

# Control of protein signaling using a computationally designed GTPase/GEF orthogonal pair

Gregory T. Kapp<sup>a,1,2</sup>, Sen Liu<sup>a,1,3</sup>, Amelie Stein<sup>a</sup>, Derek T. Wong<sup>b,c</sup>, Attila Reményi<sup>c,4</sup>, Brian J. Yeh<sup>c,d</sup>, James S. Fraser<sup>c</sup>, Jack Taunton<sup>e</sup>, Wendell A. Lim<sup>e</sup>, and Tanja Kortemme<sup>a,5</sup>

<sup>a</sup>Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA 94158; <sup>b</sup>Graduate Program in Bioengineering, University of California, Berkeley and San Francisco, CA 94158; <sup>c</sup>Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94158; <sup>d</sup>Graduate Program in Chemistry and Chemical Biology, University of California, San Francisco, CA 94158; and <sup>e</sup>Howard Hughes Medical Institute and Department of Molecular and Cellular Pharmacology, University of California, San Francisco, CA 94158

Edited by David Baker, University of Washington, Seattle, WA, and approved January 27, 2012 (received for review September 6, 2011)

Signaling pathways depend on regulatory protein-protein interactions; controlling these interactions in cells has important applications for reengineering biological functions. As many regulatory proteins are modular, considerable progress in engineering signaling circuits has been made by recombining commonly occurring domains. Our ability to predictably engineer cellular functions, however, is constrained by complex crosstalk observed in naturally occurring domains. Here we demonstrate a strategy for improving and simplifying protein network engineering: using computational design to create orthogonal (non-crossreacting) protein-protein interfaces. We validated the design of the interface between a key signaling protein, the GTPase Cdc42, and its activator, Intersectin, biochemically and by solving the crystal structure of the engineered complex. The designed GTPase (*orthoCdc42*) is activated exclusively by its engineered cognate partner (*orthoIntersectin*), but maintains the ability to interface with other GTPase signaling circuit components *in vitro*. In mammalian cells, *orthoCdc42* activity can be regulated by *orthoIntersectin*, but not wild-type Intersectin, showing that the designed interaction can trigger complex processes. Computational design of protein interfaces thus promises to provide specific components that facilitate the predictable engineering of cellular functions.

computational modeling and design | signal transduction | synthetic biology

Most approaches to engineering cellular systems with new functions have taken advantage of the relative ease with which DNA elements can be used to control gene expression (1–3). In contrast, few studies have attempted to directly engineer protein-protein interaction networks. Recent pioneering examples include engineered control of input/output relationships in protein circuits (4, 5), protein-based logic gates (6), and control of protein activity in biological processes by light (7–9). Essentially all of these approaches create fusions of existing modular protein elements to yield diverse functions (10).

Nonetheless, our ability to create new functions by domain recombination is constrained by the toolkit of domains that are naturally available. Reuse of the same or closely related domains can yield undesired or unanticipated crosstalk, complicating the ability to predictably modify function within the context of a complex cellular protein interaction network. A potential solution to this problem would be to modify protein-protein interfaces directly by tuning interaction affinity and specificity as well as by creating orthogonal protein pairs (11). In its simplest form, an orthogonal pair consists of two engineered proteins that specifically interact with each other, but avoid significant crosstalk with their native wild-type counterpart proteins (Fig. 1A). Such orthogonal interactions are useful for achieving predictable biological control in a variety of contexts. For example, orthogonal interactions could be used to insulate a desired functional pathway from another competing process. Orthogonal protein pairs could also allow more precise control if they can be specifically triggered by

a small molecule to rapidly activate their function. One approach to engineering orthogonal systems is to borrow molecular components from a different organism. However, components from other organisms might not properly interface with existing cellular machinery and require further engineering to control multi-component cellular pathways. Instead, it may be advantageous to “rewire” existing protein interactions to create orthogonal pairs that can be externally controlled. Such protein network engineering strategies are not only useful to reengineer cells to perform new functions, but also to delineate the existing functional interaction networks.

Computational design has been successfully applied to many protein engineering applications (11, 12), including design of proteins with new or altered protein-protein interactions (11, 13, 14). A clear next challenge is to design protein interfaces to create orthogonal proteins that can perform and control complex biological functions in the context of cells and organisms. Here we describe such a proof-of-concept application of computational protein design, which generated an engineered pair of interacting proteins that is orthogonal to the wild-type proteins. The orthogonal interaction can be specifically triggered by a small molecule, and can interface with existing cellular components to control complex biological responses both in an *in vitro* reconstituted system and in mammalian cells.

## Results

**The GTPase Model System and Design Principles.** As our model system we chose the interactions of Rho-type GTPases that function as binary switches in signal transduction networks controlling key biological functions such as establishment of cell polarity and cell motility *via* regulation of the actin cytoskeleton (15). GTPases control signaling by cycling between the GDP-bound, inactive state, and the GTP-bound, active state that can bind to down-

Author contributions: G.T.K., S.L., A.S., W.A.L., and T.K. designed research; G.T.K., S.L., A.S., D.T.W., A.R., B.J.Y., and J.S.F. performed research; D.T.W., B.J.Y., J.S.F., and J.T. contributed new reagents/analytic tools; G.T.K., S.L., A.S., D.T.W., A.R., B.J.Y., J.S.F., J.T., and T.K. analyzed data; and G.T.K., S.L., A.S., J.S.F., and T.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, [www.pdb.org](http://www.pdb.org) (PDB ID code 3QBV).

See Commentary on page 5140.

<sup>1</sup>G.T.K. and S.L. contributed equally to this work.

<sup>2</sup>Present address: Omnicore, Inc., San Francisco, CA 94158.

<sup>3</sup>Present address: Institute of Molecular Biology, Medical Science College, China Three Gorges University, Yichang 443002, China.

<sup>4</sup>Present address: Department of Biochemistry, Eötvös Loránd University, 1117 Budapest, Hungary.

<sup>5</sup>To whom correspondence should be addressed. E-mail: [kortemme@cgl.ucsf.edu](mailto:kortemme@cgl.ucsf.edu).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1114487109/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1114487109/-DCSupplemental).

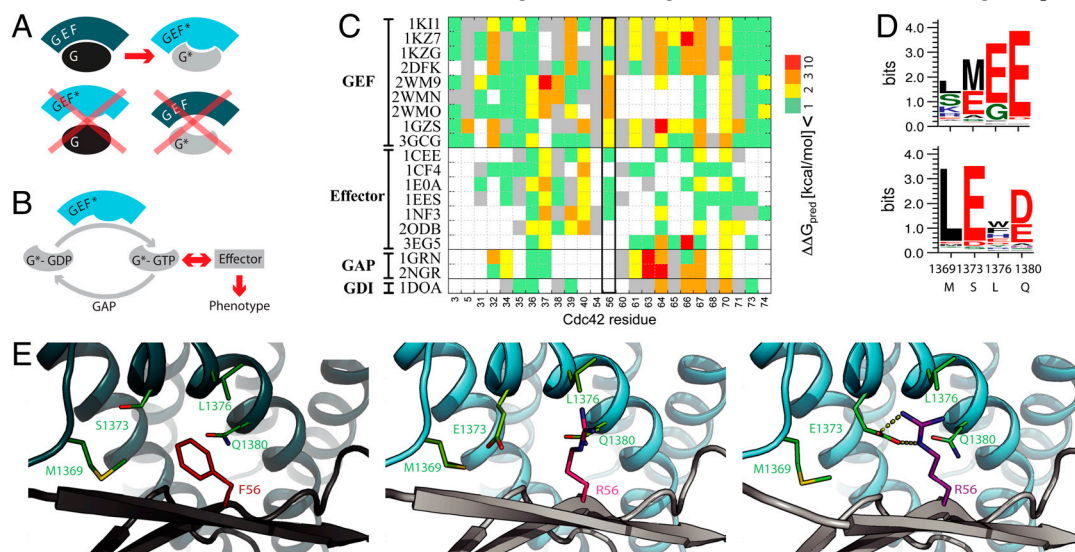
stream effector proteins and propagate signaling information (Fig. 1B). The GTPase switch (16) is modulated by accessory proteins: GTPase Activating Proteins (GAPs) accelerate the hydrolysis of GTP to GDP (inhibiting signal transduction), GTPase Exchange Factors (GEFs) accelerate the exchange of GDP and GTP (promoting signal transduction), and GDP Dissociation Inhibitors (GDIs) modulate the distribution of cytosolic and membrane-bound pools of GTPase. Therefore, the core GTPase signaling circuit comprises interactions of the GTPase with several different binding partners (GAP, GEF, GDI, and effector). Because the main regulatory process activating the GTPase switch involves regulation of the interaction between the GTPase and GEFs (17), we chose one such interaction, between the GTPase Cdc42 and a Cdc42-specific GEF Intersectin (ITSN), as our target.

We sought to design a functionally orthogonal interaction between a mutant Cdc42 (*ortho*Cdc42) and a mutant ITSN (*ortho*ITSN) that is buffered from the native pair (Cdc42<sup>WT</sup> and ITSN<sup>WT</sup>). This process should create two functionally interacting cognate pairs (Cdc42<sup>WT</sup>/ITSN<sup>WT</sup> and *ortho*Cdc42/*ortho*ITSN) and two noncognate pairs with no observable interaction specificity (Cdc42<sup>WT</sup>/*ortho*ITSN and *ortho*Cdc42/ITSN<sup>WT</sup>) (Fig. 1A). To allow the newly created cognate pair to perform its biological function, we constrained the design to minimally perturb residues implicated in the recognition interfaces of GDP, GTP, GAPs, effectors, and GDIs by Cdc42. Cellular activation of Cdc42 leads to considerable changes in cell morphology via induction of actin polymerization through the effector WASP. We thus expected that activation of *ortho*Cdc42 by *ortho*ITSN should be able to trigger WASP binding in vitro and morphology changes in cells.

**Computational Design Strategy.** We first sought to identify residues in Cdc42 that affect the binding interface between Cdc42 and ITSN, but do not affect interactions with other known Cdc42 binding partners. We performed a computational alanine scan (18) on 19 co-complex structures of Cdc42 with its binding partners (nine GEFs, two GAPs, seven effectors, and one GDI), to estimate the contribution of each interface residue to binding

each partner (Fig. 1C, *SI Appendix*, Fig. S1A). The alanine scan identified position 56 as the main candidate that affects GEF binding without perturbing the interactions with other binding partners. Additionally the F56 sidechain is spatially separated from the nucleotide binding and catalytic sites of Cdc42, suggesting that changes at F56 should not affect the affinity of Cdc42 for GDP or GTP.

Next we wanted to identify appropriate mutations around F56 in the Cdc42/ITSN interface that would lead to a functional interaction between *ortho*Cdc42 and *ortho*ITSN without introducing crosstalk between the noncognate pairs. An initial application of our previously developed computational second-site suppressor protocol (19) to the structure of Cdc42 and ITSN (PDB ID: 1KI1) (20) suggested that almost all substitutions of F56 would be destabilizing to the interaction (*SI Appendix*, Fig. S1B). However, none of the predicted compensatory changes of neighboring residues in ITSN was specific to the identity of the mutated amino acid modeled at position 56. Failure to correctly predict precise details of sidechain-sidechain interactions is a known problem with computational design approaches that leave the protein backbone fixed, such as the original second-site suppressor protocol (11, 19, 21, 22). Thus, we applied a recently developed flexible backbone design method, RosettaBackrub (23), to predict residue changes on ITSN that would compensate for mutations at Cdc42 F56. Although the flexible backbone design method has been benchmarked on existing data (23, 24), this application represents the first forward engineering test of the method's efficacy (see *Methods*). Using a flexible backbone ensemble created with RosettaBackrub, we allowed all possible mutations (except cysteine) at the four sites in ITSN adjacent to Cdc42 position 56 (1,369; 1,373; 1,376; 1,380) and searched for specific interactions with a variant residue at the 56 site of Cdc42. The predicted sequence logo in Fig. 1D shows that the mutation of position 1,373 in ITSN from serine to glutamate was distinctly enriched when F56 in Cdc42 was mutated to arginine [the residue with the largest predicted difference between destabilization in the noncognate and stabilization in the cognate pair (*SI Appendix*,



**Fig. 1.** Strategy for computational design of an orthogonal signaling interaction. (A) Schematic representation of design requirements for orthogonality: the interface between the GTPase Cdc42 (G) and ITSN (GEF) is modified to generate a pair G\*/GEF\* with new specificity. (B) Simplified schematic representation of the core GTPase signaling circuit to define the design requirements for a functional G\*/GEF\* pair that interfaces correctly with other cellular components (GAP and effector proteins that are required for phenotypic output). (C) Computational alanine scanning. Shown are the estimated effects on binding energy of replacing each residue in the Cdc42/ITSN interface (PDB code 1KI1) with alanine in the context of 19 co-complex structures of Cdc42 with partner proteins (white indicates residues not in the interface in the respective structure). Altering position F56 of Cdc42 mainly affects interaction with GEFs. (D) Comparison of fixed backbone (top) and flexible backbone (bottom) computational design predictions for four residues in ITSN (wild-type residues are indicated on the x axis) in the vicinity of position 56 of Cdc42 for a F56R mutation. (E) Model of designed *ortho*Cdc42/*ortho*ITSN interface from fixed (middle) and flexible (right) backbone modeling compared to the wild-type complex (left). Gray: Cdc42; Teal: ITSN; shown in sticks are the five designed interface residues. Small backbone changes modeled by backrub motions (0.53 Å Cα rmsd) allowed the sidechains of R56 and E1373 to adopt conformations that can form hydrogen bonds (dashed lines).

Fig. S1B)]. This interaction replaces the hydrophobic interactions (F56-L1376) observed in the original pair with a defined polar interaction (R56-E1373) in the designed pair (Fig. 1E). Importantly, a computational model of the complex of Cdc42 (F56R) and ITSN (S1373E) now showed specific hydrogen bonds formed between these two engineered sidechains that were not observed with fixed backbone simulations performed under identical conditions (Fig. 1E). Changing only one of these interacting residues in either Cdc42 (F56R) or ITSN (S1373E) is predicted to significantly destabilize the interactions between noncognate pairs. In the following, the specific Cdc42 (F56R) and ITSN (S1373E) variants are named *ortho*Cdc42 and *ortho*ITSN, respectively. (For designed variants other than the *ortho*Cdc42/*ortho*ITSN pair, see *SI Appendix, Results, Table S1*).

**In Vitro Nucleotide Exchange Activity and Binding Affinity.** We first determined the ability of the *ortho*ITSN DH-PH domains to catalyze nucleotide exchange in *ortho*Cdc42 by following the dissociation (Fig. 2A) and association (Fig. 2B) of fluorescently labeled nucleotide analogs. *ortho*ITSN specifically catalyzed exchange in *ortho*Cdc42 but not in Cdc42<sup>WT</sup>. Similarly, exchange in *ortho*Cdc42 was only catalyzed by *ortho*ITSN but not by ITSN<sup>WT</sup>. These results demonstrate that only one substitution in each protein is sufficient to engineer an *ortho*Cdc42/*ortho*ITSN pair that is indeed functionally orthogonal in vitro. This result is remarkable, given that such dramatic switches in protein-

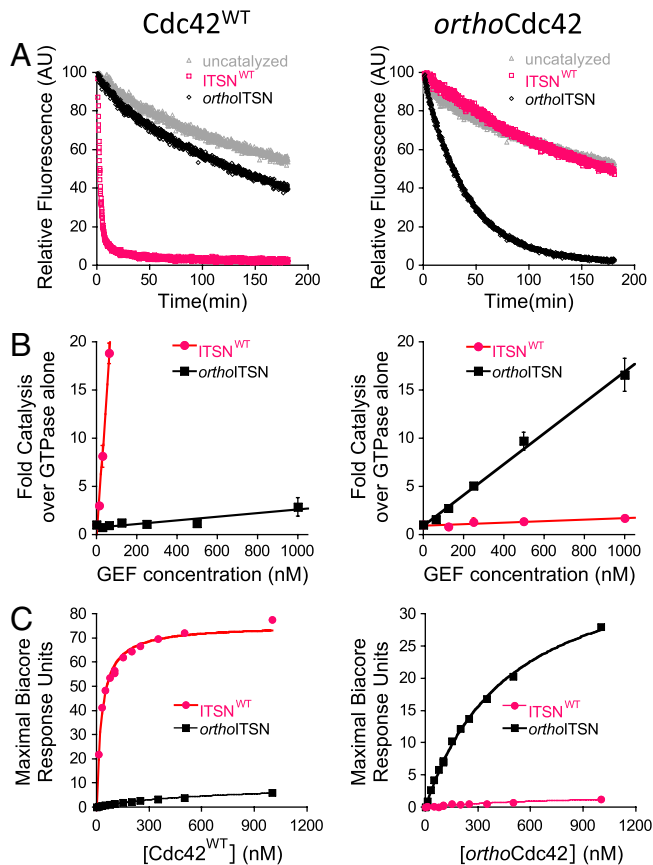
protein interaction specificity often require many changes (19, 21, 25). In our case, other modeled substitutions in ITSN, such as M1369L, did not change the exchange activity in *ortho*Cdc42/*ortho*ITSN (*SI Appendix, Table S1*). Moreover, the ITSN Q1380E mutation (a prominent prediction of the fixed backbone protocol, Fig. 1D) was not active towards *ortho*Cdc42 in combination with S1373E (*SI Appendix, Results, Fig. S2*), further confirming the importance of the specific R56-E1373 interaction.

While the designed mutations essentially eliminated cross-reactivity with the wild-type partners in noncognate complexes (Fig. 2B), *ortho*ITSN was a weaker nucleotide exchange catalyst for *ortho*Cdc42 compared to ITSN<sup>WT</sup> for Cdc42<sup>WT</sup>. To explain this weaker activity, we analyzed both the stability of the engineered variants and their binding affinity. Neither mutation significantly destabilized the engineered proteins, as indicated by similar apparent melting temperatures monitored using circular dichroism (*SI Appendix, Fig. S3*). However, the weaker functional interactions were consistent with direct binding affinity measurements of cognate and noncognate Cdc42 and ITSN complexes determined by surface plasmon resonance (Fig. 2C, *SI Appendix, Fig. S4*). The interaction between Cdc42<sup>WT</sup> and ITSN<sup>WT</sup> had a  $K_D$  of  $29 \pm 2$  nM, similar to that determined in a previous study (33 nM) (26). The  $K_D$  of *ortho*Cdc42 and *ortho*ITSN was  $478 \pm 22$  nM, approximately 16-fold weaker. Importantly, essentially no binding was observed under our conditions between the noncognate Cdc42<sup>WT</sup>/*ortho*ITSN or *ortho*Cdc42/ITSN<sup>WT</sup>, directly demonstrating the physical origin of the orthogonal relationship between cognate pairs.

**Structural Basis of the Designed Specificity.** To assess the accuracy of the design model, we determined the crystal structure of the complex between *ortho*Cdc42 and the DH-PH domains of *ortho*ITSN (Fig. 3, *SI Appendix, Results, Fig. S5, Table S2*). The structure confirms the engineered salt bridge interaction between the sidechains of R56 in *ortho*Cdc42 and E1373 in *ortho*ITSN (Fig. 3B, *SI Appendix, Fig. S5A*). However, there are notable downstream rearrangements of sidechains extending up to about 10 Å from the designed site, where sidechains of N39 and Y40 in *ortho*Cdc42 essentially switch positions (Fig. 3C, *SI Appendix, Fig. S5B*), concomitant with backbone changes in the interface.

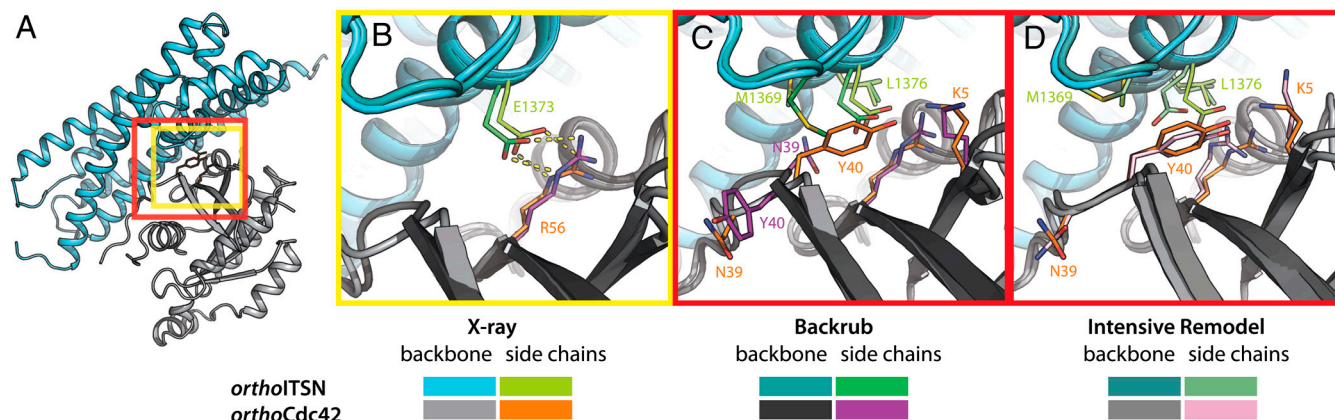
While the RosettaBackrub prediction successfully captured the defined interaction between the two designed residues by allowing small backbone adjustment and brought the backbone conformation slightly closer to that of the designed structure (*SI Appendix, Fig. S6B*), it had not captured the larger conformational change accompanying the sidechain rearrangements around Y40. Such conformational changes are a possible reason for the reduced biochemical activity in our case, and are also likely to occur more generally in response to designed mutations in interfaces. We thus tested whether a new remodeling protocol (*SI Appendix, Fig. S6, Methods, Results*) that switches between diversifying conformations and intensifying sampling, while iterating between energy functions using soft and hard repulsive forces, could model such interface changes. Intensive sampling around the designed interface site indeed yielded a conformation (the lowest energy structure in one of six resulting clusters) that was very close (0.56 Å Cα rmsd in the region of interest) to the solved crystal structure of the design and recapitulated the experimentally observed switch in the sidechains positions of N39 and Y40 (Fig. 3D).

**Interactions with Other GTPase Binding Partners.** The substitution in *ortho*Cdc42 was designed to minimize effects on other known binding partners of the GTPase (Fig. 1C). One of the most important interactions in the Cdc42 activation cycle is the binding of GTP-bound Cdc42 to the effector protein WASP, which allows for activation of the Arp2/3 complex, and induces actin polymerization. A second key interaction is with GAPs that accelerates



**Fig. 2.** The designed interaction is orthogonal in vitro. In (A)–(C), Cdc42<sup>WT</sup> is shown on the left and *ortho*Cdc42 on the right. Pink: data for ITSN<sup>WT</sup>; black: data for *ortho*ITSN. (A) Catalysis of nucleotide exchange by ITSN<sup>WT</sup> and *ortho*ITSN, monitored by dissociation of fluorescent mant-GDP from Cdc42<sup>WT</sup> and *ortho*Cdc42. Gray: intrinsic exchange in Cdc42 in the absence of any ITSN. (B) Catalysis of nucleotide exchange from initial rates of mant-GDP association at varying GEF concentrations. Data represent averages and standard deviations from at least three experiments. (C). Binding affinity monitored by Surface Plasmon Resonance equilibrium analysis (*SI Appendix, Fig. S4*).





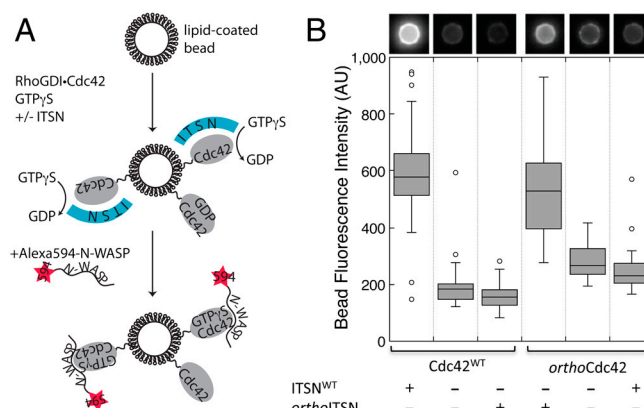
**Fig. 3.** The crystal structure of the *orthoCdc42/orthoITSN* complex confirms the designed interaction, but also highlights requirements for advanced flexible-backbone remodeling protocols. (A) Overview of the structure of the designed complex between *orthoCdc42* (gray) and the *orthoITSN* DH domain (teal). Boxes highlight the location of the designed site near the center of the protein-protein interface (yellow) as well as the area of backbone and side-chain rearrangements (red), magnified in (B–D). Sidechain and backbone colors are as indicated in the figure. (B) Comparison of the R56-E1373 interaction in the backrub flexible-backbone computational model (as in Fig. 1E, right) and in the crystal structure of the designed *orthoCdc42/orthoITSN* complex. Dashed lines represent hydrogen bonds. (C, D) Comparison of the network of residues surrounding the designed site that were rearranged to accommodate the mutations, as predicted by the backrub model (C) and the intensive remodeling protocol (D, details in *SI Appendix, Results*) vs. their observed position in the crystal structure of the designed complex. The remodeling protocol (D) was able to capture both sidechain and backbone conformational changes in the crystal structure of *orthoCdc42/orthoITSN* that were missed by the initial backrub predictions (C).

the hydrolysis of GTP bound to GTPases. Consistent with the design strategy, *orthoCdc42* binds to a fragment of N-WASP (residues 201–321) (although with an approximately fourfold weaker  $K_D$  than *Cdc42*<sup>WT</sup>, *SI Appendix, Fig. S7A*), and p50RhoGAP can enhance nucleotide hydrolysis in *orthoCdc42* (*SI Appendix, Fig. S7B*). Full-length *orthoCdc42* (containing a prenylated C-terminal CAAX motif) can also bind the Guanine Dissociation Inhibitor RhoGDI (*SI Appendix, Fig. S7C*). In addition to the interaction with ITSN, *Cdc42* has intrinsic specificity for other exchange factors, which is preserved in *orthoCdc42* (*SI Appendix, Results, Table S3*). Taken together, these results suggest that *orthoCdc42* can still interact with core components of the GTPase signaling circuit, and that the designed substitutions in *orthoCdc42* and *orthoITSN* have not introduced new and undesirable crosstalk with other known GTPases and GTPase signaling circuit components (*SI Appendix, Table S3*).

**In Vitro Reconstitution of a Partial Signaling Pathway.** The biochemical analysis above suggests that the engineered substitutions of *orthoCdc42* and *orthoITSN* have generated a new protein pair that does not interact with the wild-type proteins, but where *orthoCdc42* maintains binary interactions with other *Cdc42* regulation factors. To test the function of the designed pair in the context of a larger *Cdc42* pathway, we used an in vitro assay with purified components to monitor N-WASP recruitment to lipid-coated beads (27) (Fig. 4A). This assay mimics activation of membrane-bound *Cdc42* by GEF-catalyzed nucleotide exchange and subsequent interaction of GTP-bound *Cdc42* with the effector N-WASP. As designed, the localization of fluorescently labeled N-WASP (residues 137–502) to the surface of lipid-coated beads increased only in the presence of the *Cdc42*<sup>WT</sup>/*ITSN*<sup>WT</sup> or the *orthoCdc42/orthoITSN* cognate pairs, but not with the noncognate pairs (Fig. 4B). Kolmogorov-Smirnov testing of the bead fluorescence intensity distributions indicated that these differences were significant ( $p < 1.5 \times 10^{-6}$  for each condition, three independent experiments with at least 20 individual beads counted per experiment). Consistent with the previously noted weaker affinity of the designed pair, the required concentration of *orthoITSN* was higher (2.5  $\mu$ M) than *ITSN*<sup>WT</sup> (1  $\mu$ M) in each respective condition.

**Pathway Activity with Designed Components in Mammalian Cells.** We next tested whether the designed *orthoCdc42/orthoITSN* pair, de-

spite its lower exchange activity and weakened affinity compared to the wild-type complex, still functions in endogenous signaling networks of GTPases and GEFs in mammalian cells. We coupled the designed protein-protein interaction with a small molecule-based inducible localization system similar to that described in (28). Using this method, the cell-permeable small molecule Rapamycin can be added to recruit FK506 binding protein (FKBP)-linked ITSN to the plasma membrane by inducing Rapamycin-mediated binding of FKBP to FK506-rapamycin-binding (FRB) protein, which is localized to the membrane using the membrane-targeting domain from the Lyn protein (Fig. 5A). Activated *Cdc42* is known to induce the formation of filopodia in NIH 3T3 mouse fibroblast cells (29), as well as lamellipodia by activating the GTPase Rac through interaction with the IRSp53 protein (30). Thus, increasing the local ITSN concentration near



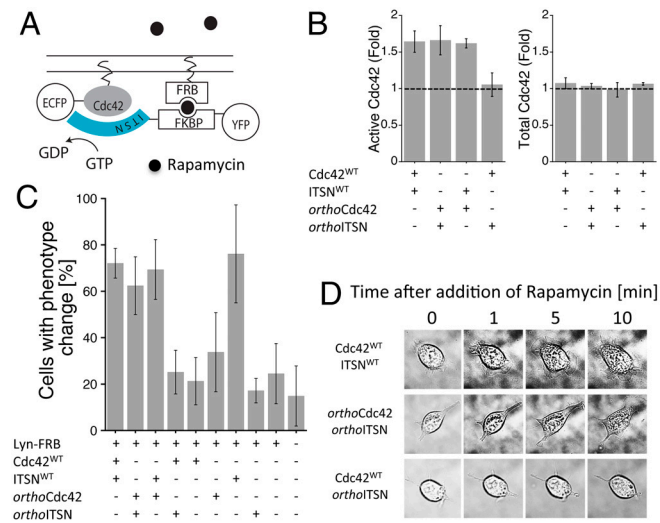
**Fig. 4.** The designed *orthoCdc42/orthoITSN* interaction mediates specific GTPase activation and effector binding in an in vitro reconstituted system. Alexa 594 labeled N-WASP (residues 137–502) translocation to a lipid-coated glass bead is specifically increased in the presence of a cognate interaction between *Cdc42* and ITSN. (A) Schematic illustrating the assay and the order of addition of the components. (B) The total fluorescence intensity of individual beads relative to the background was measured, and the distributions of the fluorescence intensities from multiple beads ( $n > 23$  for each condition) are shown in box plot representation. Boxes enclose the first and third quartile of the distribution and display a line at the median; whiskers extend outward no more than 1.5 times the size of the box and data points outside this range are drawn individually. A representative bead image is shown above each condition.



the membrane should lead to nucleotide exchange and activation of membrane-localized inactive Cdc42, which in turn activates Cdc42 signaling to induce cell morphological changes. In this way, because Cdc42 activation should be triggered by Rapamycin-dependent ITSN recruitment, any change in cellular phenotype can be observed in the same background before and after the addition of the small molecule.

We first determined whether *ortho*ITSN could activate *ortho*Cdc42 in cells by measuring the levels of activated and total Cdc42 before and after the addition of Rapamycin (see *Methods*). *ortho*ITSN indeed activated *ortho*Cdc42, but not Cdc42<sup>WT</sup>, as expected (Fig. 5B). Overall, the activation of *ortho*Cdc42 by *ortho*ITSN was similar to the activation of Cdc42<sup>WT</sup> by ITSN<sup>WT</sup>, and the active Cdc42 was at the highest level in the first 60–90 s after the addition of Rapamycin (SI Appendix, Fig. S84). Finally, to determine whether activation of *ortho*Cdc42 by *ortho*ITSN could result in morphological changes (filopodia and/or lamellipodia) in NIH 3T3 cells, we counted cells that showed induced morphological changes after the addition of Rapamycin (Fig. 5C) using fluorescence microscopy of living cells (Fig. 5D). Consistent with the Cdc42 activation assay (Fig. 5B), increased filopodia/lamellipodia were observed in cells transfected with either the *ortho*Cdc42 and *ortho*ITSN designed pair or the wild-type pair, but not with the noncognate Cdc42<sup>WT</sup>/*ortho*ITSN pair. Similarly, transfection of *ortho*ITSN in the presence of the Rapamycin recruitment system but in the absence of *ortho*Cdc42 resulted in considerably less phenotypic change.

We note that these assays (as also apparent in Fig. 5B) cannot determine orthogonality with respect to the other noncognate pair *ortho*Cdc42/ITSN<sup>WT</sup>, as Rapamycin-induced localization of ITSN<sup>WT</sup> most likely leads to activation of endogenous Cdc42<sup>WT</sup>. Consistent with this idea, transfection with the other noncognate pair *ortho*Cdc42/ITSN<sup>WT</sup> had levels of morphological change



**Fig. 5.** The *ortho*Cdc42/*ortho*ITSN pair is functional in mammalian cells. (A) Schematic representation of the cell-based assay using a Rapamycin-based recruitment system (FRB, FKBP) to colocalize fluorescently tagged GTPase and GEF constructs at the membrane. (B) Fold increase in active Cdc42 (comparing samples with and without addition of Rapamycin for 60 s) from lysed NIH 3T3 cells measured with a G-LISA assay (left). The total Cdc42 loaded in the G-LISA assay was determined by an ELISA assay, and is also shown in fold change, again comparing samples with and without Rapamycin addition (right). All samples had Lyn-FRB transfected. Error bars represent the standard deviation of three experiments. (C) Percentage of NIH 3T3 cells that showed morphological changes (filopodia/lamellipodia) after addition of Rapamycin, determined by live cell microscopy. All samples had Lyn-FRB transfected. Error bars represent the standard deviation of three experiments. The total numbers of counted cells for each condition, from left to right, are: 103, 111, 133, 120, 57, 62, 55, 50, 71, and 84. (D) Representative images of cell morphological changes upon Rapamycin addition.

similar to the cognate pairs. Furthermore, transfection of ITSN<sup>WT</sup> alone (but including the membrane-recruiting construct Lyn-FRB) shows an equivalent level of morphological change. Taken together, including additional results monitoring morphological changes by impedance (SI Appendix, Results, Fig. S8B), the cellular assays indicate that the designed *ortho*Cdc42/*ortho*ITSN interaction functions within cells to trigger production of filopodia/lamellipodia.

## Discussion

In this work, we used advanced computational protein design methods to reengineer a signaling circuit by direct modification of an interaction interface; this approach stands in contrast to previous work that either engineered expression control at the gene level or recombined existing modular protein domains. We show that the designed proteins function orthogonally in vitro and trigger responses in cells. Therefore, the engineered interacting orthogonal pair still interfaces with existing cellular machinery to direct changes in cell morphology, a complex phenotypic outcome.

Engineering orthogonality of specific interactions, while at the same time maintaining correct interfaces with existing machinery, is challenging in multiple respects. The orthogonality of the designed interaction is remarkable, given that it was achieved with only one residue change on either partner, but it comes at the price of reduced affinity. Detailed structural analysis of designed proteins is critical for evaluating inaccuracies in the design model. The defined interaction of the designed R-E pair in a central interface location, on which our predictions were based, was correctly captured in the model. However, deviations further away from the designed site illustrate the difficulty of predicting energetics and conformations of interacting residues, in particular polar networks in protein interfaces. It is not unlikely that the different conformations of the polar interaction network (SI Appendix, Fig. S5B) are approximately isoenergetic and that small changes in the surroundings, including long-range effects, can cause population shifts resulting in coordinated conformational changes. It may be difficult to predict these changes computationally in part because the relative free energy differences may be small. In this context, it is remarkable that a new intensive backbone remodeling protocol is capable of sampling conformations close to the observed structure (Fig. 3D). Currently, the Rosetta energy does not distinguish between these models, and structural clustering is necessary to reveal the diversity of the sampled conformations (SI Appendix, Fig. S6).

It is difficult to find sites in multifunctional proteins such as GTPases that can be engineered without pleiotropic consequences on many interactions or detrimental effects on function altogether. In fact, position 56, identified here by computational design as the major engineerable site (Fig. 1), may be one of a few sites that can be mutated in Cdc42 without dramatically affecting multiple partner interactions. F56 of Cdc42 has previously been implicated as a residue that defines the specificity of Cdc42 for various GEFs including ITSN (20, 31, 32). In contrast to previous studies that switched between existing interaction preferences, however, our design has created a different specificity. This finding prompts the question of whether the F56R and S1373E substitutions are present in any other existing GTPase-GEF interactions. Of the 23 Rho subfamily GTPases in the human genome, none have arginine at the position equivalent to F56 (33). In the 66 characterized human GTPase exchange factor sequences, only five have glutamate at the position equivalent to S1373. All five have either been shown to not catalyze exchange in Cdc42, or are members of the Lbc subfamily that in general does not catalyze exchange in Cdc42 (34, 35). These results suggest that the substitutions designed by computational methods are unique.

Almost every protein is involved in a number of interactions with different binding partners. The ability to design new speci-

cities into target interfaces without affecting other interactions is useful both for the biological interrogation of protein interactions and for the design of circuits that could produce new biological behaviors. This study indicates that computational methods can become an essential tool for the design of new protein interfaces. Improving computational design methodologies, including approaches to more accurately model structural and sequence plasticity in interfaces (11), will allow protein engineers and synthetic biologists to create new interactions of increasing complexity and specificity.

## Methods

**Computational Protein Interface Design.** The crystal structure of Cdc42<sup>WT</sup>/ITSN<sup>WT</sup> (PDB ID: 1K11) (20) was used as starting conformation for structure-based computational protein design. Computational alanine scanning was performed as described (18). For fixed backbone design, we used the computational second-site suppressor protocol as described (19) (*SI Appendix, Fig. S1B*). These simulations aimed to identify substitutions in one protein that are significantly destabilizing to the complex formed with the wild-type partner protein but can be compensated for by complementary changes in the partner. Flexible-backbone protein design used RosettaBackrub (23, 36) and the sequence tolerance protocol developed in (23, 24). One hundred low-scoring backrub structures were generated from the starting structure of Cdc42<sup>WT</sup>/ITSN<sup>WT</sup>, and used as a backbone ensemble in design simulations to determine sequence tolerated at the Cdc42/ITSN interface. In the design step, the amino acid identity at Cdc42 position 56 was fixed but the residue was allowed to change its rotameric conformation, and the four neighboring residues (M1369, S1373, L1376, Q1380) in ITSN were allowed to change to any other residues (designed) except cysteine. The intensive flexible-backbone design and remodeling strategy (*SI Appendix, Results, Fig. S6*) begins with modeling the F56R and S1373E mutations, followed by backbone diversification using RosettaBackrub (36) and kinematic closure (KIC) methods (37), and final intensified sampling and refinement using KIC. Soft and hard repulsive forces are iterated similar to a recently described protocol for protein folding (38). Simulation details and all Rosetta command lines are given in *SI Appendix, Methods*.

**Protein Biochemistry.** All in vitro assays except the N-WASP translocation experiments used soluble forms of the GTPases (residues 1–179 in Cdc42) lacking the C-terminal prenylation sites. All exchange factor sequences were derived from human or mouse cDNA and encoded both the DH and PH domains (*SI Appendix, Table S4*). Proteins for in vitro experiments were expressed and purified from *Escherichia coli*, and nucleotide dissociation and association assays were performed as detailed in *SI Appendix, Methods*. Cdc42–ITSN binding affinities were determined by surface plasmon resonance (SPR) experiments similar to those described in Smith, et al. (26), and the N-WASP translocation assay was performed as described by Co, et al. (27). (For more details on protein in vitro assays see *SI Appendix, Methods*).

**Crystallography.** Crystals were grown at room temperature as hanging drops above a well of 100 mM Tris pH 7.5, 25% PEG 3350, 150 mM ammonium sulfate, and 1 mM DTT. Crystals were harvested using a solution of 20% glycerol and 17% PEG 3350 as a cryoprotectant. Details on data collection, analysis and structure determination are given in *SI Appendix, Methods*. The PDB model was deposited as: 3QBV.

**Cell-Based Assays.** The Cdc42 G-LISA Kit (Cytoskeleton) was used to detect active GTP-bound Cdc42 in NIH 3T3 cells, and an ELISA assay was used to measure the total Cdc42 loaded (*SI Appendix, Methods*). For live cell fluorescence microscopy, NIH 3T3 cells were cultured in 8-well Lab-Tek II Chambered Coverglass wells. After serum starvation, pictures were taken on a Nikon Eclipse Ti Microscope with a 60X or 100X objective at 37 °C (*SI Appendix, Methods*).

**ACKNOWLEDGMENTS.** We thank Orion Weiner, Anselm Levskaya, Ben Rhau, and Alex Watters for helpful suggestions, Colin Smith and Shane O'Connor for help with design simulations, Kris Kuchenbecker and Peter Hwang for help with SPR, Farid Ahmad for help with crystallography, and James Onuffer, Benjamin Rhau, and Jason Park for help with cell-based assays. T.K. is supported by awards from the National Science Foundation (MCB-CAREER 0744541, EF-0849400), the Sandler Foundation, and the UC Lab Research Program. W.A.L., J.T. and T.K. were supported by the US National Institutes of Health Roadmap Initiative (PN2 EY016546, W.A.L., principal investigator). W.A.L. was supported by awards from the National Science Foundation (EEC-0540879), and the National Institute of Health (RO1 GM062583, P50 GM081879). J.S.F. is a QB3@UCSF Fellow. A.S. is an EMBO long-term fellow.

- Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–338.
- Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403:339–342.
- Sprinzak D, Elowitz MB (2005) Reconstruction of genetic circuits. *Nature* 438:443–448.
- Yeh BJ, Rutigliano RJ, Deb A, Bar-Sagi D, Lim WA (2007) Rewiring cellular morphology pathways with synthetic guanine nucleotide exchange factors. *Nature* 447:596–600.
- Bashor CJ, Helman NC, Yan S, Lim WA (2008) Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. *Science* 319:1539–1543.
- Dueber JE, Yeh BJ, Chak K, Lim WA (2003) Reprogramming control of an allosteric signaling switch through modular recombination. *Science* 301:1904–1908.
- Wu YI, et al. (2009) A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* 461:104–108.
- Levskaya A, Weiner OD, Lim WA, Voigt CA (2009) Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* 461:997–1001.
- Leung DW, Otomo C, Chory J, Rosen MK (2008) Genetically encoded photoswitching of actin assembly through the Cdc42-WASP-Arp2/3 complex pathway. *Proc Natl Acad Sci USA* 105:12797–12802.
- Peisajovich SG, Garbarino JE, Wei P, Lim WA (2010) Rapid diversification of cell signaling phenotypes by modular domain recombination. *Science* 328:368–372.
- Mandell DJ, Kortemme T (2009) Computer-aided design of functional protein interactions. *Nat Chem Biol* 5:797–807.
- Pokala N, Handel TM (2001) Review: protein design—where we were, where we are, where we're going. *J Struct Biol* 134:269–281.
- Fleishman SJ, et al. (2011) Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science* 332:816–821.
- Grigoryan G, Reinke AW, Keating AE (2009) Design of protein-interaction specificity gives selective bZIP-binding peptides. *Nature* 458:859–864.
- Etienne-Manneville S, Hall A (2002) Rho GTPases in cell biology. *Nature* 420:629–635.
- Cherfils J, Zeghouf M (2011) Chronicles of the GTPase switch. *Nat Chem Biol* 7:493–495.
- Schmidt A, Hall A (2002) Guanine nucleotide exchange factors for Rho GTPases: turning on the switch. *Genes Dev* 16:1587–1609.
- Kortemme T, Baker D (2002) A simple physical model for binding energy hot spots in protein-protein complexes. *Proc Natl Acad Sci USA* 99:14116–14121.
- Kortemme T, et al. (2004) Computational redesign of protein-protein interaction specificity. *Nat Struct Mol Biol* 11:371–379.
- Snyder JT, et al. (2002) Structural basis for the selective activation of Rho GTPases by Dbl exchange factors. *Nat Struct Biol* 9:468–475.
- Joachimik LA, Kortemme T, Stoddard BL, Baker D (2006) Computational design of a new hydrogen bond network and at least a 300-fold specificity switch at a protein-protein interface. *J Mol Biol* 361:195–208.
- Sammond DW, Eletr ZM, Purbeck C, Kuhlman B (2010) Computational design of second-site suppressor mutations at protein-protein interfaces. *Proteins* 78:1055–1065.
- Smith CA, Kortemme T (2011) Predicting the tolerated sequences for proteins and protein interfaces using RosettaBackrub flexible backbone design. *PLoS One* 6:e20451.
- Smith CA, Kortemme T (2010) Structure-based prediction of the peptide sequence space recognized by natural and synthetic PDZ domains. *J Mol Biol* 402:460–474.
- Levin KB, et al. (2009) Following evolutionary paths to protein-protein interactions with high affinity and selectivity. *Nat Struct Mol Biol* 16:1049–1055.
- Smith WJ, et al. (2005) A Cdc42 mutant specifically activated by interseitin. *Biochemistry* 44:13282–13290.
- Co C, Wong DT, Gierke S, Chang V, Taunton J (2007) Mechanism of actin network attachment to moving membranes: barbed end capture by N-WASP WH2 domains. *Cell* 128:901–913.
- Inoue T, Heo WD, Grimley JS, Wandless TJ, Meyer T (2005) An inducible translocation strategy to rapidly activate and inhibit small GTPase signaling pathways. *Nat Methods* 2:415–418.
- Krugmann S, et al. (2001) Cdc42 induces filopodia by promoting the formation of an IRSp53:Mena complex. *Curr Biol* 11:1645–1655.
- Ladwein M, Rottner K (2008) On the Rho'd: the regulation of membrane protrusions by Rho-GTPases. *FEBS Lett* 582:2066–2074.
- Gao Y, Xing J, Streuli M, Leto TL, Zheng Y (2001) Trp(56) of rac1 specifies interaction with a subset of guanine nucleotide exchange factors. *J Biol Chem* 276:47530–47541.
- Karnoub AE, et al. (2001) Molecular basis for Rac1 recognition by guanine nucleotide exchange factors. *Nat Struct Biol* 8:1037–1041.
- Colicelli J (2004) Human RAS superfamily proteins and related GTPases. *Science Signaling Knowledge Environment* 2004:RE13.
- Miki T, Smith CL, Long JE, Eva A, Fleming TP (1993) Oncogene ect2 is related to regulators of small GTP-binding proteins. *Nature* 362:462–465.
- Glaven JA, Whitehead IP, Nomanbhoy T, Kay R, Cerione RA (1996) Lfc and Lsc oncoproteins represent two new guanine nucleotide exchange factors for the Rho GTP-binding protein. *J Biol Chem* 271:27374–27381.
- Smith CA, Kortemme T (2008) Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction. *J Mol Biol* 380:742–756.
- Mandell DJ, Coutsiar EA, Kortemme T (2009) Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nat Methods* 6:551–552.
- Khatib F, et al. (2011) Algorithm discovery by protein folding game players. *Proc Natl Acad Sci USA* 108:18949–18953.



# SUPPLEMENTARY MATERIALS

## Control of protein signaling using a computationally designed GTPase/GEF orthogonal pair

Gregory T. Kapp, Sen Liu, Amelie Stein, Derek T. Wong, Attila Reményi, Brian Yeh, James S. Fraser, Jack Taunton, Wendell A. Lim, and Tanja Kortemme

## Supplementary Methods

### Computational specificity redesign

Initial fixed backbone specificity redesign used the computational second site suppressor protocol described previously (1). Flexible backbone specificity redesign employed the sequence tolerance protocol described in (2), using Rosetta revision r33982 and the command lines below.

*Generate an ensemble of structures with the backrub method:*

```
PATH/TO/ROSETTA/backrub.EXECUTABLE -database PATH/TO/ROSETTA_DATABASE
-s 1kil.pdb -resfile <RESFILE_NAME> -ex1 -ex2 -ex1aro -ex2aro
-extrachi_cutoff 0 -out:prefix <PREFIX_NAME> -mute core.io.pdb.file_data
-backrub:ntrials 10000 -score:weights standard_NO_HB_ENV_DEP.wts
-backrub:minimize_movemap <MOVE_MAP> -nstruct <NUMBER OF MODELS>
```

*Scan for tolerated sequences:*

```
PATH/TO/ROSETTA/sequence_tolerance.EXECUTABLE
-database PATH/TO/ROSETTA_DATABASE
-s <BACKRUB_INPUT_STRUCTURE_NAME> -resfile <RESFILE_NAME>
-ex1 -ex2 -ex1aro -ex2aro -extrachi_cutoff 0 -score:ref_offsets HIS 1.2
-seq_tol:fitness_master_weights 1 1 1 2
-ms:generations 30 -ms:pop_size 200 -ms:pop_from_ss 1
-ms:checkpoint:prefix <NAME> -ms:checkpoint:interval 200
-ms:checkpoint:gz -score:weights standard_NO_HB_ENV_DEP.wts
-out:prefix <NAME>
```

### Flexible-backbone design with intensive structural remodeling

Design and remodeling used Rosetta revision r42980, the steps described in SI Results, and Rosetta command lines as below:

### Soft repulsive fixed backbone design

#### Without ligand:

```
PATH/TO/ROSETTA/fixbb.EXECUTABLE -database PATH/TO/ROSETTA_DATABASE
-s 1ki1.pdb -resfile <RESFILE_NAME> -ex1 -ex2 -ex3 -ex4 -extrachi_cutoff 0
-score:weights soft_rep_design -out:prefix <PREFIX> -nstruct 30 -overwrite
-out:pdb_gz
```

#### With GDP:

```
PATH/TO/ROSETTA/fixbb.EXECUTABLE -database PATH/TO/ROSETTA_DATABASE
-s <1KI1_WITH_GDP> -resfile <RESFILE_NAME> -ex1 -ex2 -ex3 -ex4
-extrachi_cutoff 0 -score:weights soft_rep_design -out:prefix <PREFIX>
-nstruct 30 -overwrite -out:pdb_gz -extra_res_fa <PARAMS_FILE>
```

The `PARAMS_FILE` contains instructions for Rosetta on how to handle the ligand, including possible conformation(s). It needs to be specifically generated for each type of ligand. For all simulations reported here, the presence or absence of a GDP ligand in the structure did not lead to significantly different results in terms of backbone RMSD or side chain conformations around the mutated sites (the GDP binding site is distant from the F56R and S1373E mutations, with a distance of 9.9Å between the closest atoms).

### Backrub ensemble generation

```
PATH/TO/ROSETTA/backrub.EXECUTABLE -database PATH/TO/ROSETTA_DATABASE
-s <designed_structure> -in:file:fullatom -ex1 -ex2 -ex3 -ex4
-extrachi_cutoff 0 -resfile <RESFILE> -out:prefix <PREFIX> -overwrite
-out:pdb_gz -backrub:ntrials 10000 -nstruct 1 -out:path test
-mute core.io.pdb.file_data -pivot_residues 330 319 318 212 211 51 324 310
317 316 315 314 60 234 53 66 67 68 69 80 230 231 171 24 322 20 21 23 320 40
41 289 323 3 321 5 4 7 6 9 8 328 281 285 327 201 205 204 208 325 329 306 307
77 76 75 74 73 72 71 70 107 79 78 10 39 38 58 17 16 19 54 57 56 37 36 35 52
55 333 168 326 292 293
```

-pivot\_residues determines which residues may be used as pivots by Backrub. This list restricts the pivots to 10Å around the designed residues, using Rosetta's internal residue numbering which is sequential across all chains, starting at 1.

### Soft repulsive KIC



```

PATH/TO/ROSETTA/loopmodel.EXECUTABLE -database PATH/TO/ROSETTA_DATABASE
-in:file:fullatom -loops:loop_file <LOOP_FILE> -loops:refine refine_kic
-in:file:native <BACKRUBBED_STRUCTURE>
-loops:input_pdb <BACKRUBBED_STRUCTURE> -score:weights soft_rep_design
-out:prefix <PREFIX> -overwrite -out:pdb_gz -nstruct 1
-out:path <OUT DIR> -vicinity_sampling false -loops:neighbor_dist 6
-ex1 -ex2 -ex3 -ex4 -extrachi_cutoff 0

```

#### *Hard repulsive KIC with vicinity sampling*

```

PATH/TO/ROSETTA/loopmodel.EXECUTABLE -database PATH/TO/ROSETTA_DATABASE
-in:file:fullatom -loops:loop_file <LOOP_FILE> -loops:refine refine_kic
-in:file:native <SOFT_KIC_DECOY> -loops:input_pdb <SOFT_KIC_DECOY>
-out:prefix <PREFIX>_ -overwrite -out:pdb_gz -nstruct 1 -out:path <OUT DIR>
-loops:neighbor_dist 6 -ex1 -ex2 -ex3 -ex4 -extrachi_cutoff 0

```

#### *Clustering*

```

PATH/TO/ROSETTA/cluster.EXECUTABLE -database PATH/TO/ROSETTA_DATABASE
-l <LIST OF DECOYS> -cluster:radius 0.7 -in:file:fullatom
-ignore_unrecognized_res -native 1kil.pdb -nooutput -exclude_res 1 2 3 4 5 6
7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33
34 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67
68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93
94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114
115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133
134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152
153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171
172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190
191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209
210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228
229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247
248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266
267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285
286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304
305 306 307 308 309 310 311 312 313 314 324 325 326 327 328 329 330 331 332
333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351
352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370
371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389
390 391 392 393

```

-exclude\_res specifies the residues to be ignored for RMSD calculation – only those in the flexible loops are considered here. This uses Rosetta’s internal residue numbering, which is sequential across all chains, starting at 1.

## **Plasmids**

All constructs used in this paper are listed in **SI Table S4**. All sequence substitutions were made using the QuikChange mutagenesis system (Stratagene). All sequences were verified by DNA sequencing.

## **Protein expression and purification**

Proteins were expressed as His6 fusion proteins in the Rosetta2 strain of *E. coli* (EMD Biosciences) using a 3 hour induction with IPTG (Isopropyl  $\beta$ -D-1-thiogalactopyranoside). Cells were lysed by sonication, His6 tagged proteins were bound to Ni-NTA resin (Qiagen) and eluted in 50 mM Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 250 mM imidazole, pH 8.0. The His6 tag was cleaved by room temperature incubation with a His6-tagged TEV (Tobacco Etch Virus) protease, followed by removal of the protease and free His6 tags using a second Ni-NTA purification. GTPases were further purified using a SourceQ ion exchange column (Amersham).

GTPase concentrations were determined by the Coomassie Plus system (Pierce). Exchange factor concentrations were determined by absorbance at 280nm using extinction coefficients calculated using the method of Pace *et al.* (3).

## ***In vitro* nucleotide exchange assays**

For nucleotide dissociation assays, purified GTPases were loaded with mantGDP (methylantraniloyl-GDP, Molecular Probes) by incubation with a ten-fold molar excess of mantGDP for 30 minutes at room temperature in a buffer of 20mM Tris pH 7.6, 200mM NaCl, 1mM DTT, and 10mM EDTA. Nucleotide loading was quenched by addition of 10-fold molar excess of MgCl<sub>2</sub> above the EDTA concentration and excess nucleotide was removed using NAP-5 columns (GE Healthcare) equilibrated in Exchange Assay Buffer (20 mM Tris, 50 mM NaCl, 10 mM MgCl<sub>2</sub>, 1% glycerol, 1 mM DTT, pH 7.5). Dissociation of mantGDP from GTPases was measured in a SpectraMax Gemini XS (Molecular Devices) fluorescence multi-well plate reader (25°C, excitation: 360 nm, emission: 440 nm). Solutions were pre-equilibrated at 25°C for 10 minutes, and the reaction was initiated by transferring pre-mixed GEF/GDP to mantGDP-bound GTPases. Final concentrations were 1  $\mu$ M mantGDP-bound GTPase, 1  $\mu$ M GEF, 200  $\mu$ M GDP in Exchange Assay Buffer.

For nucleotide association assays, GTPase and GEF were mixed with Exchange Assay Buffer to a final



concentration of 0.5  $\mu$ M for GTPase and varying concentrations of GEF. The solutions were equilibrated for 10 minutes before the addition of mantGDP to a final concentration of 400 nM to start the reaction. Reaction progress was monitored by fluorescence as above. Rates were determined by linear fits to the initial rates of exchange (4). The fold catalysis was determined by dividing the catalyzed rate by the uncatalyzed rate (for the GTPase alone without GEF).

### **Circular dichroism (CD) spectroscopy**

CD data were collected on each protein (Cdc42 and ITSN, WT and variants) at concentrations close to 10 $\mu$ M on an Aviv CD spectrophotometer. CD data collection was done in a buffer of 10 mM sodium phosphate, pH 7.0, and 100 mM NaCl, in a 0.2 cm cuvette. Samples were cooled to 4°C and then heated to 90°C and the ellipticity at 222 nm recorded at 3°C increments. Ellipticity was converted to mean residue ellipticity (MRE).

### **Surface plasmon resonance**

All experiments were performed on a Biacore T100 instrument using a running buffer of HBS-P (0.01 M HEPES, 0.15 M NaCl, pH 7.4, 0.005% v/v Tween 20) with the addition of 50  $\mu$ M EDTA. Roughly 600 response units (RU) of GEF were immobilized on a CM5 sensor chip (Biacore) using the amine coupling kit. Injections at a number of concentrations (0, 10, 30, 50, 75, 100, 150, 200, 250, 350, 500, 1000 nM) of analyte (Cdc42<sup>WT</sup> or *ortho*Cdc42) were used to determine the equilibrium binding affinities. Injections at concentrations above 1000nM showed evidence of non-specific binding events and were not used in the affinity determination. All injections were performed at 25 °C at a flow rate of 25  $\mu$ L/min with a 180 second association phase, a 240 second dissociation phase, a 30 second regeneration in HBS-P + 10 mM MgCl<sub>2</sub> + 1 mM GTP, and a final regeneration of 20 seconds of HBS-P + 5 mM EDTA. Equilibrium data were analyzed using Biacore Evaluation software (version 1.1.1) and the Rmax for each injection series was fit using the Steady State Affinity Fit with the offset at zero. The affinity values reported are the average of three concentration series.

### **Crystallography**

The *ortho*Cdc42 (F56R) and *ortho*ITSN DH/PH (S1373E) proteins were purified using the NiNTA resin and the His6 tags were cleaved and removed as described above. Each protein was then further purified by gel filtration over a Sephacryl S100HR column (Amersham) in a buffer of 50 mM sodium phosphate (pH 7.4) and 150 mM NaCl. Purified proteins were concentrated in a buffer of 20 mM Tris pH 7.5, 150 mM NaCl, 2 mM EDTA, and 1 mM DTT and then combined in a 1:1 molar ratio to a final concentration of 10mg/mL.

X-ray diffraction data were collected on beam 8.3.1 at the Advanced Light Source at Lawrence Berkeley National Laboratory. A single data set was collected from a crystal diffracting to 2.65 Å and processed in space group P21 with HKL2000 (5), phased by molecular replacement using AMORE (6) with 1KI1 as a search model. Rebuilding was performed manually with Coot (7) with iterative refinement using phenix.refine (8) using non-crystallographic symmetry between the two copies of the *ortho*ITSN/*ortho*Cdc42 complex present in the asymmetric unit.

### **WASP fluorescence titration**

WASP fluorescence titrations were performed as described previously (9). The W13 fragment of WASP(9) (residues 201-321) was cloned into a His6 expression vector, expressed and purified as described in the main Methods for GTPase and GEF proteins, and the His6 tag was cleaved by treatment with TEV protease. The W13 concentration was determined using an extinction coefficient ( $E_{280}$ ) of 8250 M<sup>-1</sup> cm<sup>-1</sup> (9). Purified Cdc42<sup>WT</sup> and *ortho*Cdc42 were preloaded with mantGMPPNP (Molecular Probes) as described in the main Methods for mantGDP. Proteins were diluted in 40 mM HEPES-NaOH pH 7.4, 100 mM NaCl. A 1 cm<sup>2</sup> cuvette was filled with a 200 nM solution of Cdc42•mantGMPPNP and maintained at 25 °C. The decrease in Cdc42•mantGMPPNP fluorescence with W13 addition was monitored using a Photon Technologies International (Birmingham, NJ) fluorimeter with excitation and emission set to 360 nm and 440 nm, respectively. Titrations were performed by manual injections of W13 solution using a Hamilton syringe allowing for 1 minute of mixing before averaging fluorescence emission for 2 seconds. Raw data were corrected for Cdc42 concentration and then fit as described (9).

### **GAP assay**

GTP hydrolysis by Cdc42 was tested using the EnzChek Phosphate Assay Kit (Invitrogen) and the assay protocol of Zhang *et al.* (10). Briefly, 8 μM soluble Cdc42 was combined with 5 mM MgCl<sub>2</sub>, 0.2 mM GTP (Roche), 0.2 mM 2-amino-6-mercapto-7-methylpurine riboside (MESG), and 0.5 units of purine nucleoside phosphorylase (PNP) in 50 mM Hepes pH 7.5, 0.1 mM EDTA. Inorganic phosphate released by Cdc42 coupled to the MESG by the PNP to generate a product with an absorbance at 360 nM. Absorbance readings were made using a SpectraMax Plus (Molecular Devices) reader. The addition of 1-4 nM of p50RhoGAP produced an increase in the rate of GTP hydrolysis (observed as a more rapid increase in absorbance).

### **N-WASP translocation to beads**

Full-length Cdc42<sup>WT</sup> and *ortho*Cdc42, including the C-terminal CAAX motif, were expressed as His6-tagged proteins in SF9 cells. SF9 lysates were combined with *E. coli* lysates expressing GST-tagged



bovine RhoGDI protein. The RhoGDI•Cdc42 complexes were then purified using a Ni-NTA column followed by a GST-agarose column (Amersham). For the assay, glass beads (2.3 µm diameter; Bangs Laboratories) were coated with a lipid mixture of 75% phosphatidylcholine, 20% phosphatidylserine, and 5% PIP2 and incubated with 50 µM GTPγS, 1 mM MgCl<sub>2</sub>, 1 mM DTT, and the indicated full-length Cdc42 protein (complexed with RhoGDI) and ITSN protein for 20 minutes. The final concentrations were 1 µM Cdc42•RhoGDI complex, 1 µM ITSN<sup>WT</sup> or 2.5 µM *ortho*ITSN. Fluorescent N-WASP (dEVH1 construct, residues 137-502; Alexa 594 labeled) was then added to a final concentration of 1µM and allowed to localize for an additional 20 minutes. The beads were fixed and imaged on an Olympus IX70 microscope at 60x magnification. The fluorescence intensity of individual beads was measured by determining the total integrated fluorescence of a 2.5 µm diameter circle enclosing the bead and subtracting the fluorescence of the same circle enclosing only background fluorescence (no beads). For each condition at least 20 individual beads were measured.

### **Nucleofection**

The Cell Line 96-well Kit SE from Lonza Cologne AG was used to transfect the plasmids (all constructs are as listed in **Table S4**) into NIH 3T3 cells. The cells were cultured in DMEM (Dulbecco's Modified Eagle Medium) containing 10% bovine calf serum (BCS) to 70-80% confluency, trypsinized, spun and resuspended in the manufacturer's SE solution (20 µL for 6×10<sup>5</sup> cells). Then 20 µL of cells were mixed with 1 µg total of pre-mixed indicated plasmids in 2 µl total volume, and transferred to a 96-well Nucleocuvette Plate well. The nucleofection was performed in the 96-well Shuttle system with the standard 96-CA-137 program. After incubating at room temperature for 10 min, 80 µL of DMEM (10% BCS) were added to each well.

### **G-LISA assay**

The Cdc42 G-LISA Kit (Cytoskeleton) was used to detect active GTP-bound Cdc42 in NIH 3T3 cells. 12-well culture plates were prepared by adding 1 mL of DMEM containing 10% BCS. After the nucleofection step, for each transfected sample, the cells were transferred to two prepared wells on the 12-well culture plates with 50 µL cells per well. After 8 h of culture followed by 7 h of starvation, for each transfected sample, Rapamycin in DMEM without serum was added to a final concentration of 20 µM to the cells in one well, and the other well served as the control by adding the same volume of DMSO as Rapamycin in DMEM without serum. Then, the medium was aspirated off at the indicated time points, and G-LISA Lysis Buffer was added to lyse the cells. The lysates were flash frozen in liquid nitrogen and stored at -70°C. The G-LISA assay was performed as specified as in the manufacturer's manual, after the lysates were diluted with G-LISA Lysis Buffer containing protease inhibitors to 0.7mg/mL total protein.

### **ELISA assay**

To measure the total Cdc42 loaded for G-LISA assay, the wells of ELISA plates were first coated with Chicken Polyclonal IgY Antibody to Cdc42 (AbCam). Then the cell lysates (same lysates as above) were added into the wells for Cdc42 binding, followed by adding an HRP-conjugated anti-Cdc42 monoclonal antibody (Santa Cruz Biotechnology) to Cdc42. TMB (3,3',5,5'-tetramethylbenzidine) substrate solution was used for detection, and the absorbance at 450nm was measured after the addition of 1 M H<sub>3</sub>PO<sub>4</sub> to stop the reaction.

### **Live cell fluorescence microscopy**

After the nucleofection step, NIH 3T3 cells were cultured in 8-well Lab-Tek II Chambered Coverglass wells. After the same serum starvation process as described above, pictures were taken on a Nikon Eclipse Ti Microscope with a 60X or 100X objective at 37 °C. Rapamycin was added to a final concentration of 20 μM as above.

### **xCELLigence assay**

The xCELLigence System (Roche Applied Science) can be used to monitor cell morphological changes in real time without the incorporation of labels. The electrode impedance, which is defined as cell index (CI) values, is correlated with the change of cell morphology (11). The xCELLigence system E-plate wells were coated with 40 μg/mL Fibronectin (Sigma-Aldrich) for 1 h at 37 °C. After washing with phosphate-buffered saline (PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4), the wells were blocked with 0.5% bovine serum albumin (BSA) solution in PBS for 20 min at 37 °C. 100 μL DMEM containing 10% BCS were added to each well after washing with PBS, and then the E-plates were equilibrated in an incubator (37 °C, 5% CO<sub>2</sub>). After the NIH 3T3 cells were transfected with indicated plasmids as described in the main Methods, 15 μL of transfected cells were transferred to each prepared E-plate well. The E-plates were then placed on the xCELLigence machine for real-time recording every 3 min. After eight hours, the medium in each well was changed to 100 μL of DMEM without serum for starvation. After 7 h of serum starvation, the data recording frequency was changed to 15 seconds and Rapamycin (Sigma-Aldrich) in DMEM without serum was added to each well to the final concentration of 20 μM.

## Supplementary Results

### Other designed ITSN variants using fixed and flexible backbone design

**Table S1** lists other predicted ITSN variants tested both in the context of Cdc42<sup>WT</sup> and *ortho*Cdc42. In addition to the ITSN S1373E variant, two other ITSN mutations were tolerated in the designed interface; the mutations were either predicted in sequences designed using flexible backbone simulations (M1369L) or observed in several ITSN homologs and predicted to be favorable (L1376I). In contrast, the ITSN Q1380E mutation, which was the most frequently observed mutation in fixed backbone design simulations (**Fig. 1D**), was not active towards *ortho*Cdc42 (**Fig. S2**) when tested in combination with S1373E (which was present simultaneously, although much less frequently, in fixed backbone simulations). These results illustrate the difficulty in correctly predicting the precise details of polar interactions in protein interfaces, in particular when using fixed backbone approaches.

The L1376W substitution in ITSN, which was predicted in flexible backbone simulations to be favorable for the interaction with Cdc42<sup>WT</sup> (**Fig. S1C**), and also to be tolerated by *ortho*Cdc42 (**Fig. 1D**) was not active in either context (**Fig. S2**), possibly because of steric incompatibilities that result from overpacking.

### Interactions with other GTPases and GTPase binding partners

The DH/PH domains of the exchange factor Dbs catalyze exchange in Cdc42<sup>WT</sup> but not in *ortho*Cdc42 (**SI Table S3**). One exchange factor, the DH domain of PREX1, is able to catalyze nucleotide exchange in both Cdc42<sup>WT</sup> and *ortho*Cdc42, whereas others, such as the DH/PH domains of the exchange factors Tiam1 and Trio (the N-terminal DH/PH) do not catalyze exchange in either Cdc42<sup>WT</sup> or *ortho*Cdc42 (**SI Table S3**). These results suggest that our design was successful in specifically perturbing the interaction with ITSN and the highly similar GEF Dbs.

Because ITSN is an exchange factor specific for Cdc42, ITSN does not catalyze nucleotide exchange in the Rac1 or RhoA GTPases, and *ortho*ITSN does not change this intrinsic specificity (it does not catalyze exchange in either Rac1 or RhoA, **SI Table S3**).

### Transferability of the designed mutations to a related GTPase/GEF interaction

Given the unique nature of the engineered complementary F56R/S1373E substitutions in the Cdc24/ITSN interface, we asked whether they could be transferred into another GTPase-GEF interface to make that interaction orthogonal with respect to the original wild-type binding partners and possibly other related GTPases and GEFs. However, simply porting the R-E pair to the structurally equivalent positions in the

Rac1-Tiam1 interface was not successful in compensating for the detrimental effect of each of the mutations alone (**SI Fig. S9**). This result is not surprising, given the intricate nature of coupled residue-residue interactions in proteins and protein-protein interfaces. Interestingly, repeating our flexible backbone design prediction protocol as illustrated in **Figure 1D** on the crystal structure of the Rac1-Tiam1 complex corroborates this result and does not show enrichment for negatively charged amino acids at the Tiam1 position equivalent to 1373 in ITSN (**SI Fig. S10**).

### Structural Analysis

The overall backbone C $\alpha$  RMSD values between the *ortho*Cdc42/*ortho*ITSN and the Cdc42<sup>WT</sup>/ITSN<sup>WT</sup> complex structure (PDB ID: 1KI1) are 0.58 Å and 0.47 Å for the Cdc42 molecules and the ITSN DH domains, respectively (**SI Fig. S5A**). However, there are regions with larger deviations in the backbone, in particular the loop around Y40 in Cdc42, which rearranges to accommodate side chain movements triggered by the mutations, as discussed in the main manuscript (**Fig. 3**). This region has a C $\alpha$  RMSD of 1.998 Å in the 36-43 loop. While the DH domain of ITSN forms the interface with Cdc42, the PH domain of *ortho*ITSN (which is spatially distant from the interface and thus the site of mutation) has poor density in our complex structure for a significant portion of the mainchain, and the orientation of the ITSN PH domain relative to the Cdc42 molecule differs in the two crystal structures, indicating possible domain-domain flexibility in solution. The different orientation in our structure results in a change to the crystal lattice.

The designed complex structure has defined electron density in the active site of *ortho*Cdc42 (**SI Fig. S5C**). In the course of refining the *ortho*Cdc42/*ortho*ITSN structure, this density persisted through simulated annealing omit and kicked map calculations. Based on the shape of the electron density, we modeled a GDP molecule bound in the active site of the *ortho*Cdc42/*ortho*ITSN, bound in the same orientation as Cdc42<sup>WT</sup> without any associated exchange factor (PDB ID: 1AN0, **SI Fig. S5D**). This electron density was surprising because the side chain of A59 is in essentially the same conformation as A59 in the Cdc42<sup>WT</sup>/ITSN<sup>WT</sup> structure (PDB ID: 1KI1, **SI Fig. S5E**). The position of A59 triggered by GEF binding is assumed to displace the Mg<sup>2+</sup> ion necessary for binding a GDP molecule. In contrast, the conformation observed in the Cdc42<sup>WT</sup> structure (PDB ID: 1AN0, **SI Fig. S5D**) is compatible with Mg<sup>2+</sup>-binding. The active site residues of Cdc42 bound to ITSN would not clash directly with GDP. Rather, the direct clash of A59 with the Mg<sup>2+</sup> atom, which likely disfavors GDP binding, leads to higher GDP mobility and an increased probability of dissociation. The B-factors for the GDP molecule were higher than the surrounding protein atoms (~80 vs 54 for the protein). Since the resolution is not sufficient to refine occupancies or to observe any correlated structural changes in the protein, the simplest explanation



is that the density represents a relatively disordered GDP molecule bound in the active site of *ortho*Cdc42. In the Cdc42<sup>WT</sup>/ITSN<sup>WT</sup> complex structure, an electron rich sulfate molecule is modeled in a position that overlaps with the placement of the beta-phosphate of the putative GDP (**SI Fig. S5E**). However, structure factors for this complex were not deposited, so electron maps cannot be calculated to determine if the density extends beyond the sulfate position. Currently we cannot distinguish whether the GDP molecule present in the *ortho*Cdc42/*ortho*ITSN is due to the engineered mutations, an intermediate conformation in nucleotide exchange, or would also be observed in the wild-type complex.

### **Flexible-backbone structure remodeling**

To test whether the observed conformational changes in the *ortho*Cdc42/*ortho*ITSN interface could be computationally recapitulated, we implemented an initial version of a flexible backbone design and remodeling protocol intended to predict significant structural changes in response to designed mutations (**Figure S6A**). This protocol uses two general concepts: The first is switching between steps that diversify backbone conformations and steps that focus sampling in certain regions of conformational space. This idea has been used successfully in protein structure refinement (12). Different backbone remodeling algorithms employing “backrub” (13) and “kinematic closure” (KIC) (14) moves allow us to diversify the conformations of the protein as well as to determine regions surrounding the designed positions that are particularly flexible and thus more likely to change upon mutation (12). The second concept is interleaving soft and hard repulsive forces, which enables us to model conformational changes that initially appear unfavorable, but may be accommodated by subsequent refinement steps using intensified sampling in defined regions. Successful application of this concept has recently been reported in protein structure refinement (15).

Modeling of the *ortho*Cdc42/*ortho*ITSN complex used the general protocol outlined in **Figure S6A** with the following steps and simulation details:

1. Design, soft potential: We introduced the two designed mutations, F56R and S1373E, into the template structure (PDB ID: 1K11), keeping the backbone fixed and repacking the side chains in a 10 Å radius around the mutations. This step used the Rosetta all atom energy function with soft repulsive forces, which allow slightly unfavorable conformations of the side chains to still be accepted.
2. Initial backbone diversification, hard potential: Backbone diversification of the initial design model employed backrub moves (13) in a 10 Å radius around the mutations using the Rosetta all atom energy function with hard repulsive forces. This step generated an ensemble of 1200 structures with slight variations in the backbone that may accommodate the designed residues, while reducing possible steric clashes from the soft repulsive design step. The overall structural variation in this

ensemble is low, however, and none of the structures is very close in backbone RMSD to the *ortho*Cdc42/*ortho*ITSN crystal structure (**Figure S6B**).

3. Aggressive backbone diversification, soft potential: To generate larger diversity, the next diversification step employed kinematic closure (KIC) refinement moves with soft repulsive forces. Sampling was focused on two loop regions (36-44 in Cdc42 and 1365-1370 in ITSN) in a 10 Å radius of the mutated positions. To determine these focus regions in an unbiased fashion, we selected regions that showed the largest conformational variability in initial backbone diversification simulations of the template structure (PDB ID: 1KI1, **Figure S6C**). This selection criterion follows the rationale described in (12) that regions with the largest simulated diversity in initial models are often the regions that show the largest deviation from the native structure. **Figure S6D** shows that many of the models resulting from the intense KIC backbone diversification step moved closer to the *ortho*Cdc42/*ortho*ITSN crystal structure.
4. Intensification and refinement, hard potential. To intensify sampling around low-energy conformations identified in the previous step, the final simulation step employed KIC refinement moves in the same variable regions as in step (3) with hard repulsive forces using vicinity sampling, which restricts the sampled backbone angles to those similar to the input model. 600 input models were selected from the 10,000 decoys from step (3), using the following criteria: The 10,000 decoys were first binned by RMSD to the flexible regions in the template (so only information of the template structure was used, as in a general application the structure of the target, the designed complex, will not be known). Then, to represent the diversity of conformations sampled in step (3), we selected representative decoys from each bin such that the number of selected decoys scales logarithmically with the total number of structures in that bin, always choosing the lowest-energy decoys (**Figure S6E**). The intensification and refinement step adapts both backbone and side chains to remove clashes or unfavorable conformations that may have arisen in the preceding soft remodeling step. Application of KIC moves as well as repacking of the surrounding side chains and minimization of backbone and side chain torsion degrees of freedom resulted in considerably lower Rosetta energies (**Figure S6F**, compare with **Figure S6D**).
5. Clustering. **Figure S6F** shows that, while conformations very close ( $< 1$  Å backbone RMSD) to the crystal structure of the designed complex are sampled, they cannot be distinguished by energy from other sampled conformations. However, clustering the 2400 models resulting from step (4) by the C $\alpha$  RMSD of the two flexible regions (between all pairs of models, again not considering information from the solved structure) clearly identifies a conformation close to the designed crystal structure for two of the six dominant clusters (**Figures S6G, Fig. 3** in the main manuscript).

### Cell-based assays

In all cell-based assays, we transiently transfected NIH 3T3 cells with combinations of plasmids encoding FKBP-ITSN, Cdc42 and Lyn-FRB, as indicated (**Fig. 5** in the main manuscript). ITSN and Cdc42 constructs were additionally tagged with fluorescent proteins (**SI Table S4**) to assess expression levels and localization. Prenylated Cdc42 is expected to be localized to the plasma membrane or bound to Rho-GDI. The ITSN-FKBP construct remained predominantly cytoplasmic until the addition of the small molecule Rapamycin.

We also measured induced morphological changes using the label-free xCELLigence system (Roche). Cells transiently transfected with indicated plasmids adhere to E-plate wells covered with gold electrodes. Morphological changes of cells (due to the activation of Cdc42) cause a change of the electrode impedance, which is detected in real time and displayed as changes of cell index (CI) values. Although small, we observed reproducible differences for the different transfected cognate and non-cognate Cdc42/ITSN pairs that were overall consistent with the cell-based results on Cdc42 activation and morphological changes discussed in the main manuscript. As shown in **SI Figure S8B**, we first observed similar spikes for all samples in the first 30 seconds after the addition of Rapamycin; these spikes were likely dominated by the change of environment since they were still seen when only medium (including DMSO) without rapamycin was added. The most noticeable differences are between 30 seconds and 90 seconds after the addition of Rapamycin. During that time, the signals of presumed negative samples (the non-cognate Cdc42<sup>WT</sup>/*ortho*ITSN pair and transfections without the recruiter domain Lyn-FRB) decreased noticeably. In contrast, the signal for the designed cognate pair *ortho*Cdc42/*ortho*ITSN (with Lyn-FRB) stays high for about 90 seconds, which was also true for the positive control, Cdc42<sup>WT</sup>/ITSN<sup>WT</sup> (with Lyn-FRB). At longer times, the signal for the designed pair appears to decrease faster than that of the wild-type cognate pair. Decreasing signals at longer times could be caused by retraction of cells as seen in **SI Figure S8C**, although we observe similar retraction also for the wild-type pair. As discussed in the main manuscript, we observe a positive signal for the other non-cognate pair *ortho*Cdc42/ITSN<sup>WT</sup>, likely due to the presence of endogenous Cdc42<sup>WT</sup>. Taken together, the XCELLigence results indicate that: First, transfection of Cdc42<sup>WT</sup> has the most significant signal changes when ITSN<sup>WT</sup> was recruited to the membrane, but not with *ortho*ITSN; and second, recruitment of *ortho*ITSN induced more significant signal changes when *ortho*Cdc42 was present than with Cdc42<sup>WT</sup>.

## Supplementary Figures

### Figure S1: Additional modeling and design simulations.

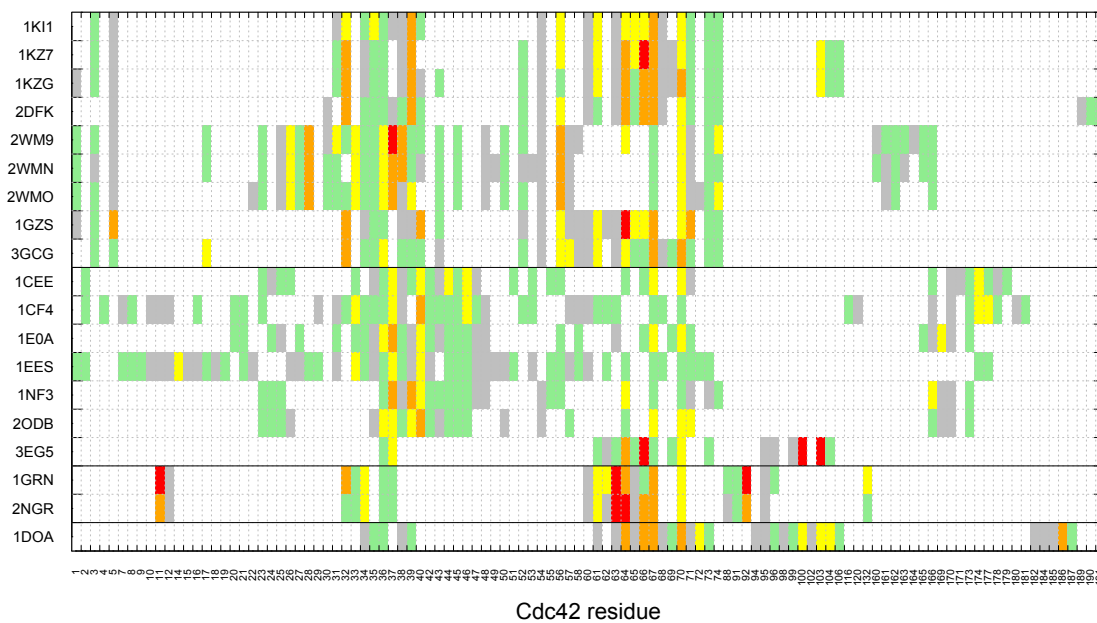
**(A)** Computational alanine scanning. Shown are the estimated effects on binding energy of replacing each residue in Cdc42 with alanine in the context of 19 co-complex structures of Cdc42 with partner proteins. Representation is as shown in Figure 1C in the main manuscript, but results are shown for all Cdc42 residues (instead of just Cdc42 residues in the interface with ITSN taken from PDB ID 1KI1). White blocks mean missing or non-interface residues.

**(B)** Application of fixed backbone computational second site suppressor design, as described in (1). F56 of Cdc42 was computationally mutated to all amino acids (except cysteine) and the effect on complex destabilization was computed (red bars,  $\Delta\text{score (destabilization)} = \text{score (complex Cdc42(mutant)/ITSN}^{\text{WT}}) - \text{score (complex Cdc42}^{\text{WT}}/\text{ITSN}^{\text{WT}})$ ). In a second simulation, the residues on ITSN in the vicinity of position 56 on Cdc42 are designed in the presence of the single mutation on Cdc42 to compensate for the change, and again the effect on the complex binding energy was estimated (blue bars,  $\Delta\text{score (compensation)} = \text{score (complex Cdc42(mutant)/ITSN(designed))} - \text{score (complex Cdc42(mutant)/ITSN}^{\text{WT}})$ ). Amino acids at position 56 (x axis) are ordered by the  $\Delta\Delta\text{score (black bars)} = \Delta\text{score (compensation)} - \Delta\text{score (destabilization)}$ .

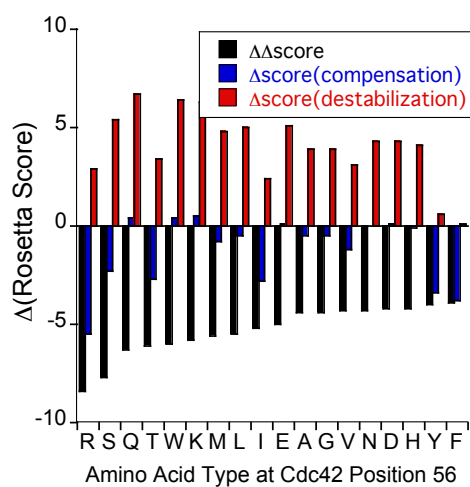
**(C)** Flexible backbone computational design predictions (**Methods**) for the four residues in ITSN neighboring position 56 of Cdc42<sup>WT</sup>. Simulations are exactly as shown in Figure 1D in the main manuscript, except that the Cdc42 does not contain a modeled F56R mutation. Figure was prepared with WebLogo.

**Figure S1**

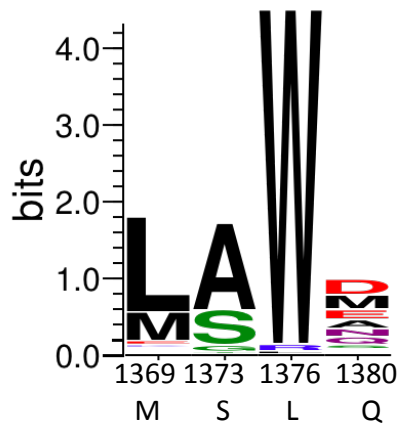
**A**



**B**



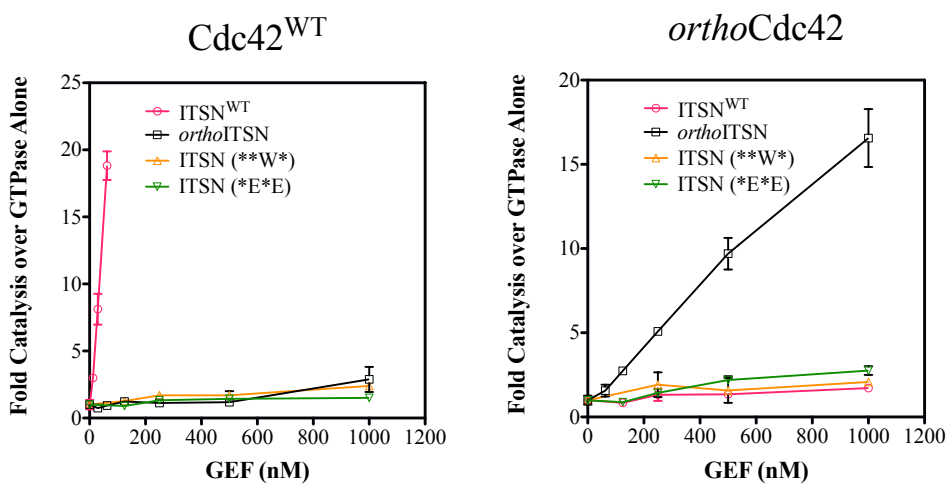
**C**





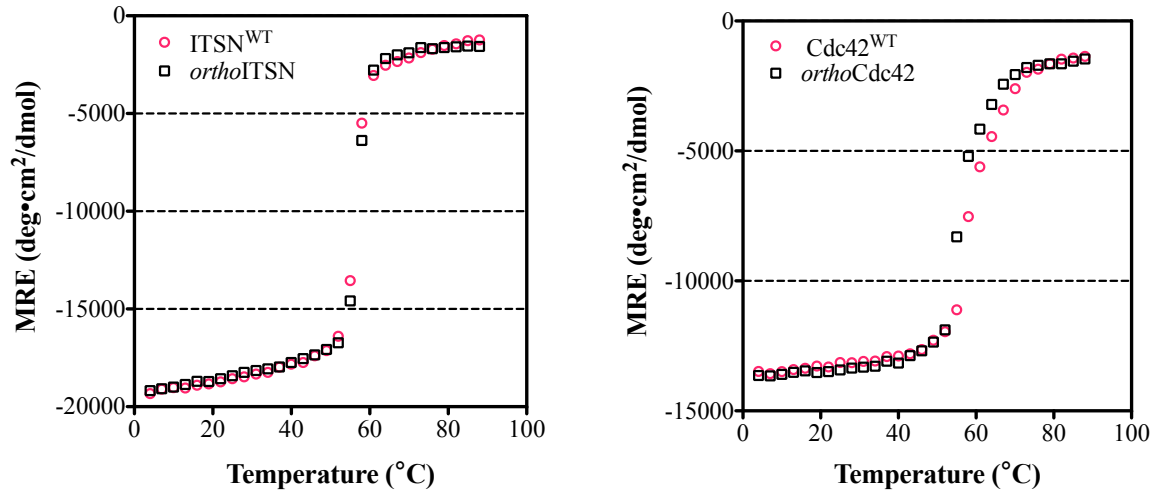
### Figure S2: Nucleotide exchange for additional designed ITSN variants.

Catalysis of nucleotide exchange in Cdc42<sup>WT</sup> (left) and *ortho*Cdc42 (right) by different ITSN variants. ITSN variants were predicted from flexible backbone design simulations (**Figure 1D** in the main manuscript). Shown is the fold increase of initial rates of mant-GDP association to Cdc42 at varying GEF concentrations over Cdc42 alone. Data represent averages and standard deviations from three experiments. Mutation(s) in ITSN: *ortho*ITSN represents S1373E, \*\*W\* L1376W, \*E\*E S1373E-Q1380E.



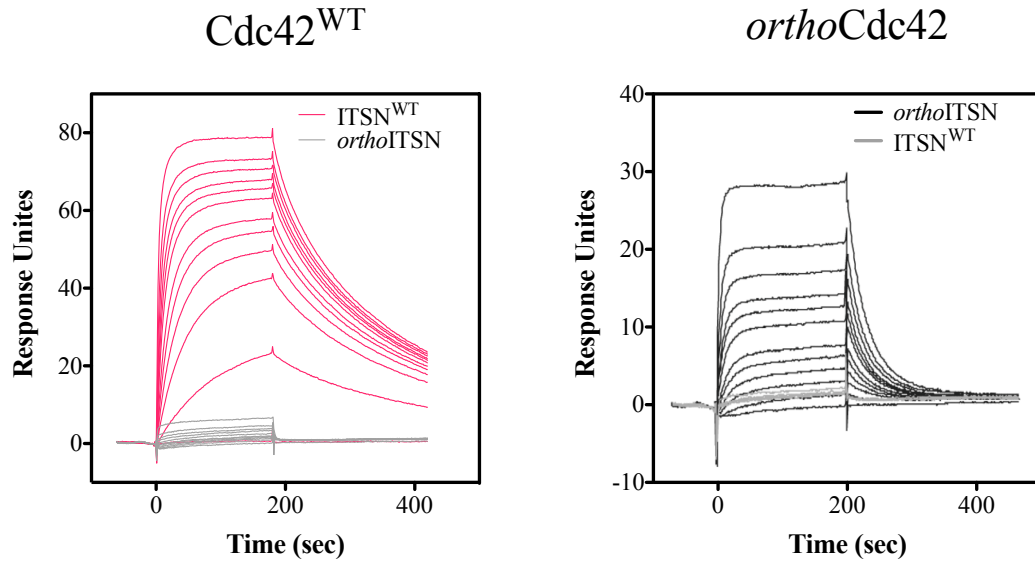
### Figure S3: Thermal denaturation by circular dichroism.

Shown is the temperature dependence of the circular dichroism signal (mean residue ellipticity, MRE) at 222 nm for Cdc42 (left) and ITSN (right). Because thermal melts are irreversible for both proteins, the curves cannot be interpreted in terms of equilibrium denaturation, but indicate similar onset of melting for wild-type (pink symbols) and engineered variants (black symbols).



### Figure S4: Raw biacore sensorgram data.

Shown are raw data from Cdc42<sup>WT</sup> (left) or *ortho*Cdc42 (right) flowing over a cell with immobilized ITSN<sup>WT</sup> (left: pink curves; right: grey curves) and a separate flow cell with immobilized *ortho*ITSN (left: grey curves; right: black curves) for different Cdc42 concentrations (10 nM to 5  $\mu$ M). Response units at saturation for each Cdc42 concentration were used to derive the equilibrium binding data shown in Figure 2C in the main manuscript. Curves for cognate pairs are pink (left) and black (right). Curves for non-cognate pairs are grey.



**Figure S5: Additional analysis of the structure of the designed *ortho*Cdc42/*ortho*ITSN complex.**

(A) The salt bridge interaction between the F56R and S1373E sidechains is revealed in a  $\sigma$ A-weighted simulated annealed composite omit electron density map calculated using the final model and contoured at  $1\sigma$ . ITSN in teal with green side chains, Cdc42 in gray with orange side chains, density contours in yellow.

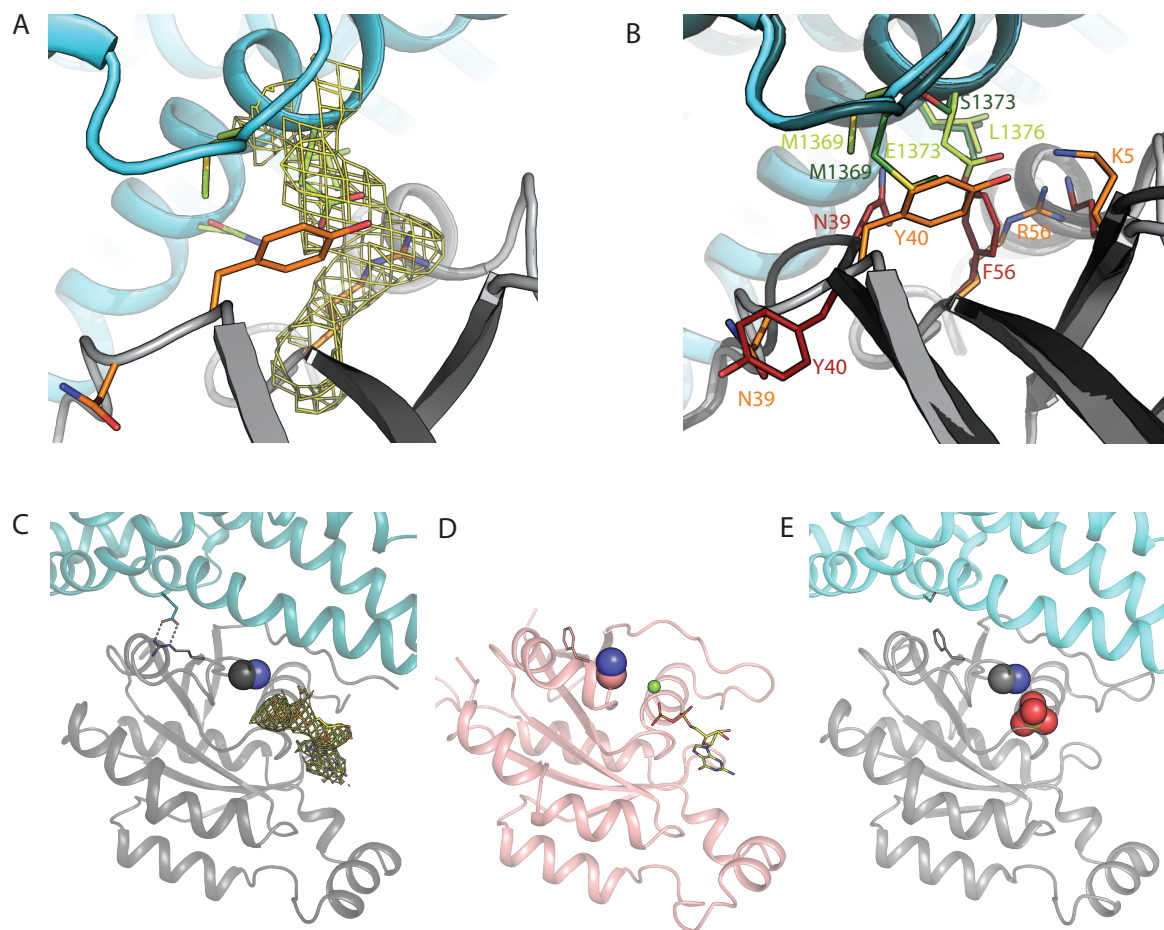
(B) Details of side chain changes, comparing the Cdc42<sup>WT</sup>/ITSN<sup>WT</sup> complex (light gray with orange side chains / teal with light green side chains) to the engineered *ortho*Cdc42/*ortho*ITSN complex (dark gray with firebrick side chains / deep teal with dark green side chains). These changes involve reorganization of a polar interaction network and a flip of the Y40 side chain about the beta strand, which requires a rearrangement of M1369 as this residue occupies the position in the template structure that is taken by Y40 in the designed interface. The loop rearrangement also drastically changes the position of N39. Note that this view is in the same orientation as shown in **Fig. 3** in the main manuscript.

(C) Analysis of the GDP site in the designed *ortho*Cdc42/*ortho*ITSN complex (Green: *ortho*ITSN; dark grey: *ortho*Cdc42). The interface design mutations (S1373E, green; F56R, grey) are shown in stick representation with hydrogen bonds as dashed lines. GDP (yellow) is also shown in stick representation surrounded by electron density from a  $\sigma$ A-weighted kicked map contoured at  $1\sigma$ . The C $\beta$  atom of A59 is shown as a dark blue sphere and would clash with a Mg<sup>2+</sup> ion in the binding site.

(D) The binding site of the Cdc42<sup>WT</sup> structure (PDB ID: 1AN0, pink) is occupied by a GDP molecule (yellow sticks) and a Mg<sup>2+</sup> ion (represented as a small green sphere). The interface residue F56 is shown in sticks. The C $\beta$  atom of the A59 side chain is shown as a dark blue sphere and does not clash with the Mg<sup>2+</sup> ion.

(E) In the Cdc42<sup>WT</sup>/ITSN<sup>WT</sup> structure (PDB ID: 1KI1, cyan: ITSN<sup>WT</sup>; light grey; Cdc42<sup>WT</sup>), the C $\beta$  atom of A59 (shown as a dark blue sphere) would clash with a Mg<sup>2+</sup> ion in the binding site. The interface residues F56 and S1373 are shown in sticks. A sulfate ion (red spheres) is observed at a location that would occupy the position of the beta-phosphate of a GDP in the binding site. The presence of the SO<sub>4</sub><sup>2-</sup> ion suggests that even wild type Cdc42 can partially accommodate GDP-like ligands when bound to a GEF.

**Figure S5**

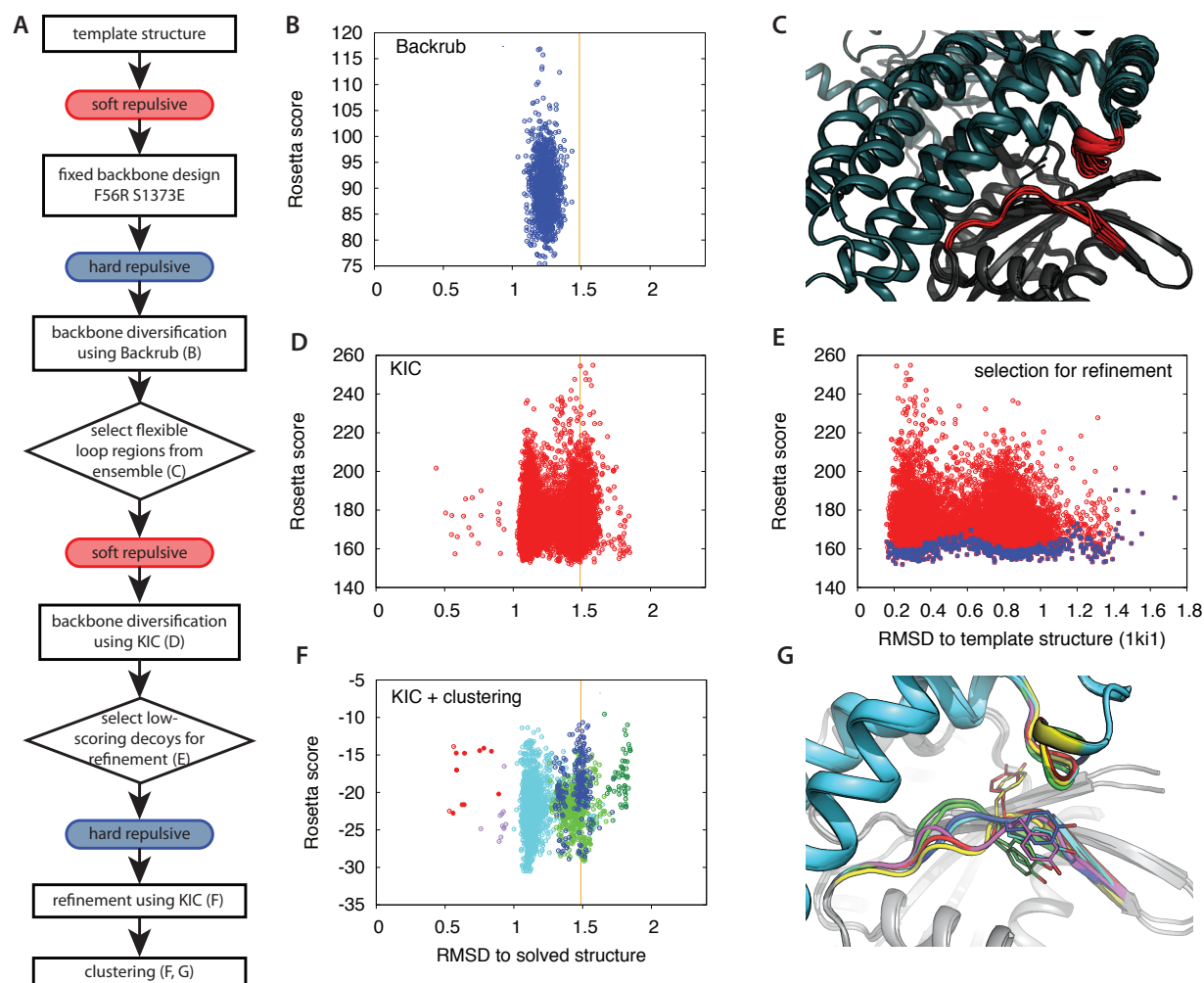




**Figure S6: Flexible backbone design and structure refinement strategy.**

(A) Flowchart of the general remodeling strategy implemented here, highlighting the alternating use of soft and hard repulsive forces, and indicating which panels provide details of simulation results for *orthoCdc42/orthoITSN*. Data in (B), (D) and (F) show the Rosetta full-atom energy versus the RMSD of the remodeled regions to the solved design structure. The RMSD between template and solved design structures is indicated by an orange line. (We note that the RMSD to the design structure was not used in selection of models or clustering). (B) Results of the initial backrub diversification simulations (step (2) in **SI Results**), with RMSDs that remain relatively close to that of the template. (C) Backrub ensemble of the template structure, from which the regions (red) sampled in the diversification simulations (step (3) in **SI Results**) were derived. Cdc42 in gray, ITSN in teal. (D) and (E) show the 10,000 decoys generated with soft repulsive KIC loop modeling on the selected flexible regions (step (3) in **SI Results**), with the RMSD to the solved structure in (D) and the RMSD to the template structure in (E). Panel (E) also highlights the decoys selected for further refinement (blue) following the log-scaling selection procedure described in **SI Results** step (4). (F) Results of hard repulsive vicinity KIC sampling (step (4) in **SI Results**), colored by the cluster each decoy was assigned to. Filled circles indicate that Y40 was in a similar conformation as observed in the crystal structure of the design (see panel G). Note that this conformation of the tyrosine side chain was only observed in cases with a low backbone RMSD with respect to the solved design structure. (G) Lowest-scoring members of each cluster (colors as in F, solved design structure in yellow), showing the position of the Y40 side chain. Cdc42 in gray, ITSN in teal.

**Figure S6**

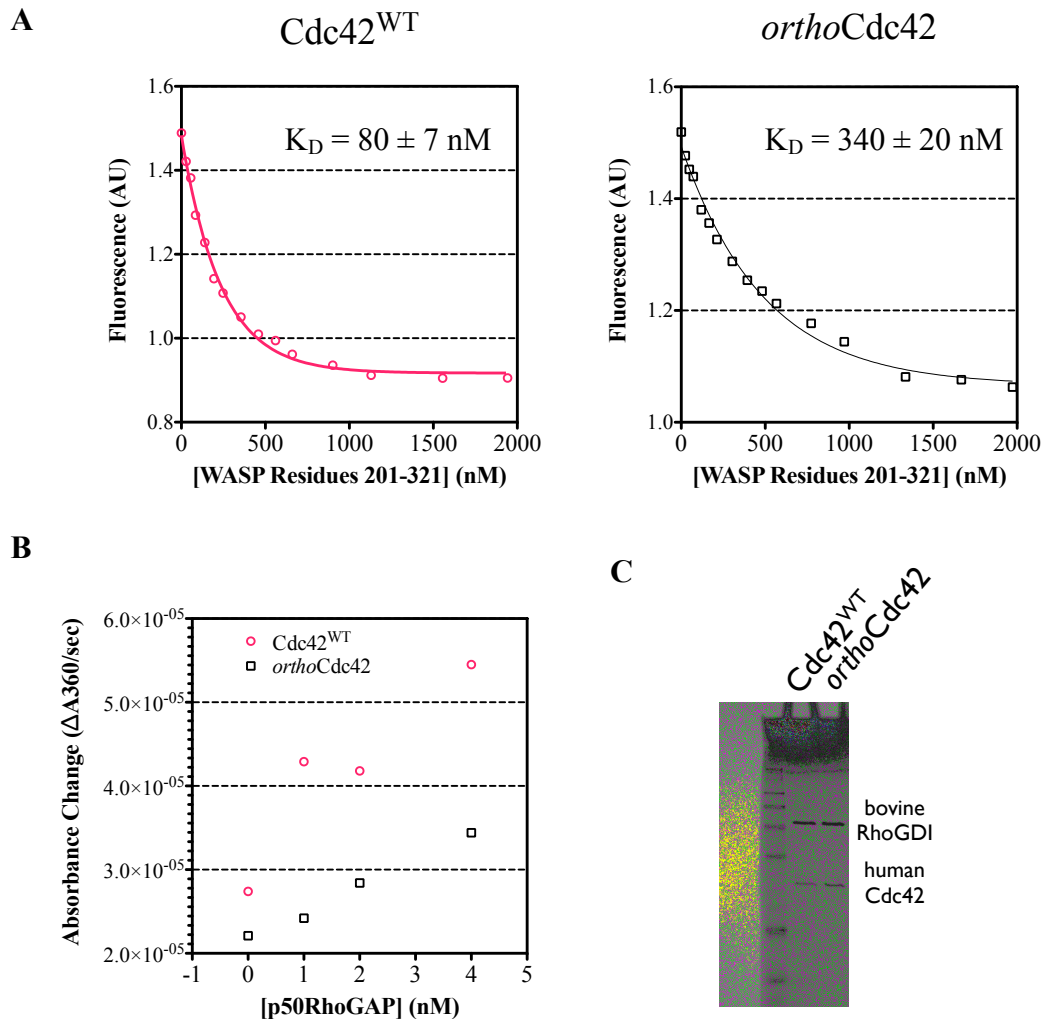


**Figure S7: Interaction of *ortho*Cdc42 with other GTPase circuit components.**

**(A)** Binding of WASP (residues 201-321) to Cdc42<sup>WT</sup> (left) and *ortho*Cdc42 (right) loaded with mant-GNPPNP, monitored by fluorescence quenching of mantGNPPNP by WASP.

**(B)** Rates of GTP hydrolysis of Cdc42 catalyzed by p50RhoGAP, as monitored by free phosphate release.

**(C)** Rho-GDI interaction. Cdc42 and RhoGDI form a stable complex and can be co-purified. Prenylated, His-tagged Cdc42 (the WT or F56R variant) were expressed in SF9 cells. GST-tagged RhoGDI was expressed in *E. coli*. SF9 and *E. coli* lysates were mixed and the Cdc42•RhoGDI complexes were purified using a Ni-NTA column followed by a GST-agarose column. Purified complexes were then run on an SDS-PAGE gel to verify that both the Cdc42 and RhoGDI proteins are present.

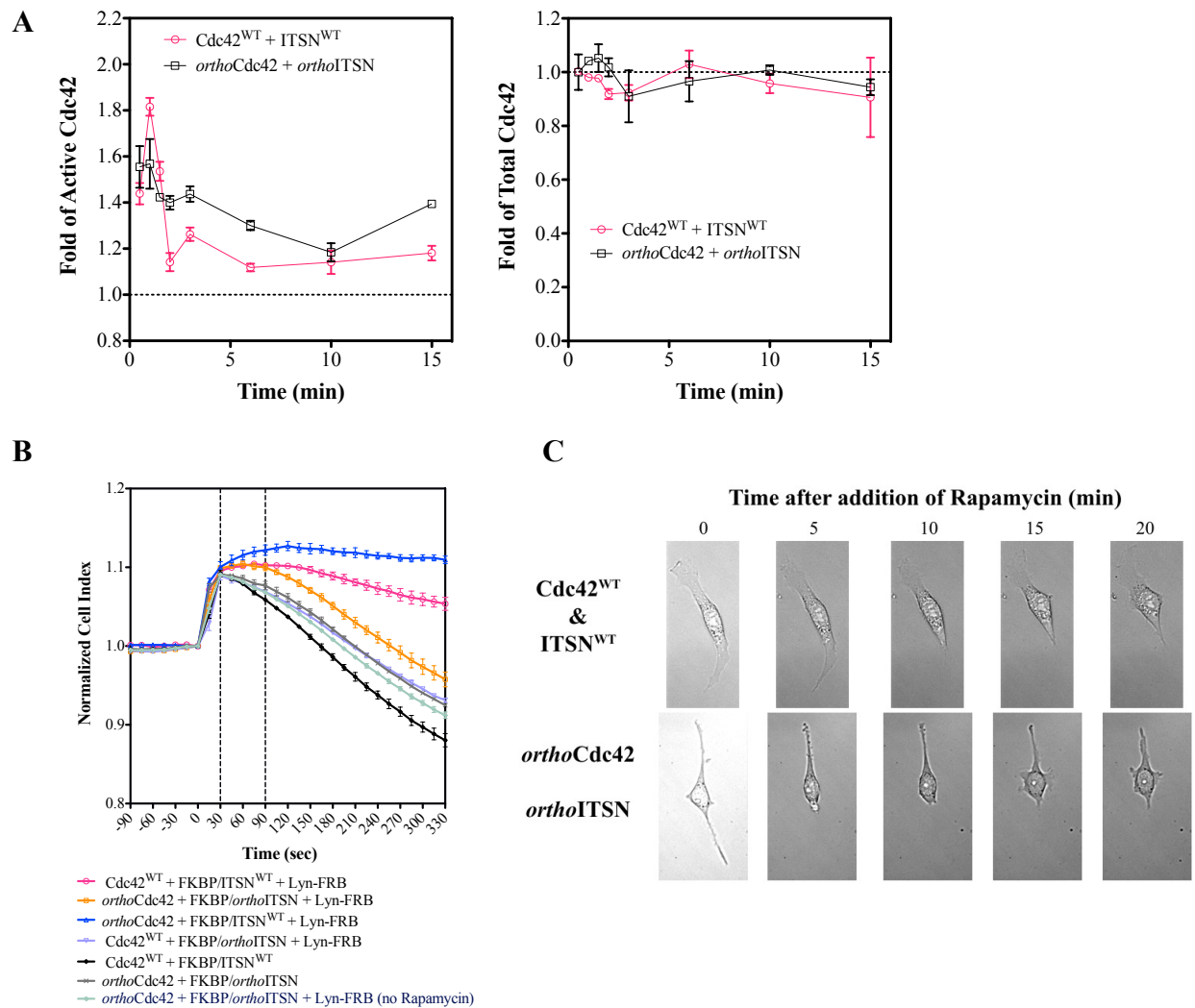


## Figure S8: Additional cell-based assays.

(A) Time-course of Cdc42 activation after Rapamycin addition monitored by the G-LISA assay (Methods), as in Figure 5B in the main manuscript (left). Shown is the fold increase comparing samples with and without addition of Rapamycin at the indicated time. The total Cdc42 loaded in the G-LISA assay was determined by using an ELISA assay (right). Error bars represent the standard deviation of three experiments.

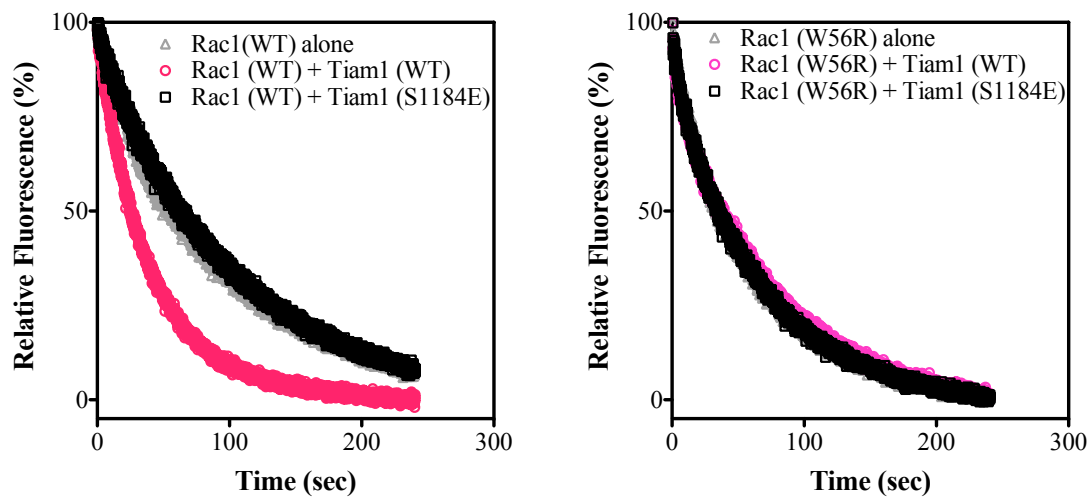
(B) Cell morphological changes monitored by the XCELLigence system (see SI Methods and SI Results). The normalized cell index reflects the measured change in impedance caused by changes in cell shape. Rapamycin was added at the 0-second time point. Error bars represent the standard deviation of three experiments.

(C) Cell retraction was observed as an additional phenotype under our experimental condition after the addition of Rapamycin.



### Figure S9: Transferability of the designed R-E interaction to the Rac1/Tiam1 interface.

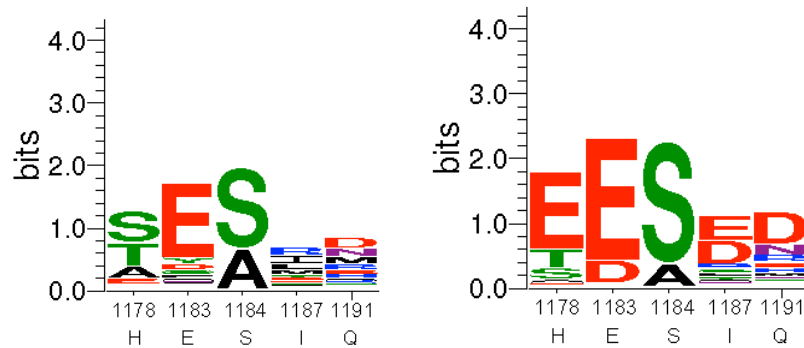
Transferring the designed substitutions from *orthoCdc42/orthoITSN* to the equivalent positions in the Rac1-Tiam1 interface does not result in the same pattern of designed orthogonality. The left graph shows mantGDP dissociation from the GTPase Rac1<sup>WT</sup> in the absence of any exchange factor (gray open triangles) and in the presence of the wild-type exchange factor Tiam1<sup>WT</sup> (pink open circles) and the designed exchange factor Tiam1\* (S1184E, using PDB numbering from PDB ID 1FOE) (black open squares). The right graph shows dissociation from the designed GTPase Rac1\* (W56R) alone (grey open triangles) and in the presence of the same two exchange factor proteins.





**Figure S10: Transferability of the designed R-E interaction to the Rac1/Tiam1 interface.**

Flexible backbone computational design predictions (**Methods**) for the five residues in Tiam1 (H1178, E1183, S1184, I1187 and Q1191) close to position 56 of Rac1 for Rac1<sup>WT</sup> (left) and Rac1\* (W56R) (right). Simulations are as shown in Figure 1D in the main manuscript, except that the backbone of the crystal structure of Rac1/Tiam1 (PDB ID: 1FOE (16)) was used as the starting conformation to create a backrub ensemble. Position S1184 is not enriched for glutamate.



## Supplementary Tables

**Table S1: Summary of designed Cdc42 and ITSN variants tested for orthogonality.**

Summary of the ability of different ITSN variant to catalyze nucleotide exchange in Cdc42<sup>WT</sup> or *ortho*Cdc42. Exchange activity was determined by mantGDP dissociation assays (<sup>a</sup>), mantGDP association assays (<sup>b</sup>), or both association and dissociation (<sup>c</sup>).

	Exchange activity with		Rationale
ITSN variant	Cdc42 <sup>WT</sup>	<i>ortho</i> Cdc42	
WT	+ <sup>c</sup>	– <sup>c</sup>	control
S1373E ( <i>ortho</i> ITSN)	– <sup>c</sup>	+ <sup>c</sup>	single substitution designed to form specific interaction with F56R in <i>ortho</i> Cdc42
M1369L, S1373E, L1376I	– <sup>a</sup>	+ <sup>a</sup>	additional substitutions predicted to be favorable at the <i>ortho</i> ITSN/ <i>ortho</i> Cdc42 site: M1369L: designed substitution (Fig. 1D) L1376I: substitution in ITSN homologs
S1373E, Q1380E	– <sup>b</sup>	– <sup>b</sup>	additional Q1380E substitution predicted (incorrectly) by the fixed backbone design protocol (Fig. 1D) to stabilize the interaction with F56R in <i>ortho</i> Cdc42
L1376W	– <sup>b</sup>	– <sup>b</sup>	substitution predicted (incorrectly) to stabilize the interaction of ITSN with Cdc42 <sup>WT</sup> (Fig. S1C)

**Table S2: Crystallographic data and refinement statistics for the structure of the designed *ortho*Cdc42/*ortho*ITSN interaction.**

	<i>ortho</i> Cdc42/ <i>ortho</i> ITSN
<b>Data collection</b>	
Resolution	50-2.65 (2.74-2.65)
Wavelength	1.115872
Space Group	P2 <sub>1</sub>
Cell angles	85.460 80.062 94.591
Cell edges	90.00 108.23 90.00
I/ $\sigma$	20.0 (2.7)
R <sub>sym</sub>	6.9% (41.2%)
Completeness	99.6% (97%)
Redundancy	3.9 (3.7)
Observed reflections	139536 (12758)
Unique reflections	35386 (3394)
<b>Refinement</b>	
Resolution	46-2.65
Reflections	35295
Free reflections	3499
R <sub>work</sub>	24.1
R <sub>free</sub>	28.4
<b>Structure</b>	
Number of atoms:	7073
Protein	6897
Solvent	120
Ligand	56
Average B-factor:	54.6
Protein	54.5
Solvent	43.0
Ligand	79.7
RMS Bonds	0.006
RMS Angles	1.001
Residues with Favored Ramachandran Angles	96.0%
Residues with Outlier Ramachandran Angles	0.0

Values in parentheses are for the highest resolution shell.

**Table S3: Summary of interactions of *ortho*Cdc42 and *ortho*ITSN with other GEFs and GTPases.**

Top table: summary of nucleotide exchange catalysis of a range of exchange factors for Cdc42<sup>WT</sup> and *ortho*Cdc42. Bottom table: exchange activity of ITSN<sup>WT</sup> and *ortho*ITSN for different GTPases. Exchange activity was determined by mantGDP dissociation (<sup>a</sup>) or both association and dissociation (<sup>b</sup>). Data for RhoG were obtained with the DH domain of ITSN<sup>WT</sup> and *ortho*ITSN only (<sup>c</sup>); all other assays used the DH-PH domain construct for ITSN.

GEF	Cdc42 <sup>WT</sup>	<i>ortho</i> Cdc42
ITSN <sup>WT</sup>	+ <sup>b</sup>	- <sup>b</sup>
<i>ortho</i> ITSN	- <sup>b</sup>	+ <sup>b</sup>
PREX	+ <sup>a</sup>	+ <sup>a</sup>
Tiam1	- <sup>a</sup>	- <sup>a</sup>
TrioN	- <sup>a</sup>	- <sup>a</sup>
Dbp	+ <sup>a</sup>	- <sup>a</sup>

GTPase	ITSN <sup>WT</sup>	<i>ortho</i> ITSN
Cdc42 <sup>WT</sup>	+ <sup>b</sup>	- <sup>b</sup>
<i>ortho</i> Cdc42	- <sup>b</sup>	+ <sup>b</sup>
Rac1	- <sup>a</sup>	- <sup>a</sup>
RhoA	- <sup>a</sup>	- <sup>a</sup>
RhoG <sup>c</sup>	- <sup>a</sup>	- <sup>a</sup>

**Table S4: All constructs used in the *in vitro* and cell-based assay**

Protein	Residues	Mutation	Origin	Vector	Digestion Sites	Resistance	Host	Usage	Ref/Note
Lyn-FRB	NA	--	--	pC4RHE	--	Amp	NIH 3T3	Cell-based assays	(17)
ITSN (DH-PH)	1229-1580	--	Human	YF-Rac1	EcoRI, BglII (ITSN)/BamHI(vector)	Kan	NIH 3T3	Cell-based assays	(17)
ITSN (DH-PH)	1229-1580	S1373E	Human	YF-Rac1	EcoRI, BglII (ITSN)/BamHI(vector)	Kan	NIH 3T3	Cell-based assays	(17)
Cdc42	1-191	--	Human	pAmCyan1-C1	BamHI, EcoRI	Kan	NIH 3T3	Cell-based assays	This study
Cdc42	1-191	F56R	Human	pAmCyan1-C1	BamHI, EcoRI	Kan	NIH 3T3	Cell-based assays	This study
ITSN (DH-PH)	1229-1580	1373E, 1380E	Human	pBY601	--	Amp	<i>E. coli</i>	Expression	This study
ITSN (DH-PH)	1229-1580	1373E	Human	pBY601	--	Amp	<i>E. coli</i>	Expression	This study
ITSN (DH-PH)	1229-1580	1376W	Human	pBY601	--	Amp	<i>E. coli</i>	Expression	This study
ITSN (DH-PH)	1229-1580	--	Human	pSH200	--	Amp	<i>E. coli</i>	Expression	Lim Lab
Cdc42	1-179	F56R	Human	pBH4	--	Amp	<i>E. coli</i>	Expression	This study
Cdc42	1-179	--	Human	pBH4	--	Amp	<i>E. coli</i>	Expression	Lim Lab
RhoGDI	1-204	C79S	Bovine	pJT159-pGEX-KG-GDI	--	Amp	<i>E. coli</i>	Expression	(18)
WASP	201-321	--	Human	pGEX-2T	--	Amp	<i>E. coli</i>	Expression	(9)
p50RhoGAP	206-439	--	Human	pBY614	--	--	<i>E. coli</i>	Expression	Lim Lab
Cdc42	1-191	--	Human	pFastBac-B	BamHI	Amp	SF9	Expression	This study
Cdc42	1-191	F56R	Human	pFastBac-B	BamHI	Amp	SF9	Expression	This study
Rac1	1-177	F78S	Human	pSH200	--	Amp	<i>E. coli</i>	Expression	Lim Lab
Rac1	1-177	W56R, F78S	Human	pSH200	--	Amp	<i>E. coli</i>	Expression	This study
RhoA	1-190	--	Human	pAD15	--	Amp	<i>E. coli</i>	Expression	Lim Lab
RhoG	1-188	--	Murine	pAD18	--	Amp	<i>E. coli</i>	Expression	Lim Lab
Tiam1	1033-1406	--	Murine	pAD1	--	Amp	<i>E. coli</i>	Expression	Lim Lab
Tiam1	1033-1406	S1184E	Murine	pAD1	--	Amp	<i>E. coli</i>	Expression	This study
PREX	63-402	--	Human	--	--	--	<i>E. coli</i>	Expression	Lim Lab
Tim	1166-1527	--	Human	pAD24	--	Amp	<i>E. coli</i>	Expression	Lim Lab
TrioN	1284-1595	--	Human	pAD2	--	Amp	<i>E. coli</i>	Expression	Lim Lab
Dbp	623-967	--	Human	--	--	Amp	<i>E. coli</i>	Expression	(19)

Lyn-FRB and YF-Rac1 were gifts from the Meyer lab. pAmCyan-C1 was bought from Clontech. pFastBac-B was purchased from Invitrogen. The DNA for the WASP fragment was graciously provided by the Wittinghofer laboratory.

## Supplementary References

1. Kortemme T, *et al.* (2004) Computational redesign of protein-protein interaction specificity. *Nat Struct Mol Biol* 11(4):371-379.
2. Smith CA & Kortemme T (2010) Structure-based prediction of the peptide sequence space recognized by natural and synthetic PDZ domains. *J Mol Biol* 402(2):460-474.
3. Pace CN, Vajdos F, Fee L, Grimsley G, & Gray T (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* 4(11):2411-2423.
4. Yeh BJ, Rutigliano RJ, Deb A, Bar-Sagi D, & Lim WA (2007) Rewiring cellular morphology pathways with synthetic guanine nucleotide exchange factors. *Nature* 447(7144):596-600.
5. Minor W, Cymborowski M, Otwinowski Z, & Chruszcz M (2006) HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr* 62(Pt 8):859-866.
6. Trapani S & Navaza J (2008) AMoRe: classical and modern. *Acta Crystallogr D Biol Crystallogr* 64(Pt 1):11-16.
7. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486-501.
8. Adams PD, *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213-221.
9. Rudolph MG, *et al.* (1998) The Cdc42/Rac interactive binding region motif of the Wiskott Aldrich syndrome protein (WASP) is necessary but not sufficient for tight binding to Cdc42 and structure formation. *J Biol Chem* 273(29):18067-18076.
10. Zhang B, Wang ZX, & Zheng Y (1997) Characterization of the interactions between the small GTPase Cdc42 and its GTPase-activating proteins and putative effectors. Comparison of kinetic properties of Cdc42 binding to the Cdc42-interactive domains. *J Biol Chem* 272(35):21999-22007.
11. Yu N, *et al.* (2006) Real-time monitoring of morphological changes in living cells by electronic cell sensor arrays: an approach to study G protein-coupled receptors. *Anal Chem* 78(1):35-43.
12. Qian B, *et al.* (2007) High-resolution structure prediction and the crystallographic phase problem. *Nature* 450(7167):259-264.
13. Smith CA & Kortemme T (2008) Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction. *J Mol Biol* 380(4):742-756.
14. Mandell DJ, Coutsiar EA, & Kortemme T (2009) Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nature methods* 6(8):551-552.
15. Khatib F, *et al.* (2011) Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences of the United States of America* 108(47):18949-18953.
16. Worthylake DK, Rossman KL, & Sondek J (2000) Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1. *Nature* 408(6813):682-688.
17. Inoue T, Heo WD, Grimley JS, Wandless TJ, & Meyer T (2005) An inducible translocation strategy to rapidly activate and inhibit small GTPase signaling pathways. *Nat Methods* 2(6):415-418.
18. Co C, Wong DT, Gierke S, Chang V, & Taunton J (2007) Mechanism of actin network attachment to moving membranes: barbed end capture by N-WASP WH2 domains. *Cell* 128(5):901-913.
19. Rossman KL, *et al.* (2002) A crystallographic view of interactions between Dbs and Cdc42: PH domain-assisted guanine nucleotide exchange. *EMBO J* 21(6):1315-1326.