPES pH 7.4, 25 mM NaCl) and 1-3 mM peptide, as previously described (Amacher et al., 2013; Amacher et al., 2011). 1 mM peptide was sufficient for most complexes, with the exception of two with poor affinity (H-iCAL36, $iCAL36_{VQD-3}$), and $iCAL36_{QDTRL}$. Crystals appeared within 2-4 d, with the exception of drops containing $iCAL36_{QDTRL}$, where crystals did not appear for approximately 30 d.

CALP complexes with the P^{-5}-substituted peptides, $iCAL36_{Q-4}$, $iCAL36_{TRL}$, and the HPV E6-derived peptides crystallized under similar conditions as CALP:iCAL36, and molecular replacement was used to determine phases (Amacher et al., 2013; Amacher et al., 2011). Reservoir and cryoprotectant buffers are listed in Table S1. $iCAL36_{VQD-3}$ and $iCAL36_{QDTRL}$ complexes crystallized under different conditions. These reservoir buffers included sodium thiocyanate pentahydrate ($Na_2S_2O_3 \cdot 5H_2O$) at pH 8.0 (Table S1).

Data were collected on beamline X6A, and data collection parameters for each complex are listed in Table S1. Data were processed using the XDS package (Kabsch, 2010; Amacher et al., 2013). Molecular replacement (MR) was performed using PHENIX (Adams et al., 2010; McCoy, 2007), with the A-protomer of CALP:iCAL36 as a search model for all structures. Peptides were omitted during molecular replacement, initial model building, and refinement, to avoid potential phase bias. This is an approach used previously (Amacher et al., 2013). In each case, clear electron density was observed for the peptide, including the substituted side-chains (data not shown). Model building and refinement were performed using PHENIX (Adams et al., 2010; McCoy, 2007), and manual refinement using COOT (Emsley et al., 2010). Structure geometry was verified using the PDB validation server (Feng et al., 1998; Laskowski et al., 1993; Vaguine et al., 1999).

CALP:peptide co-crystals formed in separate space-groups: $P1$, $P2_1$, $P2_12_12_1$, and $P6_322$ (Table S1; Figure S3). All P^{-5}-substituted co-complexes, as well as the crystal containing iCAL36$_{Q\text{-}4}$, crystallized in the $P2_12_12_1$ space group. These structures contained two molecules in the asymmetric unit. Each of these molecules closely resembled the stereochemistry of the corresponding protomer in the asymmetric unit of CALP:iCAL36 (Amacher et al., 2013). Additionally, there are crystal lattice contacts observed for peptide residues beyond the P^{-6} Arg, as reported previously (Amacher et al., 2013). All data collection and solution statistics are reported in Tables 2 and 4. Our structures reveal minimal global differences in peptide binding to CALP. We superimposed 87 C_α atoms of all CALP A-protomer molecules, with RMSD values at the level of the maximum-likelihood coordinate error (≤ 0.15 Å).

CALP:iCAL36$_{TRL}$ crystallized in space group $P1$, with 4 molecules in the asymmetric unit, and CALP:iCAL36$_{VQD\text{-}3}$ crystallized in space group $P2_1$, with 2 molecules in the asymmetric unit. The $P1$ lattice of CALP:iCAL36$_{TRL}$ contains disulfide bonds between two of the molecules, with the other two forming disulfide bonds with molecules of neighboring symmetry mates. We do not believe these disulfide bonds are indicative of dimer formation; previous work suggests CALP is a monomer in solution (Amacher et al., 2011), and that CAL dimerizes via its coiled-coil domains (Wolde et al., 2007). All data collection and refinement statistics are reported in Table 2.

CALP:iCAL36$_{QDTRL}$ crystallized in a hexagonal space group, $P6_322$, with one molecule in the asymmetric unit. Upon observing clear electron density at the Cys^{319} residue, we modeled in a sulfocysteine residue. Subsequent rounds of refinement strengthen the electron density (Figure S4). Attempts at mass spectrometry-based identification of a sulfocysteine residue were not successful, as the mass shift expected for the sulfocysteine-modified CALP was below the resolution limit of MALDI-MS; however, this moiety does not appear to affect the overall

CALP:iCAL36$_{QDTRL}$ structure. Main-chain alignment of CALP:iCAL36$_{QDTRL}$ and the A-protomer of CALP:iCAL36 yields an RMSD of 0.38 Å.

In general, electron density for low affinity peptides was weaker than for high affinity peptides (data not shown), and only peptide residues with clear density were modeled. Upon observation of high average B-factors for the HPV16 E6 peptides in both protomers of the asymmetric unit, the occupancy of all peptide residues was refined to yield an average value of 0.8. Peptides in other complexes were refined with occupancies set to 1.0. Finally, the HPV E6 peptides show crystal lattice contacts at the P^{-1} Gln residue, with a nearby molecule related by symmetry. However, agreements with the positions of the P^{-1} Gln, as well as P^{0} and P^{-2} residues, with respect to all co-crystal structures, suggest this contact does not affect the overall structure. No other peptides reveal crystal lattice contacts at residues C-terminal to the P^{-6} positions.

During analysis of the P^{-5}-substitutued peptide complexes, protomers of the asymmetric units were treated separately due to a conserved shift in the carboxylate binding loop position, as reported previously (Amacher et al., 2013). Electron density for the P^{-5} side chains is shown in Figure S6. Analysis of the CALP A-protomers gave similar results (data not shown).

SubAna Peptide Arrays

Along with iCAL36, we previously reported results for SubAna arrays of somatostatin receptor subtype 5 (SSR5; ANGLMQTSKL), vasoactive intestinal polypeptide receptor 2 (VIPR2; SFLQTETSVI), microtubulin-associated protein 4 (MAP4; LDSQIQETSI), G-protein coupled estrogen receptor-1 (GPER; QSDVRFSSAV), and breakpoint cluster region (BCR; RQSILFSTEV) C-terminal peptides (Vouilleme et al., 2010). Here, we also report the results from SubAna arrays of voltage-dependent L-type calcium channel subunit alpha-1D (CAC1D; ADEMICITTL), zona occludens protein 3 (ZO-3; GYDWGPATDL), the cystic fibrosis transmembrane conductance regulator (CFTR; TEEEVQDTRL), and G-protein coupled receptor 19 (GPR19; INSNPPNTFV) sequences (Figure S5), yielding a total of 10 SubAnas.

Supplemental References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

Amacher, J.F., Cushing, P.R., Bahl, C.D., Beck, T., and Madden, D.R. (2013). Stereochemical Determinants of C-terminal Specificity in PDZ Peptide-binding Domains: a novel contribution of the carboxylate-binding loop. J Biol Chem *288*, 5114-5126.

Amacher, J.F., Cushing, P.R., Weiner, J.A., and Madden, D.R. (2011). Crystallization and preliminary diffraction analysis of the CAL PDZ domain in complex with a selective peptide inhibitor. Acta Crystallogr Sect F Struct Biol Cryst Commun *67*, 600-603.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

Feng, J., Westbrook, J., and Berman, H. (1998). PDB Validation Server. In Rutgers University, New Brunswick, NJ NDB-407.

Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Kim, T., Tyndel, M.S., Huang, H., Sidhu, S.S., Bader, G.D., Gfeller, D., and Kim, P.M. (2011). MUSI: an integrated system for identifying multiple specificity from very large peptide or nucleic acid data sets. Nucleic Acids Res *40*, e47.

Laskowski, R., McArthur, M., Moss, D., and Thornton, J. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. *265*, 283-291.

McCoy, A.J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr D Biol Crystallogr *63*, 32-41.

Vaguine, A.A., Richelle, J., and Wodak, S.J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. Acta Crystallogr D Biol Crystallogr *55*, 191-205.

Vouilleme, L., Cushing, P.R., Volkmer, R., Madden, D.R., and Boisguerin, P. (2010). Engineering peptide inhibitors to overcome PDZ binding promiscuity. Angew Chem Int Ed Engl *49*, 9912-9916.

Wolde, M., Fellows, A., Cheng, J., Kivenson, A., Coutermarsh, B., Talebian, L., Karlson, K., Piserchio, A., Mierke, D.F., Stanton, B.A.*, et al.* (2007). Targeting CAL as a negative regulator of DeltaF508-CFTR cell-surface expression: an RNA interference and structure-based mutagenetic approach. J Biol Chem *282*, 8099-8109.