Mechanism of Tethered Agonist-Mediated Signaling by Polycystin-1

Shristi Pawnikar[1], Brenda S. Magenheimer[2,4], Ericka Nevarez Munoz[3], Robin L. Maser[2,3,4,*] and Yinglong Miao[1,*]

[1]Center for Computational Biology and Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66047; [2]Departments of Biochemistry and Molecular Biology, [3]Clinical Laboratory Sciences, and [4]The Jared Grantham Kidney Institute, University of Kansas Medical Center, Kansas City, KS 66160.

***Correspondence emails:** rmaser@kumc.edu and miao@ku.edu

Abstract

Polycystin-1 (PC1) is an important atypical G-protein-coupled receptor (GPCR) with 11 transmembrane domains and its mutations account for 85% of cases of human autosomal dominant polycystic kidney disease (ADPKD). PC1 shares multiple features with Adhesion GPCRs, including a tethered agonist (TA) from the stalk region that is suggested to activate signaling of the PC1 membrane-embedded C-terminal fragment (CTF). However, the mechanism by which a TA can activate PC1 is not known. Here, we have combined mutagenesis and cellular functional experiments with Gaussian accelerated molecular dynamics (GaMD) simulations to investigate TA-mediated activation of the wild type, stalkless and ADPKD stalk mutants of the PC1 CTF. Our GaMD simulation findings were highly consistent with the signaling experimental data. Together they revealed an allosteric signaling pathway connecting the stalk, Tetragonal Opening for Polycystins (TOP) and pore loop in activation of the PC1 CTF. Key residue interactions predicted from the GaMD simulations were validated with newly designed mutation-signaling experiments. Our complementary experiments and simulations have provided critical mechanistic insights into the TA-mediated activation of PC1 CTF, which is expected to facilitate rational drug design for the PC1 protein.

Keywords: Autosomal dominant polycystic kidney disease (ADPKD), Polycystin-1, Tethered Agonist, Activation, Mutagenesis, Cellular Signaling, Gaussian accelerated molecular dynamics.

Significance Statement

Mutations of polycystin-1 (PC1) are the major cause (85% of cases) of autosomal dominant polycystic kidney disease (ADPKD), yet the physiological and disease-related functions of PC1 remain poorly understood. A significant portion of the mutations encode PC1 with defects in maturation or reduced function that may be amenable to functional rescue. In this work, we have combined highly complementary biochemical and cellular assay experiments and accelerated molecular simulations, which revealed a novel allosteric transduction pathway in activation of the PC1 CTF. Our findings shall facilitate future rational drug design efforts targeting the PC1 signaling function.

Introduction

Autosomal dominant polycystic kidney disease (ADPKD) is caused by mutations of the *PKD1* (85% of cases) or *PKD2* (15%) gene (1), which encode the proteins, polycystin-1 (PC1) and polycystin-2 (PC2), respectively. Together, PC1 and PC2 comprise a membrane receptor-ion channel signaling complex, involved in mechanosensing, signaling and intracellular [Ca^{2+}] regulation (2-4). The PC1-PC2 complex is required for maintaining cellular homeostasis and differentiation (5, 6). Recent advancements in cryo-EM and structural studies of membrane proteins have led to the first glimpse of the PC1-PC2 complex (PDB: 6A70) (7), a heterotetramer consisting of 1 PC1 and 3 PC2 protein subunits in which the last 2 transmembrane (TM) domains of each protein form the ion channel pore (8), surrounded by the TOP (Tetragonal Opening of Polycystins) domains from the largest extracellular loop of each subunit.

PC1 is a large protein composed of 11 transmembrane (TM) domains and a short C-terminal cytosolic tail (C-tail) (9, 10). Its N-terminal, extracellular region consists of multiple binding domains that mediate cell-cell and cell-matrix interactions, Wnt binding and PC1-PC2 channel activation (9, 11-15). The C-tail initiates a number of signaling cascades (1), including the binding and activation of heterotrimeric G proteins, suggesting that PC1 functions an atypical type of G protein-coupled receptor (GPCR) (16-23). Recently, the regulation of G protein signaling by PC1 was shown to be critical for the prevention of renal cyst formation in mice (24, 25). However, we know very little about the molecular mechanism involved.

PC1 and the Adhesion family of GPCRs (ADGRs) share a conserved **G** protein-coupled receptor **p**roteolysis **s**ite (GPS) motif located just proximal to their first TM

domain that undergoes an intramolecular, self-catalyzed cleavage resulting in an N-terminal, extracellular fragment (**NTF**) and a C-terminal, membrane-embedded fragment (**CTF**) that remain non-covalently attached (26-31). GPS cleavage is known to be important for PC1 function, maturation and trafficking (27, 30, 32-35), and essential for preventing renal cystogenesis (32, 36). For the ADGRs, GPS cleavage and the association-dissociation of the NTF and CTF subunits are thought to be involved in regulation of G protein signaling (37-40). In the tethered agonist (TA) model of ADGRs, dissociation of the NTF results in exposure of a short, extracellular 'stalk' at the N-terminus of the CTF subunit which binds the extracellular loops and 7-TM bundle causing conformational changes that activate G protein signaling (41-44). We recently demonstrated that PC1 utilizes an ADGR-like TA for constitutive activation of an NFAT promoter-luciferase reporter (45). This work showed that the CTF subunit alone has greater signaling ability than full-length PC1, is dependent on the presence of the stalk and is affected by ADPKD-associated missense mutations within the stalk region. Furthermore, synthetic peptides derived from the PC1 stalk sequence could stimulate a stalkless CTF mutant to activate the NFAT reporter.

In the current study, we have combined highly complementary biochemical and cellular functional experiments and Gaussian accelerated molecular dynamics (GaMD) simulations to understand the dynamic structure-function relationships underlying the TA-mediated signaling of PC1 CTF. GaMD is an enhanced sampling computational technique that works by adding a harmonic boost potential to smooth the potential energy surface (46). GaMD greatly reduces energy barriers and accelerates biomolecular simulations by orders of magnitude. GaMD does not require pre-defined collective variables or reaction

coordinates. Compared with the enhanced sampling methods that rely on careful selection of the collective variables, GaMD is of particular advantage for studying complex biological processes such as activation of the PC1 CTF. Moreover, because the boost potential exhibits a near-Gaussian distribution, free energy profiles can be properly recovered through cumulant expansion to the second order (46). Applications of GaMD have revealed mechanisms of protein folding and conformational changes, ligand binding, protein-protein/membrane/nucleic acid interactions and carbohydrate dynamics (47).

Here, a computational model of PC1 CTF was generated by extracting the protein from cryo-EM structure of the PC1-PC2 complex (PDB: 6A70) (7). Homology modeling was applied using I-TASSER (48) to add the missing regions in PC1 CTF, including the stalk, pore loop (PL) and C-tail (**Figs. 1A** and **S1**; see details in **SI Methods**). Parallel *in vitro* cellular signaling experiments and GaMD simulations were carried out on the wild type (WT), stalkless and ADPKD-associated stalk mutants of the PC1 CTF. Most of the PC1 CTF stalk variants exhibited reduced ability in activation of the intracellular NFAT reporter compared with the WT CTF. All-atom GaMD simulations revealed highly correlated residue motions and interactions between the stalk-TOP-PL domains, which were important for activation of the PC1 CTF. A number of simulation-predicted residue interactions in the PC1 CTF were further validated by mutagenesis and cellular signaling experiments. The combined experiments and simulations have thus provided the first mechanistic insights into TA-mediated signaling by the PC1 CTF.

Results

Stalk variants of the PC1 CTF exhibited reduced signaling activity in cellular assays

We first compared the ability of different expression constructs of human PC1 CTF to activate signaling to an NFAT reporter in transiently transfected HEK293T cells (**Fig. 1**). CTF expression constructs included the WT, stalkless (ΔStalk, with deletion of the first 21 N-terminal residues of CTF) and 3 missense mutants within the stalk region (G3052R, R3063C, R3063P) that are associated with ADPKD (https://pkdb.mayo.edu) (**Fig. 1B**). Semi-quantitative Western blot analyses showed that the deletion (ΔStalk) and ADPKD missense mutations of the stalk region had minimal effect (≤ 20%) on expression levels of the PC1 CTF (**Fig. 1C-1D**). However, the mutants exhibited distinct levels of activity in the NFAT reporter assay (**Fig. 1E**). WT CTF was able to activate the NFAT reporter by 5.09-fold relative to the empty expression vector (ev) control. In comparison, the ΔStalk, G3052R and R3063C variants of CTF activated the NFAT reporter only by 2.01-, 2.46- and 2.18-fold relative to ev control, respectively. These three stalk variations thus significantly reduced activity of the PC1 CTF in the NFAT reporter activation by ~50-60% compared with the WT. Finally, the R3063P mutation resulted in only a minor reduction in NFAT reporter activation (4.71-fold relative to ev) compared to the WT CTF, which was not statistically significant ($P = 0.577$). We hypothesized that reductions observed in NFAT reporter activation of the ΔStalk, G3052R and R3063C variants of CTF were likely due to alterations in structural dynamics of the protein.

GaMD simulations revealed reduced correlations in residue motions and disrupted domain interactions in the stalk variants of PC1 CTF compared with the wild type

In parallel with the mutagenesis and cellular assay experiments, we performed all-atom GaMD simulations on the WT and the ΔStalk, G3052R and R3063C variants of PC1 CTF

(**SI Methods**). Three independent 1000 ns GaMD production trajectories were obtained on each system. The correlation matrix of residue motions was first calculated from the GaMD trajectories. In the WT CTF, high correlations were observed between the Stalk-TOP and TOP-PL domains (**Fig. 2A**). Notably, residues 3049-3074 in Stalk and 3699-3732 in TOP exhibited correlations > 0.5, and similarly for residues 3714-3880 in TOP and 4051-4080 in PL of the WT CTF. In comparison, correlations between residue motions in the stalk-TOP and TOP-PL domains were significantly reduced in the ΔStalk (**Fig. 2B**), G3052R (**Fig. 2C**) and R3063C (**Fig. 2D**) systems of the PC1 CTF compared with the WT (**Fig. 2A**). Residue correlations in the Stalk and TOP domains of the stalk variants of PC1 CTF showed values in the range of 0-0.25, being much lower than those in the WT (i.e., 0.25-0.75). Similarly, residue correlations in the TOP and PL domains of the ΔStalk and the ADPKD mutant systems also showed lower values of 0-0.35, as compared with those of the WT (0.25-0.75). The loss of correlated residue motions in the Stalk-TOP and TOP-PL domains observed in the stalk variants during GaMD simulations suggested that these domains were likely crucial for TA-mediated signaling in PC1 CTF, and therefore, were further investigated.

During the GaMD simulations, a salt bridge was frequently formed between residues R3848-E4078 in the WT CTF, but rarely observed in the ΔStalk and completely absent in the ADPKD mutant systems as shown in **Fig. S2**. Therefore, the R3848-E4078 ionic interaction appeared to play an important role in facilitating stalk-mediated signaling of the WT CTF. In addition, a hydrogen bond was formed between residue N3074 at the base of the stalk region and S3585 near the extracellular end of TM5 in two of the three

GaMD simulations of WT CTF, but was intermittently observed in the stalk variant mutants (**Fig. S3**).

Next, we calculated 2D free energy profiles of the WT and stalk variants of PC1 CTF by reweighting of the GaMD simulations (**Fig. 2E-2H**). The R3848-E4078 residue distance and the number of contacts between the Stalk and TOP domains were selected as the reaction coordinates. Two low-energy conformational states were identified from the free energy profile of the WT CTF and designated as "Closed" and "Intermediate I1" (**Fig. 2E**). In the WT CTF, the stalk region was found to interact considerably with the TOP domain, for which their number of atom contacts ranged ~40-120 (**Figs. 2E** and **S4A**). In the "Closed" state, a salt bridge was formed between R3848-E4078 with a distance of ~3.9 Å between the residue charge centers (the CZ atom in R3848 and the CD atom in E4078). The "Intermediate I1" state showed an increased number of ~60-120 contacts between Stalk-TOP domains, however, the R3848-E4078 salt bridge became broken at ~8 Å distance between the residue charge centers (**Fig. 2E**). A distinct "Open" low-energy conformational state was identified from the free energy profiles of the ΔStalk, G3052R and R3063C systems, for which the R3848-E4078 salt bridge became broken (**Fig. 2F-2H**). At the "Open" energy minimum, the number of contacts between the Stalk-TOP domains centered around 0, ~50 and ~30 in the ΔStalk (**Fig. 2F**), G3052R (**Fig. 2G**) and R3063C (**Fig. 2H**) systems, respectively.

Distinct low-energy conformations of the wild type and stalk variants of PC1 CTF

In the "Closed" low-energy conformation identified for the WT CTF, a number of polar interactions were identified between the Stalk and TOP domains, including a hydrogen

bond between N3074-S3585 (**Figs. 3B** and **S3A**), T3049-E3708 (**Figs. 3A** and **S5A**) and a salt bridge between E3068-R3700 (**Figs. 3A** and **S6A**). The distance between the OG1 atom of T3049 and the CD atom of E3708 was calculated as ~3.6 Å, and ~4.7 Å distance between the charge centers of E3068 (the CD atom) and R3700 (the CZ atom) (**Fig. 3A**). Note that the T3049-E3078 hydrogen bonding interaction was present in only the WT CTF (**Fig. S5**), while the E3068-R3700 salt bridge was formed in both the WT and R3063P mutant CTF (**Fig. S6**). The N3074-S3585 hydrogen bond was observed to form very stably in the WT CTF system, however, was intermittently present in the CTF mutants. The free energy minimum distance between the CG atom in N3074 and the OG atom in S3585 was found to be ~3.9 Å (**Fig. 3B**). Furthermore, a salt bridge was formed between R3848-E4078 from TOP and PL domains with a distance of ~3.9 Å between the residue charge centers (**Figs. 3C** and **S2A**).

In the "Open" low-energy conformational state identified from free energy profiles of the ΔStalk, G3052R and R3063C PC1 CTF, residue D3072 from the C-terminal end of the Stalk region formed a salt bridge with R3891 in the TOP domain (**Figs. 3D** and **S7**). The D3072-R3891 salt bridge interaction was formed for more extended periods of time during GaMD simulations of these three systems as compared to the WT CTF and R3063P mutant. However, the N3074-S3585 hydrogen bond became broken with ~7.7 Å distance between the residue charge centers (CG atom in N3074 and OG atom in S3585) (**Fig. 3E**). The salt bridge interaction connecting residues R3848-E4078 from TOP and PL domains also became broken with a large distance of 23.1 Å between the residue charge centers (**Fig. 3F**), although fluctuations were observed during the GaMD simulations (**Fig. S2B-S2D**).

Intermediate conformational states of PC1 CTF observed from GaMD simulations of the R3063P mutant

In addition, three independent 1000 ns GaMD simulations were performed on the R3063P ADPKD missense mutant, which showed only a minor reduction in NFAT reporter activation compared to WT PC1 CTF (**Fig. 1E**). Correlations between residues motions in the stalk-TOP and TOP-PL domains were slightly reduced in the R3063P mutant compared with the WT (**Fig. 4A**). Residue motions in the Stalk and TOP domains of the R3063P PC1 CTF system showed correlation values ranging between 0.25-0.5 as compared to correlation values of 0.25-0.75 for the WT system. Similarly, residue motions in the TOP and PL domains of the R3063P PC1 CTF system ranged between 0.25-0.5, which were slightly lower than those of the WT (0.25-0.75).

A 2D free energy profile was further calculated for the R3063P mutant of PC1 CTF regarding the number of contacts between Stalk-TOP and R3848-E4078 residue distance (**Fig. 4B**). Two low-energy conformational states were identified from the free energy profile, and designated "Intermediate I1" and "Intermediate I2". Note that the "Intermediate I1" state was observed in both WT (**Fig. 2E**) and the R3063P mutant (**Fig. 4B**) of PC1 CTF. Overall, the R3848-E4078 salt bridge became broken in these two Intermediate conformational states, but the residue distance was shorter than in the "Open" state observed with the ΔStalk, G3052R and R3063P mutant systems (Fig. 2F-H). The "Intermediate I1" state showed increased Stalk-TOP number of contacts ranging between ~90-130 (**Fig. 4B-4C**), and the distance between charge centers of residues R3848 (the CZ atom) and E4078 (the CD atom) centered around 7-9 Å (**Fig. 4B** and **4D**). The

"Intermediate I2" state, however, showed a reduced number of Stalk-TOP contacts ranging between ~40-80 (**Fig. 4B**) and the R3848-E4078 distance increased to ~15.4 Å (**Fig. 4B** and **4E**).

Experimental validation of simulation-predicted residue interactions that are important for WT stalk-mediated activation of PC1 CTF

To determine whether newly discovered residue interactions in the GaMD simulations, i.e., the N3074-S3585 hydrogen bond and R3848-E4078 salt bridge, are important for signaling of the WT CTF, these residues were mutated, one at a time, in the CTF expression construct and assessed for signaling to the NFAT reporter (**Fig. 5A**). Activation of the NFAT reporter was significantly reduced from 6.10-fold for WT CTF to 2.00-fold ($P < 0.0001$) for N3074A and 3.69-fold ($P = 0.0066$) for S3585A CTF mutants. Likewise, the ability of activating the NFAT reporter was significantly reduced for the R3848E mutant (2.15-fold, $P = 0.0001$) with respect to WT CTF, although reporter activation was only slightly reduced for the E4078R mutant (5.57-fold, $P = 0.6021$). To determine if changes in NFAT reporter activation by any of the mutants was due to loss of expression or inability to traffic, both the total (**Fig. 5B-5C**) and cell surface (**Fig. 5D-5E**) expression levels were investigated. Semi-quantitative Western blot analyses show total expression levels of the CTF mutants were reduced to 0.79 for N3074A, 0.71 for S3585A and R3848E, and 0.51 for E4078R of WT CTF level, while cell surface expression levels were increased for CTF-N3074A and -S3585A (1.20- and 1.41-fold, respectively) and decreased for R3848E and E4078R (0.51 and 0.71, respectively) relative to WT CTF.

Discussion

In this study, we have combined highly complementary site-directed mutagenesis and cellular experimental assays with all-atom GaMD simulations to decipher the TA-mediated activation mechanism of the PC1 CTF. The ΔStalk, G3052R and R3063C variants significantly reduced activation of the NFAT reporter compared with the WT CTF, unlike the R3063P mutant. Correlation matrices calculated from GaMD simulations of the WT and mutant CTF systems revealed alterations in residue motions between the stalk and TOP and the TOP and PL domains that were consistent with the results of the cellular signaling assays. Further analyses of the GaMD simulations revealed an allosteric transduction pathway connecting the stalk-TOP-PL domains, which was important for activation of the WT CTF in the "Closed" conformational state (**Fig. 6A**). This pathway and interactions between the stalk-TOP-PL domains were disrupted in the ΔStalk, G3052R and R3063C systems, leading to a distinct "Open" conformational state in the calculated free energy profiles (**Fig. 6B**). This was in contrast to the "Closed" and "Intermediate" low-energy states of the WT and R3063P PC1 CTF. Newly identified residue interactions, N3074-S3585 and R3848-E4078, that were predicted to be important in the GaMD simulations were further validated by mutagenesis and cellular signaling experiments. Therefore, our experiments and simulations combined have provided important mechanistic insights into the TA-mediated activation of the PC1 CTF signaling activity.

GaMD simulations suggested a large number of contacts formed between the Stalk region and TOP domain in the WT CTF. Removal of the stalk region and single residue mutations G3052R and R3063C decreased the number of Stalk-TOP contacts and disrupted the R3848-E4078 salt bridge interaction between the TOP-PL domains. These observations

suggested an allosteric mechanism, by which the stalk interaction with the TOP domain could trigger signal activation being translated via the R3848-E4078 salt bridge interaction between the TOP-PL domains and ultimately to the C-tail of PC1, which contains the G protein activation motif (49).

GaMD simulations also revealed a hydrogen bond interaction between N3074 located at the base of the Stalk and S3585 in the extracellular domain of TM5 in WT CTF, which was intermittently present in the CTF mutants. The importance of this interaction in stalk-mediated signaling of PC1 CTF was experimentally validated. Substitution of either residue with Ala to prevent their polar interaction resulted in a significant reduction in activation of the NFAT reporter. Together with the apparent weakening or loss of this interaction in the GaMD simulations of stalk variants, these results suggested that the N3074-S3585 hydrogen bond was likely important in anchoring the base of the Stalk for interactions with the TOP domain. Such a proposed role for the N3074-S3585 interaction could be replaced by residue-residue contacts between D3072-R3891 that are more prominent in the ΔStalk, G3052R and R3063C mutants compared to WT CTF and the R3063P mutant (**Figs. 6** and **S7**). Furthermore, retention of the N3074-S3585 and E3068-R3700 interactions by the R3063P mutant were able to compensate for its stalk residue substitution and promote a high number of functional Stalk-TOP contacts, thus explaining its only slightly reduced ability in activating the NFAT reporter. Relevant to these findings, it was intriguing that both residues N3074 and E3068 were reported to be sites of ADPKD-associated mutations (PKD mutation database, https://pkdb.mayo.edu) and were recently shown to alter signaling to the NFAT reporter when mutated (45).

The importance of the R3848-E4078 salt bridge interaction between the TOP-PL domains was also experimentally tested through mutation-signaling experiments. The salt bridge was disrupted by mutating either residue to an oppositely charged residue (R3848E and E4078R), which was expected to cause repulsion between the TOP domain and PL. While the R3848E mutant showed significantly decreased activation of the NFAT reporter as compared to the WT, the signaling ability of the E4078R CTF mutant was only slightly decreased (**Fig. 5A**). Since the total and cell surface expression levels of CTF-E4078R were lower than WT CTF, one might argue that this mutant's activity relative to its expression level was actually greater than WT CTF. In this regard, residue mutations that could even increase signaling capability of ADGRs have been identified (50, 51). A lack of correlation between surface expression levels and signaling activity of ADGRs has been also reported for GPR133/ADGRD1 (52). A possible explanation for the apparent lack of effect of the E4078R substitution in the PL on NFAT reporter activation could be an ability of the E4078R mutant to re-establish a fortuitous interaction with one of the many polar residues in the TOP domain, which maintained allosteric signaling of the Stalk-TOP interactions to the PL. On the other hand, the PL does not include any basic residues that would enable mutation of residue R3848 in the TOP domain to negatively charged Glu to interact, thus disrupting its interaction with the PL and resulting in the significantly reduced activity of the R3848E mutant. Regardless, these findings further support a model where the TOP-PL interactions are involved in stalk-/TA-mediated activation of signaling in PC1 CTF.

Shared features between PC1 and the ADGRs, including auto-proteolytic cleavage at the GPS motif and homology of the GAIN domain, led to the proposal and investigation

of whether a TA is involved in PC1 CTF-mediated activation of an NFAT luciferase reporter (45). The current study reveals that the PC1 CTF utilizes a novel allosteric activation mechanism for its TA-mediated signaling activation. To date, neither the molecular mechanism of TA-mediated activation nor the binding pocket of the TA have been established for ADGRs, although a number of observations are supportive of an allosteric activation mechanism (53, 54). The involvement of extracellular loops and ends of TM domains in either binding of a tethered ligand or contributing to activation of signaling has been reported for a number of ADGRs including Gpr56/ADGRG1, Lphn1/ ADGRL1, GPR133/ADGRD1, and GPR64/ADGRG2 (51, 52, 55, 56). Recently, the binding pocket for the tethered ligand of protease-activated receptor 4 (PAR4) was also determined to involve residues in the TM3 and TM7 and ligand binding was regulated by extracellular loop 3 (57).

In summary, we have revealed important Stalk-TOP and TOP-PL domain interactions, which played a significant role in the activation mechanism of the PC1 CTF. Our complementary mutagenesis and cellular signaling experiments and GaMD simulations on both the WT and signaling-deficient mutant CTF proteins have provided important insights into the mechanism of PC1 activation at an atomic level, which are expected to facilitate future rational drug design efforts targeting the PC1 signaling protein.

Materials and Methods

We have performed transient transfections of HEK293T cells with WT and mutant CTF expression constructs of PC1 to assess their ability to activate signaling to an NFAT promoter luciferase reporter following a previously established protocol (58). PC1 CTF

mutants, including the ΔStalk (a 21-residue deletion of the CTF stalk region), single residue substitutions (the ADPKD-associated mutations G3052R, R3063C and R3063P) and simulation suggested mutations N3074A, S3585A, R3848E, and E4078R, were generated by site directed mutagenesis. To evaluate the potential effect of mutations in the PC1 CTF, total and cell surface expression levels were determined from total cell lysates and surface biotinylated proteins, respectively, from the transfected cells using SDS-PAGE and Western blot analysis. GaMD simulations were performed on the WT, ΔStalk, G3052R, R3063C and R3063P systems of the PC1 CTF, following a previously established protocol (59). The CHARMM36m parameter set (60) was used for the protein and lipids. NAMD2.12 (61) was used for initial energy minimization, thermalization and 20 ns conventional MD equilibration. The output files of cMD simulations using NAMD were converted to AMBER format using *ParmEd* tool AMBER package (62). GaMD simulations implemented in GPU version of AMBER 18 (46) were performed for the PC1 CTF and its mutant systems. The GaMD simulations involved a short conventional MD of 12 ns, 48 ns GaMD equilibration and finally three independent 1 μs GaMD production runs with randomized initial atomic velocities for each system. Details of mutagenesis, cellular signaling assays, surface biotinylation, Western blot analysis, GaMD simulations, system setup, and simulation analysis are provided in **SI**.

Acknowledgments

This work used supercomputing resources with allocation award TG-MCB180049 through the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562, and project M2874

through the National Energy Research Scientific Computing Center (NERSC), which is a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231, and the Research Computing Cluster at the University of Kansas. This work was supported in part by National Institutes of Health (DK123590), Department of Defense CDMRP PRMRP Discovery Award (PR160710/W81XWH-17-1-0301), and Pilot Grant funding from the School of Health Professions at KU Medical Center to R.L.M. and the startup funding in the College of Liberal Arts and Sciences at the University of Kansas (to Y.M.).

Author Contributions

R.M. and Y.M. designed research; S.P., E. N. M., and B. S. M. performed research; S.P., B. S. M., R.M. and Y.M. analyzed data; and S.P., R.M. and Y.M. wrote the paper.

Competing Interest Statement

No competing interests.

References

1. P. C. Harris, V. E. Torres, Genetic mechanisms and signaling pathways in autosomal dominant polycystic kidney disease. *The Journal of clinical investigation* **124**, 2315-2324 (2014).
2. S. M. Nauli *et al.*, Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. *Nature genetics* **33**, 129-137 (2003).
3. S. M. Nauli *et al.*, Endothelial cilia are fluid shear sensors that regulate calcium signaling and nitric oxide production through polycystin-1. *Circulation* **117**, 1161-1171 (2008).
4. Z. Pedrozo *et al.*, Polycystin-1 Is a Cardiomyocyte Mechanosensor That Governs L-Type Ca2+ Channel Protein Stability. *Circulation* **131**, 2131-2142 (2015).
5. A. Boletta, G. Germino, Role of polycystins in renal tubulogenesis. *Trends Cell Biol* **13**, 484-492 (2003).
6. C. Piperi, E. K. Basdra, Polycystins and mechanotransduction: From physiology to disease. *World journal of experimental medicine* **5**, 200-205 (2015).
7. Q. Su *et al.*, Structure of the human PKD1-PKD2 complex. *Science* **361** (2018).
8. Z. Wang *et al.*, The ion channel function of polycystin-1 in the polycystin-1/polycystin-2 complex. *EMBO reports* **20**, e48336 (2019).
9. R. Sandford *et al.*, Comparative analysis of the polycystic kidney disease 1 (PKD1) gene reveals an integral membrane glycoprotein with multiple evolutionary conserved domains. *Human molecular genetics* **6**, 1483-1489 (1997).
10. N. Nims, D. Vassmer, R. L. Maser, Transmembrane domain analysis of polycystin-1, the product of the polycystic kidney disease-1 (PKD1) gene: evidence for 11 membrane-spanning domains. *Biochemistry* **42**, 13035-13048 (2003).
11. S. Kim *et al.*, The polycystin complex mediates Wnt/Ca signalling. *Nature cell biology* 10.1038/ncb3363 (2016).
12. B. S. Weston, C. Bagneris, R. G. Price, J. L. Stirling, The polycystin-1 C-type lectin domain binds carbohydrate in a calcium-dependent manner, and interacts with extracellular matrix proteins in vitro. *Biochimica et biophysica acta* **1536**, 161-176 (2001).
13. O. Ibraghimov-Beskrovnaya *et al.*, Strong homophilic interactions of the Ig-like domains of polycystin-1, the protein product of an autosomal dominant polycystic kidney disease gene, PKD1. *Human molecular genetics* **9**, 1641-1649 (2000).
14. B. S. Weston, A. N. Malhas, R. G. Price, Structure-function relationships of the extracellular domain of the autosomal dominant polycystic kidney disease-associated protein, polycystin-1. *FEBS letters* **538**, 8-13 (2003).
15. K. Ha *et al.*, The heteromeric PC-1/PC-2 polycystin complex is activated by the PC-1 N-terminus. *Elife* **9**, e60684 (2020).
16. S. C. Parnell *et al.*, The polycystic kidney disease-1 protein, polycystin-1, binds and activates heterotrimeric G-proteins in vitro. *Biochemical and biophysical research communications* **251**, 625-631 (1998).
17. T. Yuasa, A. Takakura, B. M. Denker, B. Venugopal, J. Zhou, Polycystin-1L2 is a novel G-protein-binding protein. *Genomics* **84**, 126-138 (2004).

18. P. Delmas *et al.*, Constitutive activation of G-proteins by polycystin-1 is antagonized by polycystin-2. *The Journal of biological chemistry* **277**, 11276-11283 (2002).
19. S. C. Parnell *et al.*, Polycystin-1 activation of c-Jun N-terminal kinase and AP-1 is mediated by heterotrimeric G proteins. *The Journal of biological chemistry* **277**, 19566-19572 (2002).
20. S. Puri *et al.*, Polycystin-1 activates the calcineurin/NFAT (nuclear factor of activated T-cells) signaling pathway. *The Journal of biological chemistry* **279**, 55455-55464 (2004).
21. P. Delmas *et al.*, Gating of the polycystin ion channel signaling complex in neurons and kidney cells. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **18**, 740-742 (2004).
22. Y. Wu *et al.*, Galpha12 is required for renal cystogenesis induced by Pkd1 inactivation. *Journal of cell science* **129**, 3675-3684 (2016).
23. T. Hama, F. Park, Heterotrimeric G protein signaling in polycystic kidney disease. *Physiological genomics* **48**, 429-445 (2016).
24. B. Zhang, U. Tran, O. Wessely, Polycystin 1 loss of function is directly linked to an imbalance in G-protein signaling in the kidney. *Development (Cambridge, England)* **145** (2018).
25. S. C. Parnell *et al.*, A mutation affecting polycystin-1 mediated heterotrimeric G-protein signaling causes PKD. *Human molecular genetics* **27**, 3313-3324 (2018).
26. C. P. Ponting, K. Hofmann, P. Bork, A latrophilin/CL-1-like GPS domain in polycystin-1. *Current biology : CB* **9**, R585-588 (1999).
27. F. Qian *et al.*, Cleavage of polycystin-1 requires the receptor for egg jelly domain and is disrupted by human autosomal-dominant polycystic kidney disease 1-associated mutations. *Proc Natl Acad Sci U S A* **99**, 16981-16986 (2002).
28. A. Kurbegovic *et al.*, Novel functional complexity of polycystin-1 by GPS cleavage in vivo: role in polycystic kidney disease. *Molecular and cellular biology* **34**, 3341-3353 (2014).
29. R. L. Maser, J. P. Calvet, Adhesion GPCRs as a paradigm for understanding polycystin-1 G protein regulation. *Cellular signalling* **72**, 109637 (2020).
30. D. Arac *et al.*, A novel evolutionarily conserved domain of cell-adhesion GPCRs mediates autoproteolysis. *The EMBO journal* **31**, 1364-1378 (2012).
31. S. Promel, T. Langenhan, D. Arac, Matching structure with function: the GAIN domain of adhesion-GPCR and PKD1-like proteins. *Trends in pharmacological sciences* **34**, 470-478 (2013).
32. Y. Cai *et al.*, Altered trafficking and stability of polycystins underlie polycystic kidney disease. *J Clin Invest* **124**, 5129-5144 (2014).
33. V. G. Gainullin, K. Hopp, C. J. Ward, C. J. Hommerding, P. C. Harris, Polycystin-1 maturation requires polycystin-2 in a dose-dependent manner. *The Journal of clinical investigation* **125**, 607-620 (2015).
34. X. Su *et al.*, Regulation of polycystin-1 ciliary trafficking by motifs at its C-terminus and polycystin-2 but not by cleavage at the GPS site. *Journal of cell science* **128**, 4063-4073 (2015).

35. M. Trudel, Q. Yao, F. Qian, The Role of G-Protein-Coupled Receptor Proteolysis Site Cleavage of Polycystin-1 in Renal Physiology and Polycystic Kidney Disease. *Cells* **5** (2016).
36. S. Yu *et al.*, Essential role of cleavage of Polycystin-1 at G protein-coupled receptor proteolytic site for kidney tubular structure. *Proc Natl Acad Sci U S A* **104**, 18688-18693 (2007).
37. K. J. Paavola, J. R. Stephenson, S. L. Ritter, S. P. Alter, R. A. Hall, The N terminus of the adhesion G protein-coupled receptor GPR56 controls receptor signaling activity. *The Journal of biological chemistry* **286**, 28914-28921 (2011).
38. K. J. Paavola, R. A. Hall, Adhesion G protein-coupled receptors: signaling, pharmacology, and mechanisms of activation. *Molecular pharmacology* **82**, 777-783 (2012).
39. S. Promel *et al.*, The GPS motif is a molecular switch for bimodal activities of adhesion class G protein-coupled receptors. *Cell reports* **2**, 321-331 (2012).
40. S. C. Petersen *et al.*, The adhesion GPCR GPR126 has distinct, domain-dependent functions in Schwann cell development mediated by interaction with laminin-211. *Neuron* **85**, 755-769 (2015).
41. I. Liebscher *et al.*, A tethered agonist within the ectodomain activates the adhesion G protein-coupled receptors GPR126 and GPR133. *Cell reports* **9**, 2018-2026 (2014).
42. T. Schoneberg, I. Liebscher, R. Luo, K. R. Monk, X. Piao, Tethered agonists: a new mechanism underlying adhesion G protein-coupled receptor activation. *Journal of receptor and signal transduction research* **35**, 220-223 (2015).
43. L. M. Demberg, S. Rothemund, T. Schoneberg, I. Liebscher, Identification of the tethered peptide agonist of the adhesion G protein-coupled receptor GPR64/ADGRG2. *Biochemical and biophysical research communications* **464**, 743-747 (2015).
44. H. M. Stoveken, A. G. Hajduczok, L. Xu, G. G. Tall, Adhesion G protein-coupled receptors are activated by exposure of a cryptic tethered agonist. *Proc Natl Acad Sci* **112**, 6194-6199 (2015).
45. E. N. M. Brenda S. Magenheimer, Jayalakshmi Ravichandran, and Robin L. Maser (Constitutive signaling by the C-terminal fragment (CTF) of polycystin-1 is mediated by a tethered peptide agonist.
46. Y. Miao, V. A. Feher, J. A. McCammon, Gaussian accelerated molecular dynamics: Unconstrained enhanced sampling and free energy calculation. *Journal of chemical theory and computation* **11**, 3584-3595 (2015).
47. J. Wang *et al.*, Gaussian accelerated molecular dynamics: principles and applications. *WIREs Computational Molecular Science* 10.1002/wcms.1521, e1521 (2021).
48. A. Roy, A. Kucukural, Y. Zhang, I-TASSER: a unified platform for automated protein structure and function prediction. *Nature protocols* **5**, 725-738 (2010).
49. S. C. Parnell *et al.*, The Polycystic Kidney Disease-1 Protein, Polycystin-1, Binds and Activates Heterotrimeric G-Proteinsin Vitro. *Biochemical and biophysical research communications* **251**, 625-631 (1998).

50. R. H. Purcell, C. Toro, W. A. Gahl, R. A. Hall, A disease-associated mutation in the adhesion GPCR BAI2 (ADGRB2) increases receptor signaling activity. *Hum Mutat* **38**, 1751-1760 (2017).
51. O. Nazarko *et al.*, A Comprehensive Mutagenesis Screen of the Adhesion GPCR Latrophilin-1/ADGRL1. *iScience* **3**, 264-278 (2018).
52. L. Fischer, C. Wilde, T. Schöneberg, I. Liebscher, Functional relevance of naturally occurring mutations in adhesion G protein-coupled receptor ADGRD1 (GPR133). *BMC genomics* **17**, 1-9 (2016).
53. A. Vizurraga, R. Adhikari, J. Yeung, M. Yu, G. G. Tall, Mechanisms of adhesion G protein–coupled receptor activation. *Journal of Biological Chemistry* **295**, 14065-14083 (2020).
54. Y.-Q. Ping *et al.*, Structures of the glucocorticoid-bound adhesion receptor GPR97–G o complex. *Nature* **589**, 620-626 (2021).
55. A. Kishore, R. A. Hall, Disease-associated extracellular loop mutations in the adhesion G protein-coupled receptor G1 (ADGRG1; GPR56) differentially regulate downstream signaling. *Journal of Biological Chemistry* **292**, 9711-9720 (2017).
56. Y. Sun *et al.*, Optimization of a peptide ligand for the adhesion GPCR ADGRG2 provides a potent tool to explore receptor biology. *Journal of Biological Chemistry* **296**, 100174 (2021).
57. X. Han *et al.*, PAR4 activation involves extracellular loop 3 and transmembrane residue Thr153. *Blood* **136**, 2217-2228 (2020).
58. R. L. Maser, B. S. Magenheimer, C. A. Zien, J. P. Calvet, "Transient transfection assays for analysis of signal transduction in renal cells" in Renal Disease. (Springer, 2003), pp. 205-217.
59. Y. Miao, J. A. McCammon, Graded activation and free energy landscapes of a muscarinic G-protein–coupled receptor. *Proceedings of the National Academy of Sciences* **113**, 12162-12167 (2016).
60. J. Huang *et al.*, CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature methods* **14**, 71-73 (2017).
61. J. C. Phillips *et al.*, Scalable molecular dynamics with NAMD. *Journal of computational chemistry* **26**, 1781-1802 (2005).
62. D. A. Case *et al.*, AMBER 18, University of California, San Francisco. (2018).

Figure Captions

Figure 1. Atomic structure of PC1 CTF and experimental effects of stalk region variants on signaling to NFAT reporter. (A) Atomic structure of the PC1 CTF as extracted from the cryo-EM structure of PC1-PC2 complex (PDB: 6A70). The missing regions added through homology modeling, including Stalk, Pore Loop (PL) and C-tail, are highlighted in red. The TOP and PLAT domains are shown in green and pink surfaces, respectively. (B) Sequence alignment of the stalk region in wild type (WT), stalkless (ΔStalk), and the G3052R, R3063C and R3063P ADPKD missense mutants of human PC1 CTF and structural view of PC1 stalk variants with the mutated residues highlighted as sticks. The sequence alignment shows residue numbers (superscripted), GPS cleavage site (arrow), and ADPKD-associated missense mutations/polymorphisms (red). (C) Representative western blot of WT and mutant CTF proteins from one of the experiments summarized in (D). Blot was originally probed with C20 antibody and re-probed with antibody against actin. (D) Relative, average total expression level (+SD) for WT and mutant CTF constructs from signaling transfections represented in (E). (E) Average fold NFAT reporter activation (+SD) relative to empty vector (ev) for WT and mutant CTF proteins from 3-5 separate transfection experiments with n=3 wells/expression construct/experiment. *, $p<0.005$; **, $p\leq0.0001$ relative to WT CTF levels.

Figure 2. Reduced correlations in residue motions and disrupted domain interactions in the stalk variants of PC1 CTF compared with the wild type. (A) The correlation matrix of residue motions calculated from GaMD simulations of the WT PC1 CTF. High correlations between Stalk-TOP and TOP-PL domains are highlighted in red circles. (B-D) The correlation matrix of residue motions (lower triangle) and the matrix of correlation

differences from the WT (upper triangle) for the (B) ΔStalk, (C) G3052R and (D) R3063C systems of PC1 CTF. Regions with significant differences are highlighted in blue circles. (E-H) 2D free energy profiles of the (E) WT, (**F**) ΔStalk, (G) G3052R and (H) R3063C systems of PC1 CTF regarding the number of atom contacts between the Stalk and TOP domains and the R3848-E4078 distance (the CZ atom in R3848 and the CD atom in E4078) calculated from the GaMD simulations. Residue contacts were calculated for heavy atoms with a distance cutoff of 4 Å. Important low-energy conformational states are identified, including the "Closed", "Intermediate I1" and "Open".

Figure 3. Distinct low-energy conformations of the WT and stalk variants of PC1 CTF. (A-C) Residue interactions observed in "Closed" protein conformation between (A) Stalk-TOP, (B) Stalk-TM5 and (C) TOP-PL domains. (D-F) Residue interactions observed in "Open" conformation of the Stalk variants between (D) Stalk-TOP, (E) Stalk-TM5 and (F) TOP-PL domains. The Stalk (orange cartoon), transmembrane helices (blue cartoon), TOP domain (green surface) and pore loop (magenta cartoon) are labeled in the PC1 CTF. Residue interactions are represented in ball and stick.

Figure 4. Intermediate conformational states of PC1 CTF observed from GaMD simulations of the R3063P mutant. (A) The correlation matrix of residue motions (lower triangle) and the difference matrix from WT (upper triangle) for the R3063P mutant. (B) 2D free energy profile of the R3063P mutant regarding the number of contacts between the Stalk and TOP domains and the R3848-E4078 distance calculated from the GaMD simulations. Important low-energy conformational states are identified, including the "Intermediate I1" and "Intermediate I2". (C) Structural conformation of Stalk-TOP interaction for the R3063P mutant. (D) Residue interactions between TOP-PL as observed

in the “Intermediate I1” conformational state. (E) Residue interactions between TOP-PL as observed in the “Intermediate I2” conformational state.

Figure 5. Experimental validation of GaMD simulation-predicted residue interactions between TOP-PL domains important for WT stalk-mediated activation of PC1 CTF. (A) Average fold NFAT reporter activation (+SD) relative to empty vector (ev) for WT and GaMD mutant CTF proteins from 3 separate transfection experiments with n=3 wells/expression construct/experiment. (B) Average relative total expression level (+SD) for WT and GaMD mutant CTF proteins from experiments in (A). (C) Representative Western blots of WT and GaMD mutant CTF proteins from one of the experiments represented in (B) probed with C20 antibody. Images are from two separate Western blots. Removal of intervening lanes is indicated by solid line. Blots were originally probed with C20 antibody and reprobed for actin. (D) Average relative cell surface expression level (+SD) for WT and GaMD mutant CTF proteins from 2-6 separate surface biotinylation experiments. (E) Representative Western blots of the surface biotinylation analyses of WT and GaMD mutant proteins from one of the experiments represented in (D). Supernatant (Sup) lane is from the pulldown with WT CTF. Images are from two separate blots. Removal of intervening lanes indicated by solid lines. Blots were originally probed with C20 or E8 antibody as indicated and reprobed for actin. *, $p<0.05$; **, $p\leq0.01$; ***, $p<0.005$; #, $p<0.0005$; and ##, $p<0.0001$ relative to WT CTF levels.

Figure 6. Model of the Stalk TA-mediated activation signaling of the PC1 CTF. (A) A “Closed” low-energy conformation is identified for the WT CTF, in which signal is transduced through an allosteric pathway connecting the Stalk-TOP-PL domains and ultimately to the C-tail for G protein activation. (B) Stalk variants (including ΔStalk,

G3052R and R3063C) of CTF adopt a distinct “Open” conformation, which exhibits reduced residue interactions between the Stalk-TOP domains and broken salt bridge interaction of residues R3848 and E4078 between the TOP-PL domains, leading to significantly reduced signaling activity of these PC1 mutants.

Figure 1.

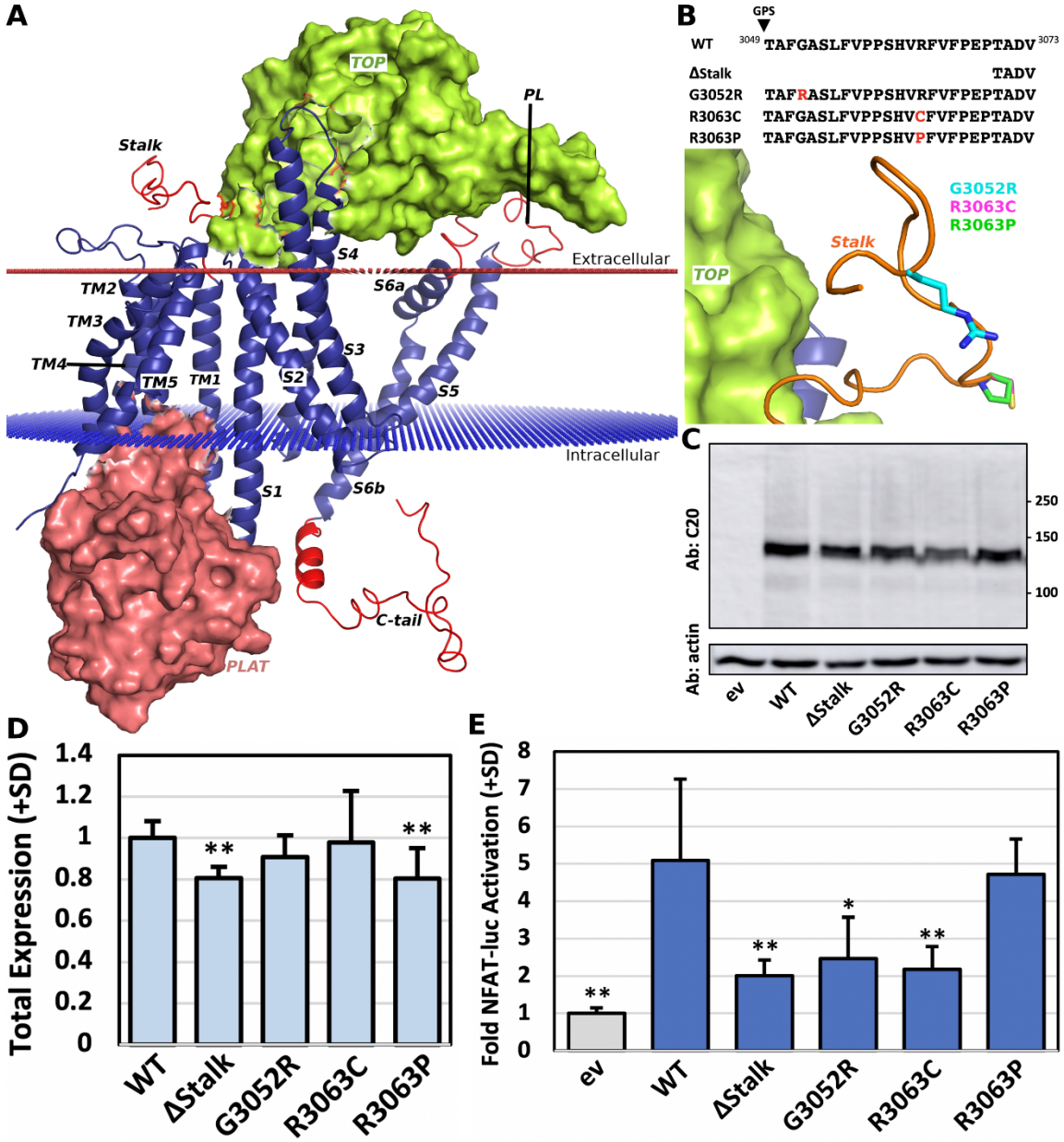

A
TOP
PL
Stalk
Extracellular
S4
S6a
TM2
TM3
TM4
TM5
TM1
S2
S3
S5
Intracellular
S1
S6b
C-tail
PLAT
B
GPS
WT 3049 TAFGASLFVPPSHVRFVFPEPTADV 3073
ΔStalk TADV
G3052R TAFRASLFVPPSHVRFVFPEPTADV
R3063C TAFGASLFVPPSHVCFVFPEPTADV
R3063P TAFGASLFVPPSHVPFVFPEPTADV
G3052R
R3063C
R3063P
Stalk
TOP
C
Ab: C20
Ab: actin
250
150
100
ev
WT
ΔStalk
G3052R
R3063C
R3063P
D
Total Expression (+SD)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
**
**
WT
ΔStalk
G3052R
R3063C
R3063P
E
Fold NFAT-luc Activation (+SD)
8
7
6
5
4
3
2
1
0
**
**
*
**
ev
WT
ΔStalk
G3052R
R3063C
R3063P

Figure 2.

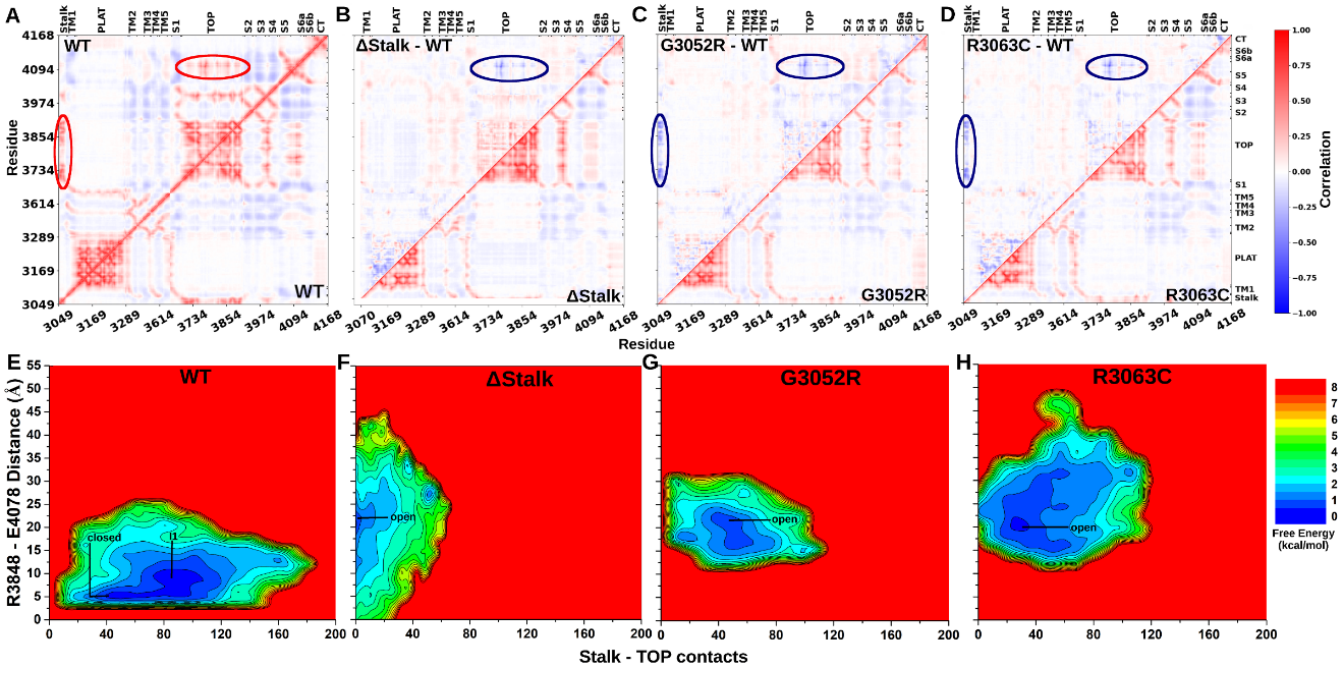

A
WT
B
ΔStalk - WT
C
G3052R - WT
D
R3063C - WT
Residue
Correlation
E
WT
closed
F
ΔStalk
open
G
G3052R
open
H
R3063C
open
R3848 - E4078 Distance (Å)
Stalk - TOP contacts
Free Energy (kcal/mol)

Figure 3.

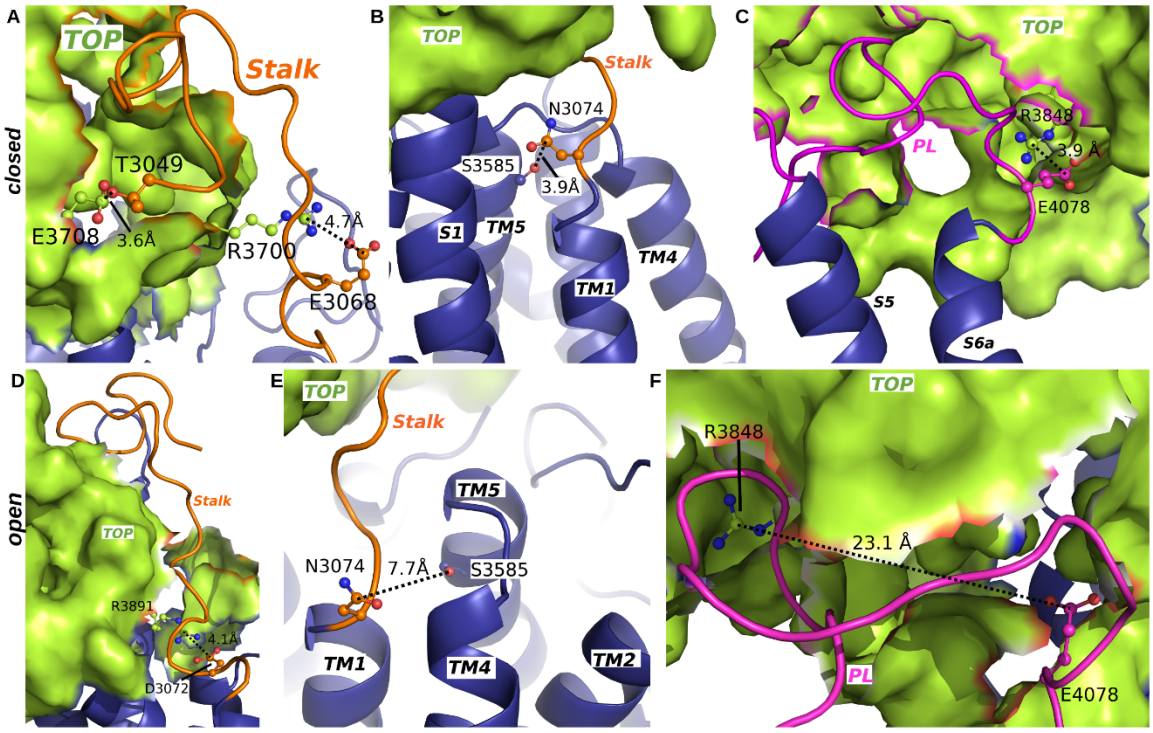

A
TOP
Stalk
closed
T3049
E3708
3.6Å
R3700
4.7Å
E3068
B
TOP
Stalk
N3074
S3585
3.9Å
S1
TM5
TM4
TM1
C
TOP
R3848
PL
3.9 Å
E4078
S5
S6a
D
open
Stalk
TOP
R3891
4.1Å
D3072
E
TOP
Stalk
TM5
N3074
7.7Å
S3585
TM1
TM4
TM2
F
TOP
R3848
23.1 Å
PL
E4078

Figure 4.

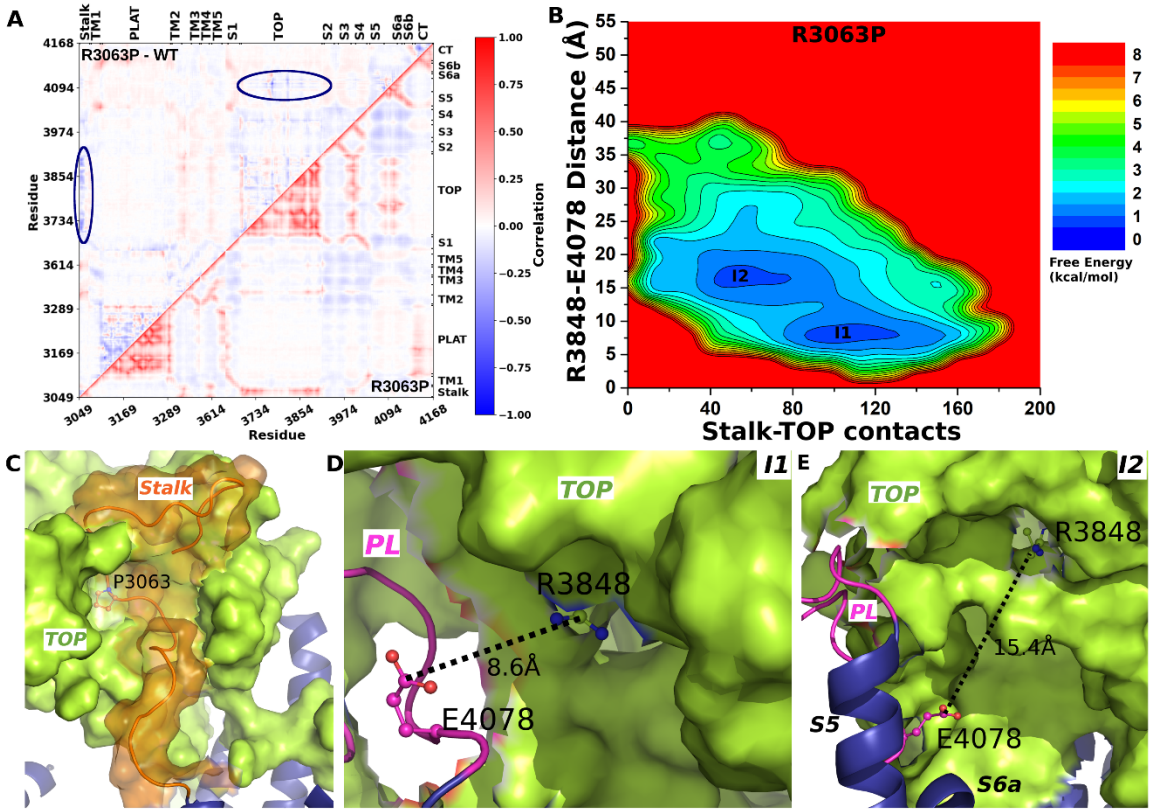

A
R3063P - WT
R3063P
Residue
Correlation
B
R3063P
R3848-E4078 Distance (Å)
Stalk-TOP contacts
Free Energy (kcal/mol)
I1
I2
C
Stalk
P3063
TOP
D
I1
TOP
PL
R3848
8.6Å
E4078
E
I2
TOP
R3848
PL
15.4Å
S5
E4078
S6a

Figure 5.

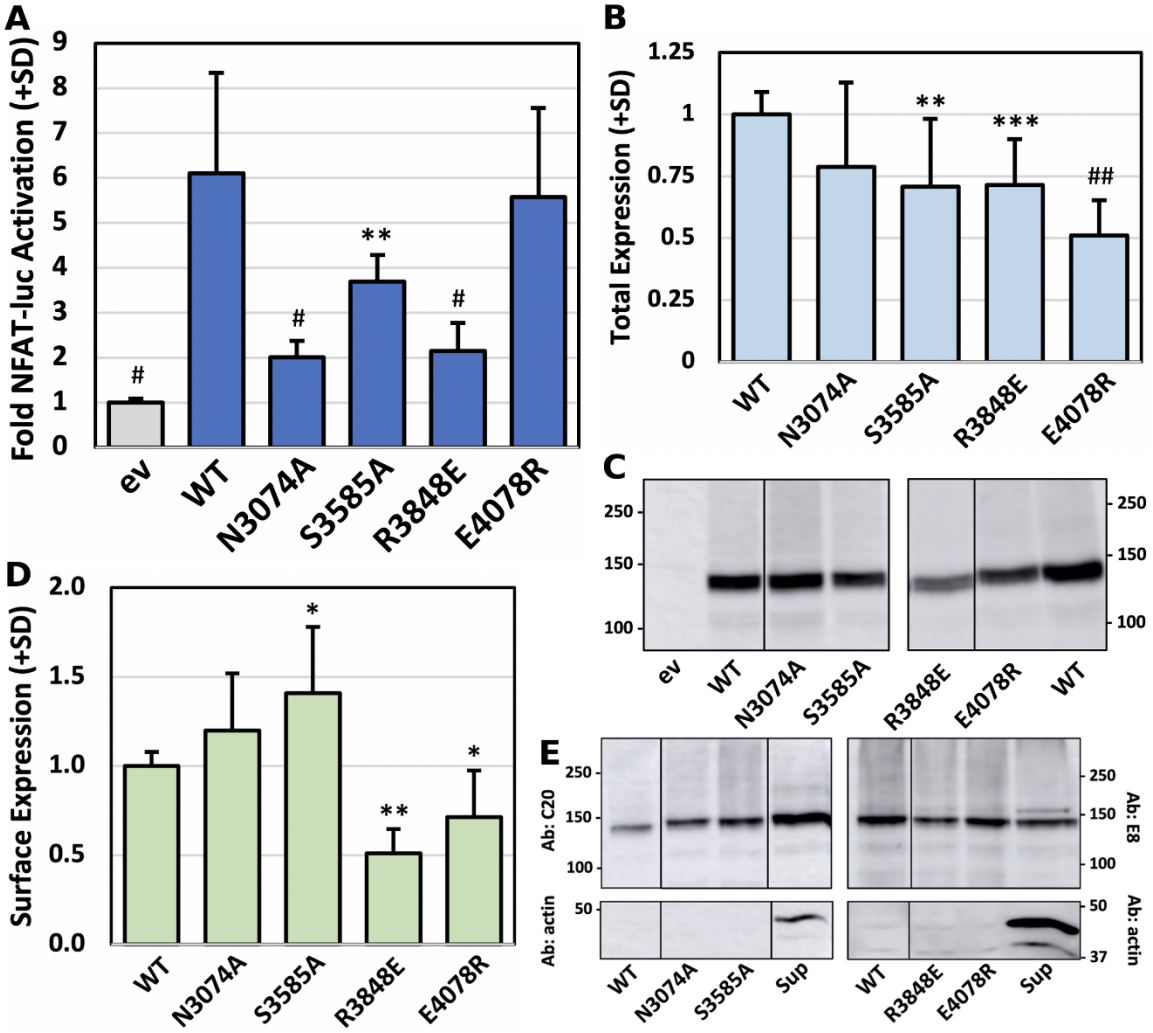
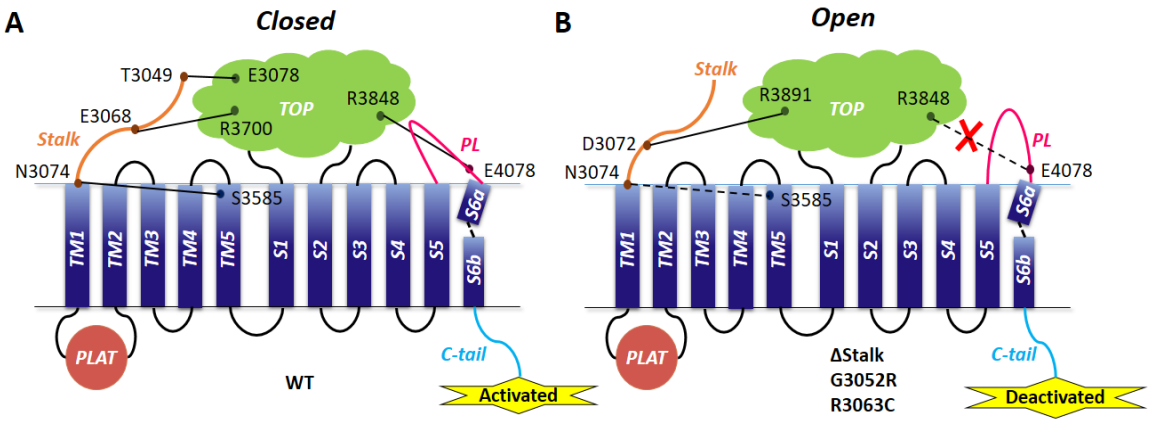

A
Fold NFAT-luc Activation (+SD)
9
8
7
6
5
4
3
2
1
0
#
#
**
#
ev
WT
N3074A
S3585A
R3848E
E4078R
B
Total Expression (+SD)
1.25
1
0.75
0.5
0.25
0
**
***
##
WT
N3074A
S3585A
R3848E
E4078R
C
250
150
100
ev
WT
N3074A
S3585A
R3848E
E4078R
WT
D
Surface Expression (+SD)
2.0
1.5
1.0
0.5
0.0
*
**
*
WT
N3074A
S3585A
R3848E
E4078R
E
Ab: C20
250
150
100
Ab: E8
Ab: actin
50
37
WT
N3074A
S3585A
Sup
WT
R3848E
E4078R
Sup

Fig. 6


A
Closed
T3049
E3078
E3068
R3700
TOP
R3848
Stalk
PL
N3074
E4078
S3585
S6a
TM1
TM2
TM3
TM4
TM5
S1
S2
S3
S4
S5
S6b
PLAT
C-tail
WT
Activated
B
Open
Stalk
R3891
TOP
R3848
D3072
PL
N3074
E4078
S3585
S6a
TM1
TM2
TM3
TM4
TM5
S1
S2
S3
S4
S5
S6b
PLAT
ΔStalk
G3052R
R3063C
C-tail
Deactivated