Title: Complete mutational mapping of a GTPase switch in vivo reveals novel allosteric regulation

Authors: Christopher J.P. Mathy[1,2,3], Parul Mishra[4], Julia M. Flynn[4], Tina Perica[1,2], David Mavor[4], Daniel N.A. Bolon[4]*, Tanja Kortemme[1,2,3,5]*

Affiliations:

1. Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco; San Francisco, CA, USA.
2. Quantitative Biosciences Institute, University of California, San Francisco; San Francisco, CA, USA.
3. The UC Berkeley-UCSF Graduate Program in Bioengineering, University of California, San Francisco; San Francisco, CA, USA.
4. Department of Biochemistry and Molecular Biotechnology, University of Massachusetts Medical School; Worcester, MA, USA
5. Chan Zuckerberg Biohub; San Francisco, CA, USA.

*Corresponding authors. Emails: Dan.Bolon@umassmed.edu (D.N.B); kortemme@cgl.ucsf.edu (T.K.)

Abstract: GTPases are highly regulated molecular switches that control signaling. Switching occurs via catalyzed GTP hydrolysis and nucleotide exchange, yet the importance of additional regulation for biological function - suggested by high GTPase conservation and centrality in protein networks - is unclear. Here we map the function of each residue in the central GTPase Gsp1/Ran in its native biological network by deep mutagenesis. 28% of 4315 assayed mutations in Gsp1/Ran show a toxic/gain-of-function (GOF) response, and 20/60 positions enriched for GOF mutations are outside the GTPase active site. Kinetic analysis shows that these sites are allosterically coupled to the active site, identifying distinct mechanisms that alter Gsp1/Ran cellular function by modulating GTPase switching. Our systematic discovery of new regulatory sites provides a functional map relevant to other GTPases such as Ras.

Main Text:

Protein switches that cycle between "on" and "off" states in response to signals are ubiquitous in biological regulation (*1*), and their misregulation is often associated with disease (*2*). In GTPases switches, interconversion between a GTP-bound on-state and a GDP-bound off-state is intrinsically slow but is accelerated by two opposing regulators: GTPase-activating (GAP) proteins that activate GTP hydrolysis and guanine nucleotide exchange factor (GEF) proteins that accelerate nucleotide replacement. While classic studies have focused on misregulation that renders GTPase switches either constitutively "on" or "off" (*3*), more recent work on GTPase mutations observed in cancer (*4*) or located in interaction interfaces with partner proteins (*5*) has suggested that cellular functions can be sensitive to even small perturbations of the kinetics of switching (*5*, *6*). Moreover, perturbations at a limited number of sites in GTPases distal from the active site (allosteric) can affect biochemical function in vitro (*5*), lead to switch overactivation (*7*, *8*) or alter cellular functions (*5*), and such allosteric sites can be targeted by small molecule inhibitors (*9*). Yet, the vast majority of prior approaches have focused on characterizing known active site regions in GTPases comprising the nucleotide binding region and the switch loops (*10*), and it is hence unclear whether additional mechanisms exist that use regulatory sites outside of these regions to control cellular GTPase function. Such mechanisms could exist since regions outside the GTPases active site are highly conserved, form diverse protein-protein interactions in cellular networks, and are targets of posttranslational modifications. However, discovering new allosteric sites relevant in vivo requires a systematic approach that studies a GTPase switch in its native biological network (*11*) that preserves the functional context of opposing regulators, posttranslational modifications, interaction partners and downstream signaling pathways (**Fig. 1A**).

Here we use deep mutational scanning (DMS) (*12, 13*) to map the function of every position in the GTPase switch Gsp1/Ran in the native context of its in vivo interaction network in *S. cerevisiae*. Gsp1/Ran uses a single pair of regulators, the GAP Rna1 and the GEF Srm1, but an extended network of adaptor and effector proteins, whose interactions with Gsp1/Ran are dependent on switch state, control diverse processes including nucleocytoplasmic transport, cell cycle progression and RNA processing (*5*). Gsp1 is highly conserved, with 82% of its amino acid sequence identical to the human homolog Ran. With some notable exceptions (*7*, *14*), prior DMS experiments have revealed a tolerance to mutations even among highly conserved proteins (*15*),

suggesting missing biological context (*14*, *16*). In contrast, for Gsp1 in its physiological network, we discover that cellular function is affected by mutations at a large number of previously uncharacterized positions outside the active site, mapping widespread sensitivity of a central GTPase to allosteric regulation.

To systematically measure the effect of all Gsp1 mutations on cellular fitness, we developed a plasmid dropout selection derived from our EMPIRIC (extremely methodical and parallel investigation of randomized individual codons) method (*17*) (**Fig. 1B**, Methods). We transformed a chromosomal GSP1 knockout strain with the wild-type (WT) GSP1 allele under the control of its native promoter on a URA selectable plasmid harboring constitutively expressed GFP, and confirmed Gsp1 protein expression via Western blot (**fig. S1A**). We introduced a library of all possible single Gsp1 mutants, also expressed from the native Gsp1 promoter, using a HIS selectable plasmid harboring constitutively expressed mCherry. We sorted for cells expressing mCherry (library plasmid) but not GFP (WT plasmid) and compared allele abundances from the initial population to the population after six generations of growth to compute fitness scores for all 19 possible single amino acid substitutions as well as WT synonymous (WT-syn) and STOP codons at every position in Gsp1 (**Fig. 1C**, Methods). This approach interrogates variant fitness both in the presence and absence of a WT copy with the potential to inform on both gain of toxic function and loss of normal function.

We categorized the fitness score of each mutation relative to the distributions of fitness scores for WT-syn and STOP codons (**Fig. 1D**, Methods). Compared to the WT-syn distribution, 48.5% of all mutations showed deleterious fitness effects, while very few mutations (15/4315 or 0.35%) were beneficial. We observed strongly deleterious mutations in the GTPase active site, which we define as the highly conserved G1-5 functional regions of the Ras superfamily of small GTPases (including the switch loops that change conformation in the GDP- and GTP-bound states) and any additional positions contacting the nucleotide (**Fig. 1C** and **fig. S2**).

The distribution of STOP codon scores (**Fig. 1E**) fell into two groups: STOP codon mutations before Gsp1 sequence position 175 had narrowly distributed fitness scores no lower than -2.90 (scores are $\log_2$-transformed changes in variant abundance relative to wild-type). In contrast, STOP mutations after position 175 had substantially lower fitness scores (down to -10.5). Residues 1-174 comprise a standard GTPase fold, whereas residues 175-219 comprise a C-terminal extension specific to the Ran subfamily (**Fig. 1C** and **fig. S2A**). Thus, the first set of STOP codon

mutants (residues 1-174) likely represent the growth defect of a null Gsp1 mutant, as internal truncations in the GTPase fold likely result in nonfunctional proteins. Mutations with worse scores than null alleles must have a functional effect more detrimental than loss-of-function, and we termed these mutations "toxic gain-of-function", or toxic/GOF. Using a definition of scores worse than the mean STOP codon mutation score of positions 1-174 by more than three standard deviations, more than half of all deleterious mutations (58.4%, and 28.4% of all mutations) were toxic/GOF. Toxic/GOF mutations were not exclusive to the active site regions defined above, but were broadly distributed across the Gsp1 structure, including in interfaces with Gsp1 partner proteins, in parts of the Gsp1 buried core, and at surface positions outside of the interaction interfaces (**Fig. 2A**).

To identify potential mechanisms underlying the unexpected prevalence of toxic/GOF mutations, we defined sequence positions that were enriched in toxic/GOF mutations. We counted the number of toxic/GOF mutations at each position and compared this empirical distribution to a null distribution parameterized according to the total number of toxic/GOF mutations in the dataset (**fig. S3**, Methods). Positions with 10 or more toxic/GOF mutations showed significant enrichment and were labeled as toxic/GOF positions. In total, 60 out of 219 Gsp1 sequence positions were toxic/GOF, including 57 residues conserved from *S. cerevisiae* Gsp1 to human Ran.

Given most substantial fitness effects observed in DMS studies are typically from mutations at positions in active sites required for function, or at positions in the protein core involved in stability, we asked whether the locations of toxic/GOF positions overlapped with the active site or the core. Only half (30/60) of the toxic/GOF positions are in the active site (**Fig. 2B**, blue) and an additional 10 positions are in the C-terminal extension. Thus, 20/60 toxic/GOF positions are at positions in the GTPase fold but distal to the active site (**Fig. 2B, red**). 16 out of the 46 active site positions are not toxic/GOF. Conversely, only 19 of 60 toxic/GOF positions are in the buried protein core (**Fig. 2C**, red), and 43 out of the 62 core positions are not toxic/GOF (**Fig. 2B**, orange). Moreover, mutations in the active site would typically be expected to ablate function and therefore lead to a loss-of-function phenotype (similar to STOP). However, we observe 517 toxic/GOF mutations in the active site compared to only 93 STOP-like mutations (**Fig. 2A**). Similarly, mutations in the protein core that destabilize Gsp1 would be expected to exhibit a moderate fitness cost as observed for STOP codons in the GTPase fold, but not be toxic/GOF. In addition, computational stability calculations (Methods) showed little correlation between predicted

destabilization and decreased fitness when including toxic/GOF mutations, and only a modest correlation for mutations in the buried core when excluding toxic/GOF mutations (**fig. S4**). Thus, the mechanism of Gsp1 toxic/GOF mutations is not satisfactorily explained by either simply the location in the active site or by destabilization of the protein.

The prevalence of toxic/GOF mutations in the C-terminal extension (**Fig. 1C, E**) provided the first evidence that the toxicity of the mutants stems from perturbed regulation: Deleting the C-terminus of Ran/Gsp1 is known to alter the balance between the switch states by stabilizing the GTP-bound form (*18*) which may explain the enrichment of cancer mutations in the C-terminus of Ran (*19*). We therefore asked whether all toxic/GOF mutations (**Fig. 4A**) perturbed Gsp1 GTPase switch function. This model would account for the toxic/GOF effects of mutations at the 40 positions in the GTPase active site or C-terminus. Of the remaining 20 toxic/GOF distal sites within the GTPase fold (**Fig. 2B**), 13 are located in the interfaces with key regulators of the GTPase switch Rna1 (GAP), Srm1 (GEF), and Yrb1; Y157 is an allosteric site previously identified to be coupled to the Gsp1 active site (*5*), consistent with the proposed effect of mutations on regulated switching; and S155 is a known phosphorylation site (*20*) neighboring the conserved G5 SAK motif in the active site (**Fig. 1C**). Four of the remaining five toxic/GOF positions are clustered in the Gsp1 structure and along with the final position (H50) and two other toxic/GOF positions (N156 and F159) form interaction networks in crystal structures of Ran/Gsp1 that extend to the Switch I and the C-terminal extension in the GDP-bound state (*21*) (**Fig. 3A**). We verified that toxic/GOF mutants at these positions indeed had severe fitness defects compared to WT or an internal STOP-codon mutant when co-expressed with WT using a yeast spotting assay (**Fig. 3B**), and that a C-terminal deletion variant was as toxic as the toxic/GOF mutations at these positions.

To examine whether toxic/GOF mutations perturbed switch function in this unexplained set of mutants, we purified and characterized pairs of toxic/GOF (F28V, F54A, F159L, and F163L) and WT-like mutants (F28Y, F54W, F159W, and F163Y) at the four Phenylalanine positions that are clustered in the structure but distal from the active site. The purified mutants were well-folded and stable (**figs. S5** and **S6)**. We then measured the rate of GEF-mediated nucleotide exchange to either GTP or GDP using recombinantly expressed and purified *S. cerevisiae* Srm1, the GEF of Gsp1 (**Fig. 3C** and Methods). All mutations except F159L had reduced or similar GEF-catalyzed nucleotide exchange rates compared to WT (**fig. S7B**). However, the exchange was dependent on the nucleotide: toxic/GOF mutants had a faster rate of exchange to GTP than to GDP while the

WT-like counterparts had a preference for GDP over GTP, identical to WT (**Fig. 3D**). Hence, toxic/GOF mutations reversed the nucleotide preference but WT-like mutations did not. We also measured GAP-catalyzed GTP hydrolysis and found that toxic/GOF mutations did not have reduced GTP hydrolysis (**fig. S7C**). We conclude that toxic/GOF mutations distal to the active site can indeed allosterically perturb the molecular function of the switch by disfavoring the GDP-bound state, while WT-like mutations at the same positions do not.

Our analyses assign functional roles to all 60 toxic/GOF positions in our dataset (**Fig. 4A**), mapping the functionally essential residues in a GTPase molecular switch (**Fig. 4B**). While the active site (nucleotide recognition sites and the GTPase switch loops) is the most common location for toxic/GOF positions, 33% of toxic/GOF positions (20/60) are outside of the active site (**Fig. 4B**). We identify several mechanisms for how these regions outside of the active site affect GTPase switching: First, 13 sites are in interaction interfaces with the key regulators Rna1 (GAP) and Srm1 (GEF), which accelerate interconversion between the GTP- and GDP-bound states, and Yrb1, the *S. cerevisiae* homolog of human RanBP1, which stabilizes the GTP-bound state of Gsp1 and increases interaction with the GAP (*18*). Second, distal positions in protein-protein interaction interfaces are in addition directly coupled to the switch by modulating the efficiency of GTP hydrolysis (*5*). Third, we show here that a previously unknown allosteric cluster in the structure core (**Fig. 4B**, red) is coupled to switch regulation by altering the nucleotide preference (**Fig. 3**). Finally, the toxic/GOF positions also include 4 locations of posttranslational modifications (PTMs) (*20*, *22*, *23*). Relatively small perturbations at all identified sites resulted in cellular defects consistently worse than a null mutant. Taken together, our data suggest that the effect on the rates of regulated switching between GTPase states is the key quantitative parameter dominating the functional effects of any Gsp1 mutation.

While there are no experimental studies probing the sensitivity of other GTPases under normal cellular conditions at the residue level, the mutational sensitivity of Gsp1 is predictive of many activating mutations recently reported for the human H-Ras protein in mouse-derived Ba/F3 cells (*8*) (**Fig. 4C**). 19/30 positions with activating mutations in H-Ras are also toxic/GOF positions in Gsp1. Those positions are enriched in the active site (**Fig. 4D**), whereas our Gsp1 perturbation analysis revealed additional allosteric sites and positions in regulatory partner interfaces. The additional sites may be specific to Gsp1 or may not be relevant for the overactivation phenotype screened for in the H-Ras assay. We also compared our data to a computational analysis of

GTPases based on residue-residue co-variation in multiple sequence alignments of the GTPase superfamily (*24*). Key "sector" positions identified computationally show more overlap with the Gsp1 toxic/GOF positions than the H-Ras activation data (26/49 aligned positions, versus 19/49), primarily by capturing more residues in the GTPase active site regions (**Fig. 4E)**. However, the computational sector analysis misses 23/49 toxic/GOF positions in Gsp1. This finding could indicate a lack of sensitivity or the potential for key regulatory differences between highly conserved GTPases that may be difficult to discern from sequence analysis alone but which are enabled by quantitative perturbations in the native cellular context.

A key finding of our work is the broad sensitivity of a critical molecular switch to perturbations at many allosteric regulatory sites outside the typically studied active site "switch" regions. We propose a model where this sensitivity of the switch facilitates both its responsiveness to many biological inputs and its output signaling specificity (*5*). We identify an altered switch balance as the common mechanism by which toxic/GOF mutations affect the cellular function of Gsp1. This finding suggests that the GTPase switch balance is finely tuned and that the sensitivity of this balance to mutations at many positions might explain why GTPases are so highly conserved even outside the active site regions. We further show that relatively small perturbations to the switch balance have deleterious functional consequences. This finding is consistent with results from kinetic models of ultrasensitivity, where for switches controlled by opposing regulators (**Fig. 1A**) small changes in the concentration or activity of regulators can result in sharp changes in the fraction of the switch "on" state (*25*). Our study provides insight into the emerging linkage between allosteric regulation of the switch balance, the ultrasensitivity of switches (*25*), and the functional consequences for cellular regulation (*5*). Our residue-level map of a GTPase molecular switch and the discovery of new regulatory sites opens avenues to interrogate and target GTPases controlling many essential biological processes including intracellular transport, cell growth, differentiation, and cell survival.

References and Notes

1. B. N. Kholodenko, Cell-signalling dynamics in time and space. *Nat. Rev. Mol. Cell Biol.* **7**, 165–176 (2006).
2. J. Cherfils, M. Zeghouf, Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiol. Rev.* **93**, 269–309 (2013).
3. H. R. Bourne, D. A. Sanders, F. McCormick, The GTPase superfamily: a conserved switch for diverse cell functions. *Nature*. **348**, 125–132 (1990).
4. L. J. Bugaj, A. J. Sabnis, A. Mitchell, J. E. Garbarino, J. E. Toettcher, T. G. Bivona, W. A. Lim, Cancer mutations and targeted drugs can disrupt dynamic signal encoding by the Ras-Erk pathway. *Science*. **361** (2018), doi:10.1126/science.aao3048.
5. T. Perica, C. J. P. Mathy, J. Xu, G. M. Jang, Y. Zhang, R. Kaake, N. Ollikainen, H. Braberg, D. L. Swaney, D. G. Lambright, M. J. S. Kelly, N. J. Krogan, T. Kortemme, Systems-level effects of allosteric perturbations to a model molecular switch. *Nature*. **599**, 152–157 (2021).
6. N. Jameson, M. Gavagan, J. G. Zalatan, A kinetic mechanism for systems-level behavior in GTPase signaling. *Trends Biochem. Sci.* (2022), doi:10.1016/j.tibs.2022.01.006.
7. P. Bandaru, N. H. Shah, M. Bhattacharyya, J. P. Barton, Y. Kondo, J. C. Cofsky, C. L. Gee, A. K. Chakraborty, T. Kortemme, R. Ranganathan, J. Kuriyan, Deconstruction of the Ras switching cycle through saturation mutagenesis. *Elife*. **6**, 727 (2017).
8. F. Hidalgo, L. M. Nocka, N. H. Shah, K. Gorday, N. R. Latorraca, P. Bandaru, S. Templeton, D. Lee, D. Karandur, J. G. Pelton, S. Marqusee, D. Wemmer, J. Kuriyan, A saturation-mutagenesis analysis of the interplay between stability and activation in Ras. *Elife*. **11** (2022), doi:10.7554/eLife.76595.
9. J. M. Ostrem, U. Peters, M. L. Sos, J. A. Wells, K. M. Shokat, K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature*. **503**, 548–551 (2013).
10. I. R. Vetter, A. Wittinghofer, The guanine nucleotide-binding switch in three dimensions. *Science*. **294**, 1299–1304 (2001).
11. D. Eisenberg, E. M. Marcotte, I. Xenarios, T. O. Yeates, Protein function in the post-genomic era. *Nature*. **405**, 823–826 (2000).
12. D. M. Fowler, S. Fields, Deep mutational scanning: a new style of protein science. *Nat. Methods*. **11**, 801–807 (2014).
13. A. S. Dunham, P. Beltrao, Exploring amino acid functions in a deep mutational landscape. *Mol. Syst. Biol.* **17**, e10305 (2021).
14. M. A. Stiffler, D. R. Hekstra, R. Ranganathan, Evolvability as a function of purifying selection in TEM-1 β-lactamase. *Cell*. **160**, 882–892 (2015).
15. B. P. Roscoe, K. M. Thayer, K. B. Zeldovich, D. Fushman, D. N. A. Bolon, Analyses of the effects of all ubiquitin point mutants on yeast growth rate. *J. Mol. Biol.* **425**, 1363–1377 (2013).
16. R. T. Hietpas, C. Bank, J. D. Jensen, D. N. A. Bolon, Shifting fitness landscapes in response to altered environments. *Evolution*. **67**, 3512–3522 (2013).
17. R. Hietpas, B. Roscoe, L. Jiang, D. N. A. Bolon, Fitness analyses of all possible point mutations for regions of genes in yeast. *Nat. Protoc.* **7**, 1382–1396 (2012).
18. S. A. Richards, K. M. Lounsbury, I. G. Macara, The C Terminus of the Nuclear RAN/TC4 GTPase Stabilizes the GDP-bound State and Mediates Interactions with RCC1, RAN-GAP, and HTF9A/RANBP1∗. *J. Biol. Chem.* **270**, 14405–14411 (1995).
19. J. Zhou, Y. Tan, Y. Zhang, A. Tong, X. Shen, X. Sun, D. Jia, Q. Sun, GEF-independent Ran activation shifts a fraction of the protein to the cytoplasm and promotes cell proliferation. *Molecular Biomedicine*. **1**, 18 (2020).
20. L. J. Holt, B. B. Tuch, J. Villén, A. D. Johnson, S. P. Gygi, D. O. Morgan, Global analysis of

Cdk1 substrate phosphorylation sites provides insights into evolution. *Science*. **325**, 1682–1686 (2009).
21. A. F. Neuwald, N. Kannan, A. Poleksic, N. Hata, J. S. Liu, Ran's C-terminal, basic patch, and nucleotide exchange mechanisms in light of a canonical structure for Rab, Rho, Ras, and Ran GTPases. *Genome Res.* **13**, 673–692 (2003).
22. B. T. Weinert, C. Schölz, S. A. Wagner, V. Iesmantavicius, D. Su, J. A. Daniel, C. Choudhary, Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation. *Cell Rep.* **4**, 842–851 (2013).
23. D. L. Swaney, P. Beltrao, L. Starita, A. Guo, J. Rush, S. Fields, N. J. Krogan, J. Villén, Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. *Nat. Methods*. **10**, 676–682 (2013).
24. O. Rivoire, K. A. Reynolds, R. Ranganathan, Evolution-Based Functional Decomposition of Proteins. *PLoS Comput. Biol.* **12**, e1004817 (2016).
25. A. Goldbeter, D. E. Koshland Jr, An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 6840–6844 (1981).

Acknowledgments: We would like to thank Mark Kelly for advice and training on biophysical experiments; David Lambright for generously providing kinetic analysis software; Frank DiMaio, Hahnbeom Park, and Brandon Frenz for advice on Rosetta ΔΔG calculations; members of the Kortemme lab for discussion; and Guy Riddihough of Life Science Editors for helpful comments and edits on the manuscript. **Funding**: This work was supported by a grant from the National Institutes of Health (R01-GM117189) to T.K. and D.N.A.B.; C.J.P.M. is a UCSF Discovery Fellow; T.K. is a Chan Zuckerberg Biohub investigator. **Author contributions**: C.J.P.M., D.N.A.B. and T.K. identified and developed the core questions. C.J.P.M. performed the majority of the data analysis, with contributions from D.M., T.P., P.M., D.N.A.B. and T.K., and the biophysical experiments. P.M. performed the mutational scanning experiments. J.M.F. performed the spotting assays and western blot experiments. C.J.P.M. and T.K. wrote the manuscript with contributions from the other authors. D.N.A.B. and T.K. oversaw the project. **Competing interests**: The authors declare no competing interests. **Data availability**: All relevant data and code are available in the manuscript, the supplementary materials, or on GitHub at https://github.com/cjmathy/Gsp1_DMS_Manuscript.

Supplementary Materials

Materials and Methods
Figs. S1 to S8
Table S1
Captions for Data S1 to S3
References
Data S1 to S3

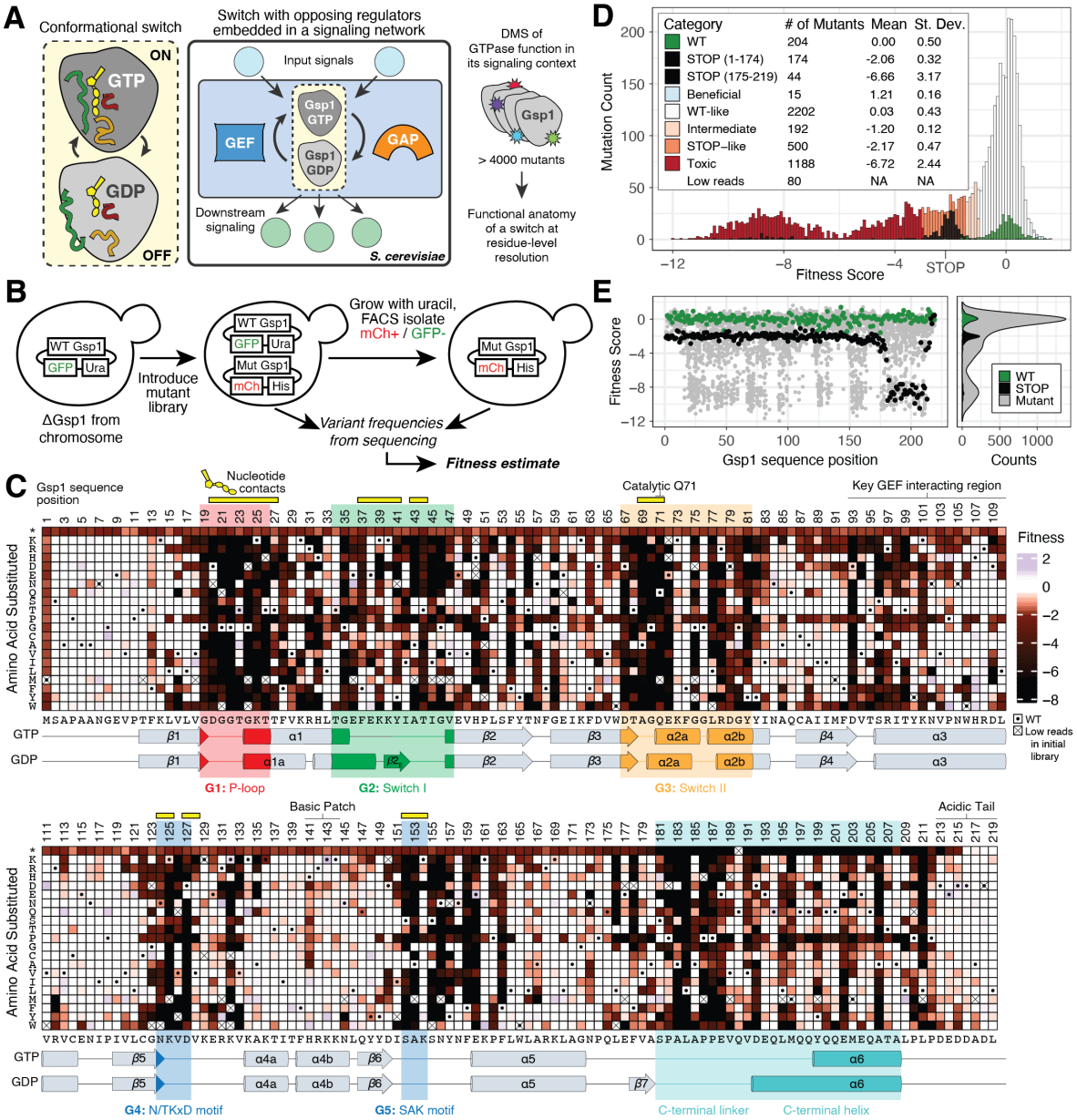

Fig. 1. *In vivo* sensitivity of the GTPase Gsp1 to all possible single amino acid substitutions. **(A)** Deep mutational scanning exhaustively probes the structure-function map of a switch in its signaling context. **(B)** Plasmid dropout selection method to quantify the effects of Gsp1 point mutations on *S. cerevisiae* fitness. **(C)** Heatmap showing quantitative fitness scores for all Gsp1 mutations after 6 generations of competitive growth. Dot indicates WT synonymous codons; X indicates mutants with low reads in the initial library outgrowth. Conserved G1-5 regions are shown in colors corresponding to structural annotations in **fig. S2**. Additional annotated functional regions include the catalytic residue Q71, the GEF interacting region, and the basic patch and acidic tail which interact in the GDP-bound structure (*20*). Positions contacting the nucleotide or magnesium cofactor are indicated by yellow bars. Secondary structure assignments for each position in the GTP- and GDP-bound states are shown below. **(D)** Histogram of scores colored by bin (Methods). Note that 37 of the STOP mutants are toxic/GOF. **(E)** Distribution of fitness scores ordered by Gsp1 sequence position, colored by mutation type: WT synonymous mutations (green), STOP codon mutations (black), and substitutions (gray).

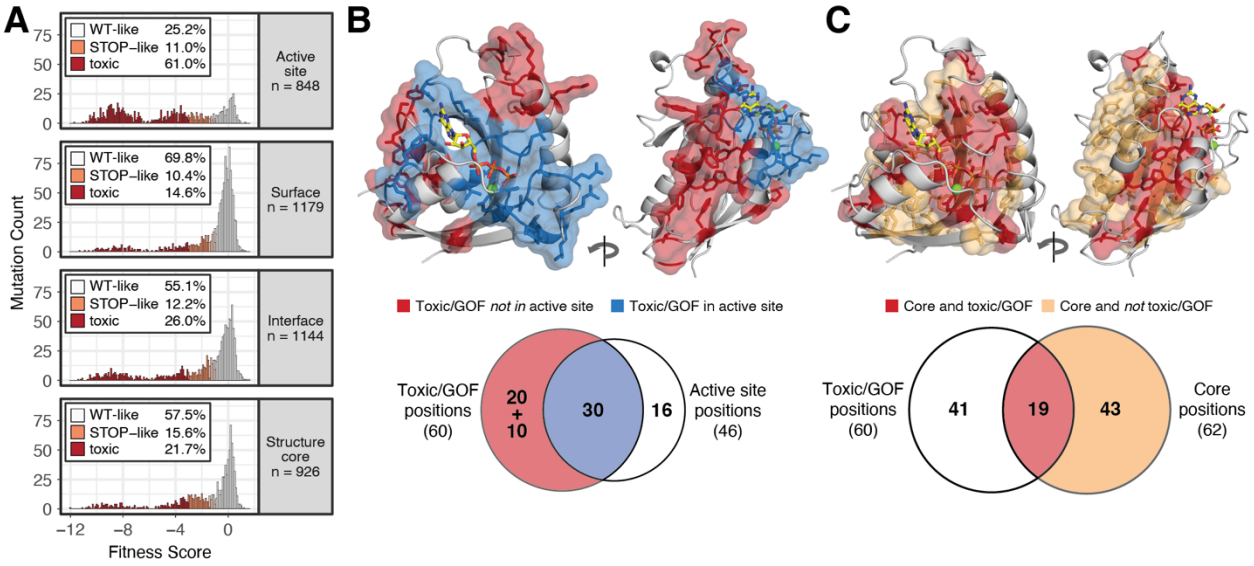

Fig. 2. Locations of toxic/GOF positions outside of the active site. (A) Histograms of fitness scores of mutants by structural regions; colors are as in **Fig. 1D** (showing only point mutations, excluding changes that are WT synonymous or to/from STOP; intermediate and beneficial mutations make up the difference to 100%). Fractions are computed within each structural region; n indicates number of mutations. **(B** and **C)** Gsp1-GTP structure (PDB ID: 3M1I) showing toxic/GOF positions in stick and surface representation (excluding the C-terminal extension). **(B)** Toxic/GOF positions in GTPase active site shown in blue, other toxic/GOF positions shown in red. Venn diagram showing overlap of toxic/GOF positions with GTPase active site positions (10 toxic/GOF positions not shown in the structure are in the C-terminal extension). **(C)** Toxic/GOF core positions shown in red, non-toxic/GOF core positions shown in orange. Venn diagram showing overlap of toxic/GOF positions with structure core positions.

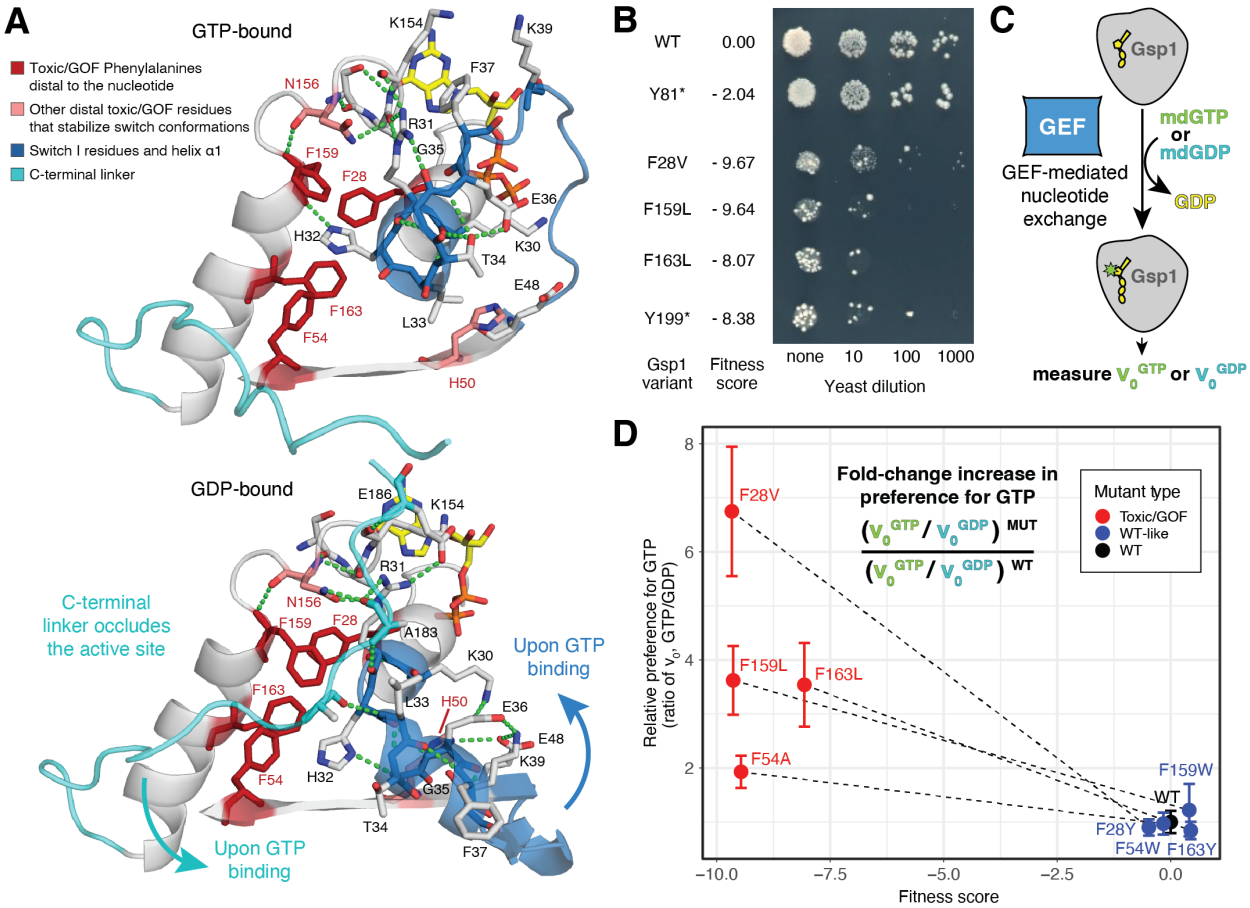

Fig. 3. Distal toxic/GOF mutations alter the balance of the switch states. (A) Structural depiction of the alternate networks of hydrogen-bonding and packing interactions in the GTP-bound (top, PDB ID: 3M1I) and GDP-bound states (bottom, PDB ID: 3GJ0). Toxic/GOF mutants characterized in panels (B) and (D) shown in red. Backbone is colored for the Switch I region (blue) and the C-terminal linker (cyan). The nucleotides are shown in yellow sticks. **(B)** Dilution series for individual Gsp1 variants expressed together with WT in *S. cerevisiae*, with corresponding fitness scores from the EMPIRIC assay. **(C)** FRET-based nucleotide exchange kinetics are measured by adding an excess of mant-labeled fluorescent nucleotide and catalytic amounts of GEF to a sample of purified Gsp1 bound to GDP (Methods). **(D)** Relative change in nucleotide preference for Phe variants, calculated as the ratio of initial rate of exchange to GTP divided by the initial rate of exchange to GDP, normalized to the wild-type ratio. Error bars represent the standard deviations of v_0 measurements propagated across the division operator (Methods).

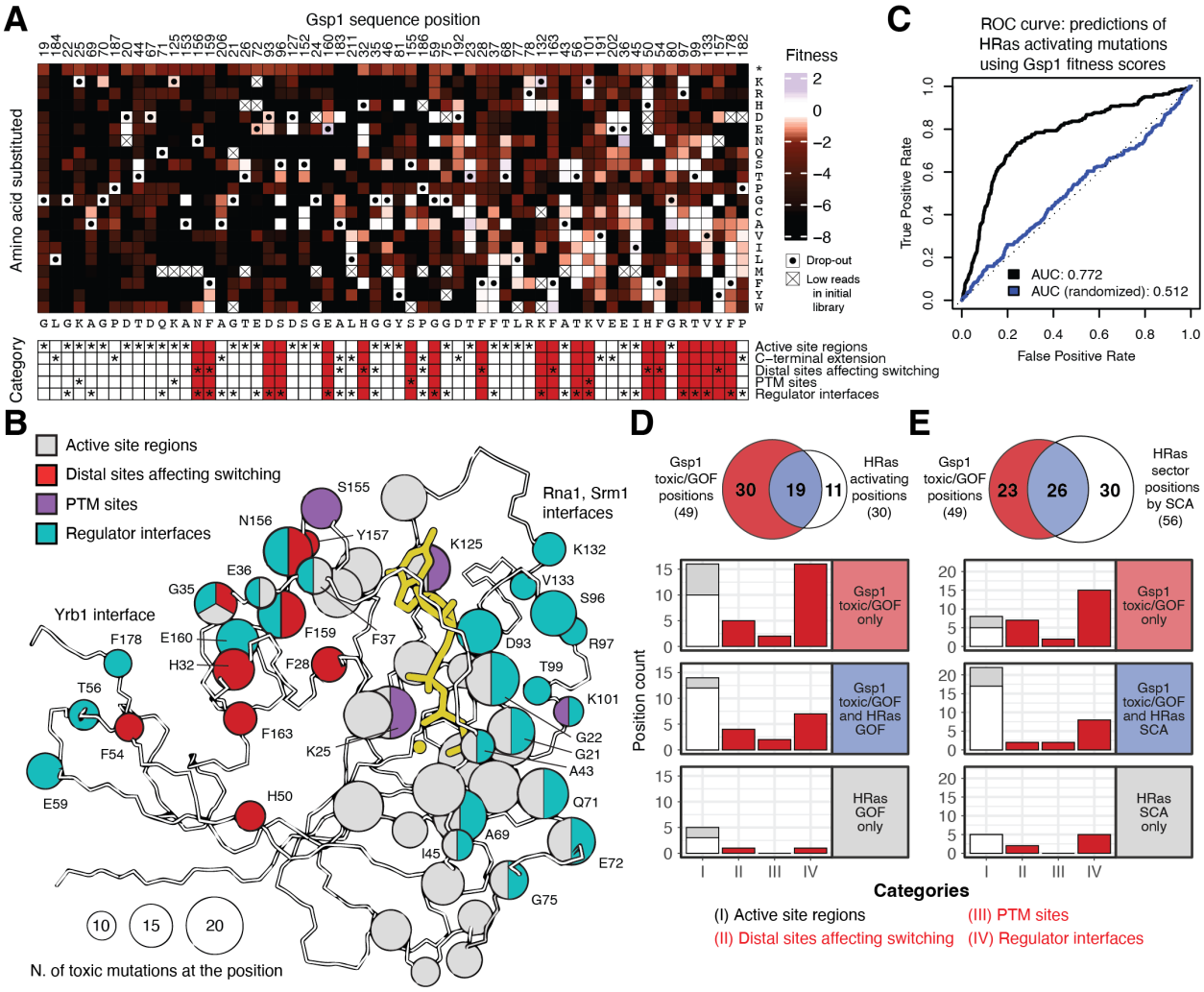

Fig. 4. Functional anatomy of the Gsp1 GTPase switch. (A) Heatmap showing fitness scores at toxic/GOF positions ordered by number of toxic/GOF mutations. WT amino acid residue shown below each column. Functional annotations (stars) are shown below and marked in red for positions outside of the active site. **(B)** Wire representation of Gsp1-GTP (PDB ID: 3M1I, residues 1-180). Toxic/GOF positions are shown in sphere representation. Sphere radius represents number of toxic/GOF mutations at each position. Spheres are colored by functional categories in (A). The nucleotide and Mg^{2+} cofactor are shown in yellow. **(C)** Receiver operating characteristic (ROC) curves and area under the curve (AUC) showing the statistical power of Gsp1 fitness scores in classifying an *H. sapiens* HRas mutant as activating, as defined by (*8*). Datasets were trimmed to the 156 sequence positions alignable for Gsp1 and HRas. **(D** and **E)** Overlap of functional sites defined as Gsp1 toxic/GOF and either (D) HRas activating or (E) comprising an HRas sector defined by statistical coupling analysis (SCA) (*23*) (**Table S1**). Bar graphs show functional annotations of positions as in (A) in each subset; a position can have multiple annotations. White bars indicate positions annotated as canonical active site regions across Ras-superfamily GTPases, while gray bars include residues in the expanded definitions for Ran/Gsp1 based on structural analysis (**fig. S2B** and Methods).