Wavelet coherence phases decode the universal switching mechanism of Ras GTPase superfamily

Zenia Motiwala[†$‡], Anand S. Sandholu[†$‡], Durba Sengupta[¶$], Kiran Kulkarni[†$*]

[†]Division of Biochemical Sciences, CSIR-National Chemical Laboratory, Pune-411008, India.

[$]Academy of Scientific and Innovative Research (AcSIR), CSIR-National Chemical Laboratory, Pune-411008, India.

[¶]Division of Physical and Material Chemistry, CSIR-National Chemical Laboratory, Pune-411008, India.

* **Corresponding author**: Dr. Kiran Kulkarni

E-mail: ka.kulkarni@ncl.res.in

ORCID: 0000-0002-7857-8932

‡ These authors contributed equally

Classification

BIOLOGICAL SCIENCES: Biophysics and Computational Biology

Keywords

Protein structural dynamics, Ras superfamily GTPases, Continuous wavelet transformation, wavelet coherence

Author contributions

ZM and ASS performed data analysis. DS assisted with MD simulation analysis. KK designed the project, supervised overall work and wrote the paper with inputs from ZM, ASS and DS.

Abstract

Ras superfamily GTPases are molecular switches which regulate critical cellular processes. Extensive structural and computational studies on this G protein family have tried to establish a general framework for their switching mechanism. The current understanding of the mechanism is that two loops, Switch I and Switch II, undergo conformational changes upon GTP binding and hydrolysis, which results in alteration of their functional state. However, because of variation in the extent of conformational changes seen across the members of the Ras superfamily, there is no generic *modus operandi* defining their switching mechanism, in terms of loop conformation. Here, we have developed a novel method employing wavelet transformation to dissect the structures of these molecular switches to explore indices that defines the unified principle of working. Our analysis shows that the structural coupling between the Switch I and Switch II regions is manifested in terms of wavelet coherence phases. The resultant phase pertaining to these regions serve as a functional identity of the GTPases. The coupling defined in terms of wavelet coherence phases is conserved across the Ras superfamily. In oncogenic mutants of the GTPases the phase coupling gets disentangled, this perhaps provides an alternative explanation for their aberrant function. Although similar observations were made using MD simulations, there was no structural parameter to define the coupling, as delineated here. Furthermore, the technique reported here is computationally inexpensive and can provide significant functional insights on the GTPases by analyzing as few as two structures.

Significance

Protein-nucleotide interaction induced structural changes are the reflections of functional states of small GTPases. However, from conventional comparison of protein structures these conformational changes cannot be parameterized to directly correlate them with functional states. The wavelet transformation based technique developed here fills this lacuna and provides a novel means to define the universal switching mechanism, with quantifiable indices. The technique gives an alternate prospective on the functional modulations of the GTPase due to oncogenic mutations. The methodology is generic in nature and can be extended to other class of proteins, like kinases.

Introduction

Proteins utilize reversible transformations, such as covalent modifications (example: phosphorylation) and ligand induced conformational changes, to exert signal transduction events. For instance, Ras superfamily of GTPases act as biomolecular switches to regulates diverse biological functions such as cell migration, proliferation and fate. They achieve this switching mechanism by shuttling between GDP bound 'OFF' and GTP bound 'ON' states (Fig. 1*A*). Comparative analysis of structures in different ligand bound states has proved to be a useful technique to gain functional insights. However, in few cases, simple structure comparison proved to be inadequate to derive generalized inferences, as the observed differences between structures are either not quantifiable or they exhibit complex divergence patterns. This again can be demonstrated from the Ras superfamily of GTPases, which comprises of 5 families: Rho, Ras, Rab, Ran and Arf GTPase (1–3). All the members of this superfamily possess the conserved nucleotide binding G domain, containing a 6-stranded mixed β-sheet surrounded by five α-helices. The G-domain harbors five conserved sequence motifs termed G1-G5, which stabilize the nucleotide-protein interactions. Out of these motifs, the P loop GxxxxGKT (G1) and the two switch regions (SWI & SWII), containing the xT/Sx (G2) and DxxG (G3), interact with phosphates of the nucleotide to function as sensors for the presence of the gamma-phosphate (1, 3) . Structures of GDP and GTP bound proteins have shown that, primarily SWI and SWII regions of the protein undergo conformational changes to deploy the conserved "loaded-spring" mechanism where release of the γ-phosphate after GTP hydrolysis allows the two switch regions to relax into the GDP specific conformation (*SI Appendix* Fig. S1*A*). These conformational changes are attributed to loss of two hydrogen bonds between γ-phosphate oxygens and the main chain NH groups of the invariant Thr and Gly residues (Thr35/ Gly60 in Ras) in SWI and SWII, respectively (*SI Appendix* Fig. S1*B*). Thus it was inferred that the functional states of Ras superfamily GTPases is the consequence of the conformational changes in their SWI and SWII regions (4, 5).

However, extensive structural and molecular dynamics (MD) simulation studies suggest that there are no unique conformational signatures of SWI & SWII regions to define the functional states of the GTPases (6, 7) . This poses an important question, whether the bound nucleotide alter the dynamics of the entire structure of the protein, making it suitable for interaction with the downstream effectors/regulators or simply offer conformational space

required to achieve the function (8) ? This question becomes even more pertinent while deciphering the impact of oncogenic mutants on the functioning of GTPases, as the crystal structures of mutants at positions 12, 13 or 61 (numbering as Cdc42), show no significant conformational deviation from the *wild* type structures (Fig. 2). Several MD simulation based studies, involving PCA, QM/MM and network analysis have addressed this question from different perspectives. By and large, these studies suggest that there are distinct ensembles of SWI & SWII conformers, which could be correlated with the functional states of the Ras superfamily GTPases and mutations in P-loop and switch regions affect the equilibrium of ensembles to alter the switching mechanism (7, 15). Similar attempts have been made by analyzing the sequences of GTPases, employing Statistical Coupling Analysis (SCA), which has shown that evolutionarily conserved residues act at the functionally important regions map in a structurally connected network (16–18) . All together, MD and SCA studies suggest an allosteric model where the coupling of different structural regions dictates the dynamics and hence the functioning of the GTPases and aberrations in the coupling due to mutations would result in their abnormal functioning. However, these techniques do not clearly bring out the unified *modus operandi* of GTPases, manifested in terms of the structural allostery. Furthermore, these methods require large amount of structural (in terms of number of MD trajectories) and /or sequence data to probe the structural coupling. In this paper, we have developed novel wavelet based tool to address questions like how the unified "loaded-spring" mechanism manifest across the Ras superfamily, despite exhibiting significant sequence and structural divergence? And how oncogenic mutations alter the loaded-spring mechanism?

Continuous wavelet transformation (CWT) has emerged as a powerful tool for feature extraction purposes, which can analyze localized and intermittent oscillations in the time series (19, 20). Coherence analysis of wavelet transform of two signals, known as wavelet coherence analysis (WCA), simultaneously provides information on the relative changes between the signals and occurrence of these changes in time domain with reasonable resolution. In the current studied we have applied WCA by mapping the 3D structure of proteins into 1D signal to unravel the co-related structural differences to gain functional insights.

Using this novel technique we have analyzed the conformational variations in the Ras superfamily GTPases upon change in their bound nucleotide. Our results show that indeed there is a unified switching mechanism in the Ras superfamily, which manifests as structural coupling

between SWI & SWII regions and more importantly this can be expressed as wavelet coherence phases. Thus, the regionally entangled propagation of conformational changes in the GTPases can be delineated as WC phase coupling, which can be correlated with their functional state.

Methodology

Wavelet transformation as tool to analyze conformational waves in proteins.

Conversion of 3D protein structures to 1D signal: Mathematically, a signal is nothing but changes in space or time, in some cases both. For example, an electrocardiogram (ECG) is a representation of a signal, which shows the electrical impulse passing through the heart as a function of time. Determination of fluctuations, such as discontinuities and breakdown points, in the signal will be of prime importance in delineating the information carried by the signal. Mathematical tools like Fourier or Wavelet transform (WT) aid in dissecting the meaningful information from the signals, which are otherwise not easily detectable. For instance WT has proved to be a powerful tool in accurate detection of QRS complex, T and P segments of ECG. These parameters determine how long the electrical wave takes to pass through the heart and pace of the wave to travel from one part of the heart to the other and thus provide information on the patho-physiology. The advantages of WT are that it may decompose a signal directly according to the frequency domains, which would be then represented in the time domain. Thus, both time and frequency information of the signal is retained. In simpler words, from WT one can know with reasonable resolution "when and how much" fluctuation occurred in the signal. Here, we have used WT techniques to determine both the quantum and the location of conformational changes in proteins, occurring due to change in their functional states. Proteins are the linear polymers of amino acids with distinct three-dimensional structures, which is defines of their function. Thus, three-dimensional structures of proteins can be looked as spatial changes of amino acids along the length of the ploy-peptide chain. For many class of proteins their functional states, like substrate/ligand bound-unbound & active-inactive, could again be viewed as motions introduced in their polypeptide chains. To apply WT techniques on proteins we first transformed the three dimensional (3D) structure of the protein to a one dimensional signal (1D) by plotting the Residue Contact Order (RCO) for each of the residue against their respective positions. Mathematically, RCO of the ith residue is defined as

$$RCO_i = \frac{1}{n}\sum_{j}^{L} |i-j|\delta_{ij}$$

where n is the number of contacts between residue *i* and the others (*j*) belonging the polypeptide chain of length *L*, which lie within 3.5 Å of the ith residue. Previous reports have defined RCO by excluding the contacts between the immediate neighbors of the residues (21, 22) . However, to generate a continuum model we have included the immediate neighboring residues, as well, for the calculations. RCO captures the spatial separation between the residues, which may not have covalent interactions, hence, encodes long range interactions present in the protein structure (21, 22)

Wavelet transformation of RCO signal of the proteins: The wavelet transform decomposes a signal as a combination of a set of basis functions, obtained by means of dilation (τ) and translation (*S*) of a proto type wave ψ_0,called as mother wavelet. As a characteristic, it has sliding windows that expand or compress to capture low- and high-frequency signals, respectively (23) (Chernick, 2001). CWT of a signal *X(t)* is defined as

$$W_t(s,\tau) = \sum_{t=0}^{T-1} X(t)\psi_0^*\left(\frac{(t-\tau)\delta t}{S}\right)$$

Where ψ_0^* indicates complex conjugate of ψ_0. S represents the scale, by varying which the wavelet can be dilated or contracted. By translating along localized time position τ, one can calculate the wavelet coefficients $W_t(s,\tau)$, which describe the contribution of the scales s to the time series *X(t)* at different time positions τ (24) .

The mother wavelet, which forms the orthonormal basis for the transformation, could be of different forms. For the current study we have use Morlet mother wavelet, which is defined as

$$\psi_0 = \pi^{\frac{-1}{4}} e^{i\omega_0 t} e^{\frac{-t^2}{2}}$$

Wavelet coefficients are normalized to unit variance in order to allow direct comparisons of the wavelet coefficients. The local wavelet power, which describes how the special range of interaction of the residues in the structure varies along the residue positions, can be computed from wavelet coefficients as $P_{w(s,\tau)} = 2^s|W_t(s,\tau)|^2$. If the signal has finite length, as in the present case, wavelet transformation of it introduces zero padding effects. These are the artifacts

present in the transformation that get more pronounced at the ends of the signal. Hence the wavelet power spectrum at the edges, in the region known as cone of influence (COI), is omitted for the analysis (24). The COI is demarcated in Fig. 1*A, C, E & G* & Fig. 2 *A & C*.

Wavelet coherence analysis: Wavelet coherence analysis provides insights on the variations between two signals. As mentioned earlier, wavelets are useful in delineating the "when and how much" of the variations between signals with reasonable resolution. The strength of the covariation between two signals (here conformational changes between two structures of a given protein in different states), represented as, *x(t)* and *y(t)* signals, can be quantified from the Wavelet coherence (WC) analysis. WC can be obtained from their cross wavelet transform (w_{xy}), defined as

$$w_{xy}(s,\tau) = w_x(s,\tau)w_y^*(s,\tau)$$

Where w_x is the CWT of the first structure [*x(t)*] and w_y^* is the complex conjugate of CWT of the second structure [*y(t)*]. The power of the cross-spectrum is modulus of the CWT, $\left|w_{xy}(s,\tau)\right|$. The statistical significance of wavelet coherence is obtained by using Monte Carlo randomization techniques (25). The direction of the covariation between the signals can be obtained from the wavelet coherence phase (WCP), which can be calculated from the real, $R\left(w_{xy}(s,\tau)\right)$, and the imaginary, $I\left(w_{xy}(s,\tau)\right)$,parts of the cross wavelet transform as

$$\phi_{xy}(s,\tau) = tan^{-1}\left[\frac{I\left(w_{xy}(s,\tau)\right)}{R\left(w_{xy}(s,\tau)\right)}\right]$$

In the present situation, WCP provides propagation of the conformational variation for each of the residues as a function of their position in space. To quantify the net phase for a particular region a vector sum of all the phases, corresponding to the residues in that region, was obtained as

$$\vec{\phi}_{\mathrm{Axy}} = \vec{\phi}_{\mathrm{1xy}} + \vec{\phi}_{\mathrm{2xy}} + ...\vec{\phi}_{\mathrm{Nxy}}$$

where *1....N* are the residues from the region *A*.

The algorithm is represented as a flowchart in Fig. 1 and it is realized as a C program and a Python script. Both can be accessed on the GitHub page:

https://github.com/drkakulkarni/WaveProtein

Results and Discussion

Morlet Wavelet captures the backbone conformation of the protein. Details of the features that could be captured by the CWT depend on factors like the choice of the mother wavelet and scales chosen for the transformation. Here, we choose Morlet Wavelet (with ω_0=6) and default scales used in the python module of continuous wavelet transformation (https://pypi.org/project/pycwt/), to obtain the CWT of the protein 1D signal. Hence it is essential to validate the parameters chosen here are suitable for protein structural feature detection using CWT. To achieve this we calculated CWT of all known Rac and Cdc42 structures bound to GTP and GDP (*SI Appendix* Table S1) and sorted them using unsupervised *Kmeans* clustering (*SI Appendix Transparent methods)*. The outcome of this exercise would validate whether CWT used here are suitable enough to capture the conformational state of the GTPases. For this purpose, only main chain atoms of the proteins were used to calculate the RCOs and subsequent CWTs, as this would be sufficient to delineate the backbone conformation and at the same time mask the deviations in the structures due to mutations present in them. Our analysis show that the parameters used in calculating CWTs are optimal in capturing the conformational states of GTPases, as *Kmeans* clustering sorted the GTP and GDP bound structures with more than 80% efficiency (*SI Appendix* Table S2).

Cross wavelet phase relationships show unique structural coupling between SWI and SWII regions of Small GTPases. As discussed earlier, conformational changes in SWI & SWII are being attributed to the functional states of Ras superfamily GTPases. However, the quantum of conformational change in SWI and SWII regions, upon change in their bound nucleotides, vary across different sub-families. For example, amongst GTPases considered here, Rho, Ras, Rab and Arf, the Rho family exhibit insignificant changes in the conformation of SWI & SWII regions, whereas in the other three it is more dramatic (Fig2. *B, D, E & F*). Therefore, correlating the functional states of GTPases from their SWI & SWII conformations is really challenging. This conundrum further poses difficulties in understanding mechanistics of GTPase interactions with their regulators and effectors. Furthermore, the classical way of ascertaining their functional states (ON & OFF) in terms of loss or gain of hydrogen bonds between γ-phosphate oxygens of the GDP/GTP (*SI Appendix* Fig. S1 *B*) with the protein would be an insufficient parameter, as GTPases have been proposed to exhibit intermediate conformational states such as "ON like" state (7, 15). Hence to explore a novel parameter that would provide tangible correlation

between the structure and functional states of GTPases, we probed the structural differences between GTP and GDP bound structures of Ras superfamily of GTPases, using wavelet coherence (WC) technique. Fig. 2 shows the WC plots between the GTP and GDP bound *wild type* structures of a few members of the Ras superfamily. Since for Cdc42, KRas, Rab11 and Arf6 GTPases complete (without breaks in the chains) GDP and GTP bound structures were available, they were chosen for the analysis (*SI Appendix* Table S1).

Like the structure superpositions, even overlay of RCO plots by themselves do not show significant changes in the SWI, SWII and P-loop regions. However, since RCO is a 1D projection of 3D structure, these plots would contain the information on the intrinsic dynamical alterations, which occur due to change in functional state of the protein. Hence to unravel these latent differential features of the structures, contained in the RCO, we performed Wavelet coherence analysis of the 1D signals of the protein structures (RCO plots). Coherence analysis of two signals provides the relationship between the signals in terms of relative changes occurring in them. The advantage of using wavelet coherence analysis, instead of simple coherence (also known as direct or time domain coherence), is to simultaneously get the information on the quantum and location (residue position) of the relative changes with reasonable resolution.

Fig. 2 *A*, *C, E & G* show WC plots for GTP-GDP structures of Cdc42, KRas, Arf1 and Rab11, respectively. The y-axis of the plots corresponds to an arbitrary unit of period, which in the current context is a determinant of the range over which particular residue has 'influence' and the x-axis represents the residue numbers. In essence, the y-axis of the plot is a measure of 3D structural content of every residue mapped on 1D. The colour gradient of the plot is a measure of the strength of structural correlations between the GTP and GDP bound states, which is in the range of -1 (100% inversely correlated) to +1 (100% correlated). The arrows in the plot indicate the phase angles, measured in counter-clockwise direction. The arrows pointing right and left correspond to a zero angle (in phase) and 180°(out-of phase), respectively. Similarly, upward and downward pointing arrows correspond to part of the signal with 90° and 270° phase relationship, respectively.

Clearly, all WC plots show changes in P-loop, SWI and SWII regions and these changes appear to be propagating from P-loop to the other two regions (Fig. 2 *A*, *C, E & G*). The quantum and propagation of change in these regions, as indicated by the WC-power (color of the region), appears to be same, which hints at the structural coupling between these regions. Another

common feature in all of these plots is deviations at the C-terminus. Even though we have excluded about 5 residues at the N and C terminus of the structures, as they are floppy and disordered in many crystal structures, we see significantly correlated conformational differences at extreme regions. From these plots, absolute structural correlations of the individual regions in terms of WC power (color gradient) appears to be challenging, as they are heterogeneous across the GTPases considered here. However, given the definition of the signal as 1D map of protein structure, the physical interpretation of arrows in the present context is propagation of relative conformational change. In short, WC plots brings out the subtle structural couplings between the functional regions of the GTPases, which otherwise would require at least 100ns of MD simulations. Furthermore, the structural signals, i.e, RCO plots (Fig. 2), do not explicitly shown any structural coupling, other than minor changes in the peak heights at P-loop, SWI & SWII regions. Even the autocorrelation analysis of RCOs, with a window of 5-10 residues, fails to conclusively show any correlation between the regions.

Next, we asked whether WC analysis delineates the convergent switching mechanism, as there are no conserved conformational signatures between different GTPases families. To address this, we looked at the collective motion of residues in P-loop, SWI and SWII regions to explore the signatures of conserved structural couplings that results in an orchestrated motion of these functional regions upon change in the bound nucleotide. As discussed earlier, the phase vectors of wavelet cross correlation indicate the direction of propagation of conformational changes between GDP and GTP bound structures, we analyzed the collective propagation of the structural changes in P-loop, SWI & SWII regions in terms of phase vectors. For this we calculated the resultant of the phase vectors belonging to these regions. The magnitude of this resultant vector is a measure of net conformational change and the direction indicates the cumulative correlated change. The magnitude of these resultant vectors vary across the GTPases, as it is a function of sequence and structural content of P-loop, SWI & SWII regions, which is quite divergent. However, the direction of these vectors could serve as a parameter to identify the functional state. In other words, the net conformational propagation, in terms of angle between the vector corresponding to a particular region and the y-axis would be zero if there is absolutely no coordinated conformational change between the GDP and GTP bound states and a finite value of this angle serves as an index to compare the conformational changes across the GTPases.

Interestingly, these vector plots (Fig. 3) show that there is a well-defined phase relationship between SWI-SWII regions (Table 1, yellow highlights), irrespective of the conformational changes seen in the crystal structures (Fig. 3 *A-D*). The phase relationship between SWI-SWII seems to be conserved in Cdc42, KRas & Rab11, however significantly deviates in Arf6 (Table 1, grey highlight). Perhaps, this could be due to the role of additional regions in Arf switching mechanism, where the rearrangement of the switch regions is coupled to a variable N-terminal extension, which is autoinhibitory in the GDP-bound form and swings out to facilitate activation (26). To test whether the observed phase vector relationship is a consequence of the switching mechanism or it is an artifact of the methodology, we randomly morphed the SWI region in $Cdc42^{GDP}$ structure and plotted the WC phase vector plot between the wild type GTP bound and the GDP bound morphed structures (Fig. 3 *I & J)*. We clearly see that, in this instance, the conserved phase vector relationship is completely abolished, as the conformational changes introduced are random (Table 1, blue highlight). Thus, the phase vector plots unambiguously bring out the conserved switching mechanism, which is not apparent in otherwise conformationally divergent Ras superfamily GTPases.

Oncogenic mutations in the P-loop disrupt the structural coupling between the functional regions of the GTPases. Although it is well known that oncogenic point mutations at 12th, 13th (P-loop) and 61st (SWII, residue numbering is with respect to the KRas structure) of the GTPases (27) abrogate their GTP hydrolyzing capacity, the exact role of the mutants in disrupting the GTP hydrolysis is not clear. Several studies, employing X-ray crystallography (28, 29) and NMR (7, 15, 30) have shown that there is no significant structural differences between the *wild* type and mutant proteins, including their P-loop and switch regions (RMSD between the mutant and *wild* type structures is in the range 0.5 to 0.7Å). In the light of GTPase-GAP complex structures (31), using MD simulations and QM/MM calculations, it was proposed that disruption of GTP hydrolysis in the P-loop mutants (especially G12V/C mutants) could be due to increased flexibility at the nucleotide binding pocket, which would alter the GTP/GDP release rate (27, 32, 33). Thus, in the absences of major structural rearrangements (Fig. 3 *F* & *G*), just from structural analysis, it is difficult to explain how subtle changes at nucleotide binding pocket alter the intrinsic hydrolyzing capacity of the GTPases. The problem in deriving a generalized mechanism gets even more compounded due to structure and sequence divergences amongst the different family of GTPases at the nucleotide binding pocket. Since wavelet transformations are known to

effectively separate individual signal components, we probed this problem with our WC tool. Comparison of WC analysis of *wild* type and G12V mutant structures showed that the conserved phase coupling between SWI- SWII regions is disrupted in the mutant structure (Table 1, pink highlights). Although the G12V mutation lies in the P-loop, major de-couplings are seen between SWI-SWII regions. SWII region harbors the catalytically important Q61 residue and disentanglement in the P-loop--SWII coupling would essentially alter the Q61-solvent configuration, which could in turn affect the GTP hydrolysis. Thus WCA shows that the long-range interaction between functional regions of the GTPase are important for its biological function and minor aberrations in them would results in decoupling of the long range interactions leading to abrogation their function.

Molecular dynamic (MD) simulations validate the WC phase coupling analysis. MD simulations have proved to be reliable mathematical tools to explore long-range interactions in proteins. However, they are computationally expensive. In this paper, we have proposed that WC analysis can throw light on the long-range interactions in proteins with little computational resource and fewer structures. In our study, the conformation space sampled is limited (two structures for each protein), which could lead to a possible bias in the results presented. Therefore, to explore weather our analysis hold true even when the tool is applied to a larger conformational space, we performed a 100ns simulation for both *wild* type and mutant structures of KRas and Cdc42. The trajectories were sampled at every 200 ps and for every trajectory phase vectors were calculated using WCA tool (*SI Appendix* Fig. S3). This exercise provided a larger sample size with 500 conformers to analyze the distribution of the wavelet phase angles. The results of this operation (Fig 5.) clearly showed that even for a larger conformation space, the phase angles corresponding to SWI and SWII regions follow the trend observed with just two structures. To further enhance the confidence level of our findings, for one of structures (KRas), we performed a very long 1μs simulation and sampled the trajectories at every 1000 ps. Interestingly, even for this larger dataset, the phase angles corresponding to SWI and SWII showed the same correlation, as determined with only two conformers. The personR-value of the joint (bivariate) distribution of SWI and SWII phases for the wild type proteins lie in the range of 0.2 to 0.68 (with p-values in the range 0.0005 to 0.0018), indicating decent correlation between the distributions. However, for mutants the joint distribution appears to be broadened with relatively lower personR correlation values. It is worth noting that the perasonR values must be

taken with caution, as the exact nature of the correlation (linear or non-linear) between the distributions is not clear. Thus, from these results it is evident that wavelet analysis has a potential to unravel intrinsic structural coupling by utilizing only two structures of the protein corresponding to different functional state.

Conclusions

It is evident that function of proteins manifest in their structure and its dynamics, which is encoded in their sequence. However, subtle changes in the structure, either due to their functional states or mutations, cannot be easily comprehended from crystal structures. Perhaps, computational tools like MD simulations may turn out to be good in probing protein dynamics. However, to derive generic inferences for a family of proteins would require extensive simulation studies, which are computationally intractable and also suffer from convergence problem due to the non-linear nature of the method.

The wavelet based methodology developed here is powerful enough to bring out the subtle dynamical aspects of proteins and it is computationally inexpensive. By applying WCA, we have shown that Ras superfamily of small GTPases possess unique coupling in their functional regions despite exhibiting huge structural divergence in these regions. WCA provides a non-conventional interpretation of this coupling, which aids in unraveling the functional states of the protein and their aberrations due to mutations. Thus, this study could be a stepping stone in delineating more complex dynamical aspects of protein structures and can be extended to other class of proteins, such as kinases, which exhibit conformational changes upon ligand binding.

Acknowledgments

KK would like to acknowledge grant from Department of Biotechnology, India (BT/PR12502/BRB/10/1387/2015) and ZM would like to acknowledge fellowship from University Grants Commission, India. ASS would like to than CSIR for fellowship (HCP0008)

Declaration of Interests

The authors declare no competing interests.

References

1. L. Goitre, E. Trapani, L. Trabalzini, S. F. Retta, The ras superfamily of small GTPases: The unlocked secrets. *Methods Mol. Biol.* **1120**, 1–18 (2014).

2. A. Valencia, P. Chardin, A. Wittinghofer, C. Sander, The ras protein family: evolutionary tree and role of conserved amino acids. *Biochemistry* **30** (1991).

3. K. Wennerberg, K. L. Rossman, C. J. Der, The Ras superfamily at a glance. *J. Cell Sci.* **118**, 843–846 (2005).

4. O. Pylypenko, H. Hammich, I. M. Yu, A. Houdusse, Rab GTPases and their interacting protein partners: Structural insights into Rab functional diversity. *Small GTPases* **9**, 22–48 (2018).

5. I. R. Vetter, a Wittinghofer, The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299–304 (2001).

6. A. Kumawat, S. Chakrabarty, K. Kulkarni, Nucleotide Dependent Switching in Rho GTPase: Conformational Heterogeneity and Competing Molecular Interactions. *Sci. Rep.* (2017) https:/doi.org/10.1038/srep45829.

7. G. Pálfy, I. Vida, A. Perczel, 1H, 15N backbone assignment and comparative analysis of the wild type and G12C, G12D, G12V mutants of K-Ras bound to GDP at physiological pH. *Biomol. NMR Assign.* (2019) https:/doi.org/10.1007/s12104-019-09909-7.

8. F. Raimondi, A. Felline, G. Portella, M. Orozco, F. Fanelli, Light on the structural communication in Ras GTPases. *J. Biomol. Struct. Dyn.* (2013) https:/doi.org/10.1080/07391102.2012.698379.

9. A. A. Gorfe, B. J. Grant, J. A. McCammon, Mapping the Nucleotide and Isoform-Dependent Structural and Dynamical Features of Ras Proteins. *Structure* (2008) https:/doi.org/10.1016/j.str.2008.03.009.

10. A. Kapoor, A. Travesset, Mechanism of the exchange reaction in HRAS from multiscale modeling. *PLoS One* (2014) https:/doi.org/10.1371/journal.pone.0108846.

11. H. Li, X. Q. Yao, B. J. Grant, Comparative structural dynamic analysis of GTPases. *PLoS Comput. Biol.* (2018) https:/doi.org/10.1371/journal.pcbi.1006364.

12. S. Muraoka, *et al.*, Crystal structures of the state 1 conformations of the GTP-bound H-Ras protein and its oncogenic G12V and Q61L mutants. *FEBS Lett.* (2012) https:/doi.org/10.1016/j.febslet.2012.04.058.

13. P. Prakash, A. A. Gorfe, Lessons from computer simulations of Ras proteins in solution and in membrane. *Biochim. Biophys. Acta - Gen. Subj.* (2013) https:/doi.org/10.1016/j.bbagen.2013.07.024.

14. A. Shurki, A. Warshel, Why Does the Ras Switch "Break" by Oncogenic Mutations? *Proteins Struct. Funct. Genet.* (2004) https:/doi.org/10.1002/prot.20004.

15. R. Chandrashekar, O. Salem, H. Krizova, R. McFeeters, P. D. Adams, A switch I mutant of Cdc42 exhibits less conformational freedom. *Biochemistry* (2011) https:/doi.org/10.1021/bi2004284.

16. P. Bandaru, *et al.*, Deconstruction of the Ras switching cycle through saturation mutagenesis. *Elife* **6** (2017).

17. O. Rivoire, K. A. Reynolds, R. Ranganathan, Evolution-Based Functional Decomposition of Proteins. *PLoS Comput. Biol.* **12**, e1004817 (2016).

18. T. Teşileanu, L. J. Colwell, S. Leibler, Protein Sectors: Statistical Coupling Analysis versus Conservation. *PLOS Comput. Biol.* **11**, e1004091 (2015).

19. H. K. Lee, Y. S. Choi, Application of continuous wavelet transform and convolutional neural network in decoding motor imagery brain-computer Interface. *Entropy* (2019) https:/doi.org/10.3390/e21121199.

20. T. Sato, R. Kajiwara, I. Takashima, T. Iijima, "Wavelet Correlation Analysis for Quantifying Similarities and Real-Time Estimates of Information Encoded or Decoded in Single-Trial Oscillatory Brain Waves" in *Wavelet Theory and Its Applications*, (InTech, 2018) https:/doi.org/10.5772/intechopen.74810.

21. J. Chen, N. S. Chaudhari, Statistical Analysis of Long-Range Interactions in Proteins in *Proceedings of the 2006 International Conference on Bioinformatics & Computational Biology, BIOCOMP'06, Las Vegas, Nevada, USA*, (2006), pp. 296–302.

22. D. Kihara, The effect of long-range interactions on the secondary structure formation of proteins. *Protein Sci.* **14**, 1955–1963 (2005).

23. M. R. Chernick, Wavelet Methods for Time Series Analysis. *Technometrics* (2001) https:/doi.org/10.1198/tech.2001.s49.

24. B. Cazelles, *et al.*, Wavelet analysis of ecological time series. *Oecologia* (2008) https:/doi.org/10.1007/s00442-008-0993-2.

25. A. Grinsted, J. C. Moore, S. Jevrejeva, Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Process. Geophys.* (2004) https:/doi.org/10.5194/npg-11-561-2004.

26. E. Sztul, *et al.*, Arf GTPases and their GEFs and GAPS: Concepts and challenges. *Mol. Biol. Cell* (2019) https:/doi.org/10.1091/mbc.E18-12-0820.

27. T. Pantsar, The current understanding of KRAS protein structure and dynamics. *Comput. Struct. Biotechnol. J.* (2020) https:/doi.org/10.1016/j.csbj.2019.12.004.

28. M. G. Rudolph, A. Wittinghofer, I. R. Vetter, Nucleotide binding to the G12V-mutant of Cdc42 investigated by X-ray diffraction and fluorescence spectroscopy: Two different nucleotide states in one crystal. *Protein Sci.* (2008) https:/doi.org/10.1110/ps.8.4.778.

29. M. Kawazu, *et al.*, Transforming mutations of RAC guanosine triphosphatases in human cancers. *Proc. Natl. Acad. Sci. U. S. A.* (2013) https:/doi.org/10.1073/pnas.1216141110.

30. M. J. Smith, B. G. Neel, M. Ikura, NMR-based functional profiling of RASopathies and oncogenic RAS mutations. *Proc. Natl. Acad. Sci. U. S. A.* (2013) https:/doi.org/10.1073/pnas.1218173110.

31. S. Lu, H. Jang, R. Nussinov, J. Zhang, The Structural Basis of Oncogenic Mutations G12, G13 and Q61 in Small GTPase K-Ras4B. *Sci. Rep.* (2016) https:/doi.org/10.1038/srep21949.

32. M. G. Khrenova, V. A. Mironov, B. L. Grigorenko, A. V. Nemukhin, Modeling the role of G12V and G13V ras mutations in the ras-GAP-catalyzed hydrolysis reaction of guanosine triphosphate. *Biochemistry* (2014) https:/doi.org/10.1021/bi5011333.

33. A. Sayyed-Ahmad, P. Prakash, A. A. Gorfe, Distinct dynamics and interaction patterns in H- and K-Ras oncogenic P-loop mutants. *Proteins Struct. Funct. Bioinforma.* (2017) https:/doi.org/10.1002/prot.25317.

Figures and Tables

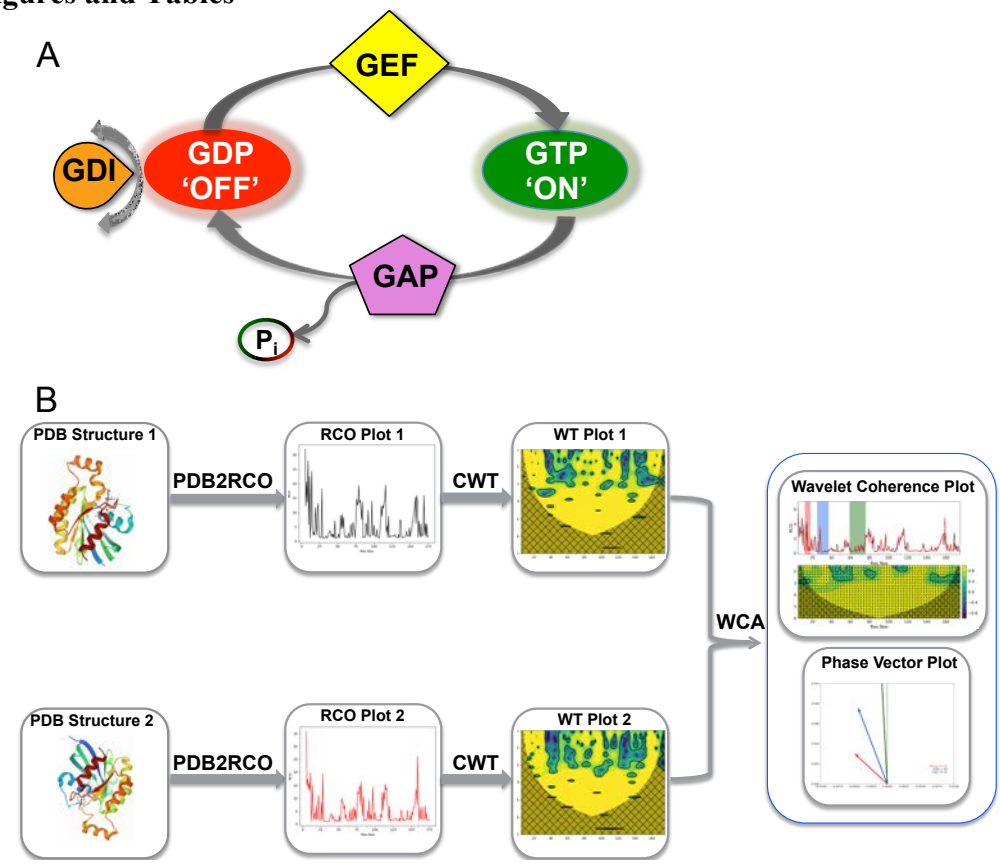

Fig. 1. (A) Switching mechanism of the Ras superfamily GTPase families. The switching between GTP bound ON and the GTP bound OFF state is regulated by Guanine Nucleotide Exchange Factors (GEFs) and GTPase Activating Proteins (GAPs); GEFs catalyze the exchanging their bound GDP with GTP and GAPs promote intrinsic GTP hydrolysis. For the Rho GTPase subfamily, has an additional level of regulation involving Guanine Dissociation Inhibitors (GDI), which sequester GDP bound Rho GTPases in the cytosol. (B) Flowchart of the wavelet based methodology developed for analyzing the GTPases.

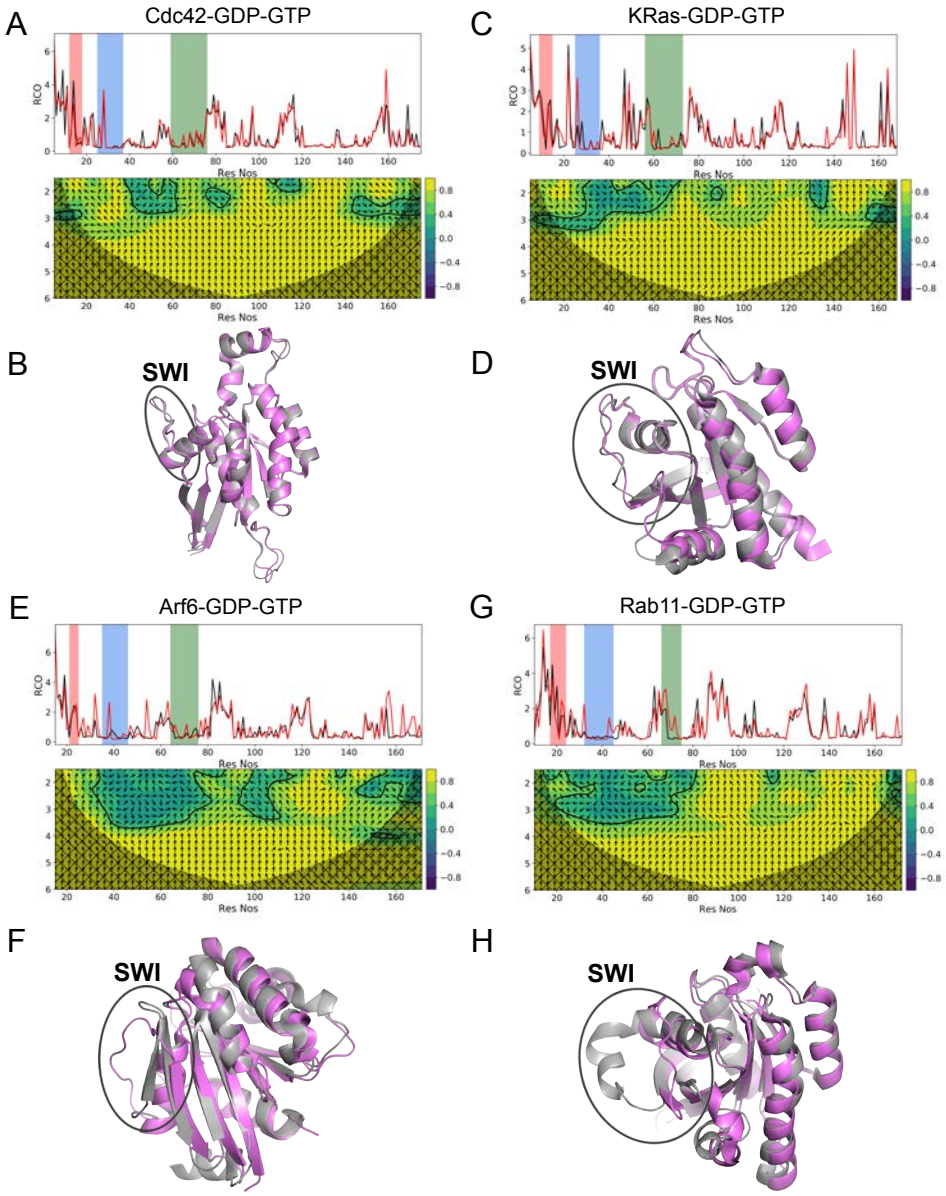

Fig. 2. Residue Contact Order (RCO) and Wavelet coherence (WC) plots of GDP and GTP bound structures and their respective wavelet coherence *(A)* RCO and WC plots of *wild type* Cdc42 GTP bound (PDB CODE: 2QRZ) and GDP bound (PDB CODE: 1AN0) structures and *(B)* their superimposition. *(C)* RCO and WC plots of *wild type* KRas GTP bound (PDB CODE: 6GOD) and GDP bound (PDB CODE: 4OBE) structures and *(D)* their superimposition. *(E)* RCO and WC plots of *wild type* Arf6 GTP bound (PDB CODE: 2J5X) and GDP bound (PDB CODE: 1E0S) structures and *(F)* their superimposition. *(G)* RCO and WC plots of *wild type* Rab11 GTP bound (PDB CODE: 3E5H) and GDP bound (PDB CODE: 2HXS) structures and *(H)* their superimposition. In all the structure superimposition images GTP bound structures are shown in violet and the GDP bound structures are shown in grey and SWI region is highlighted with a black circle.

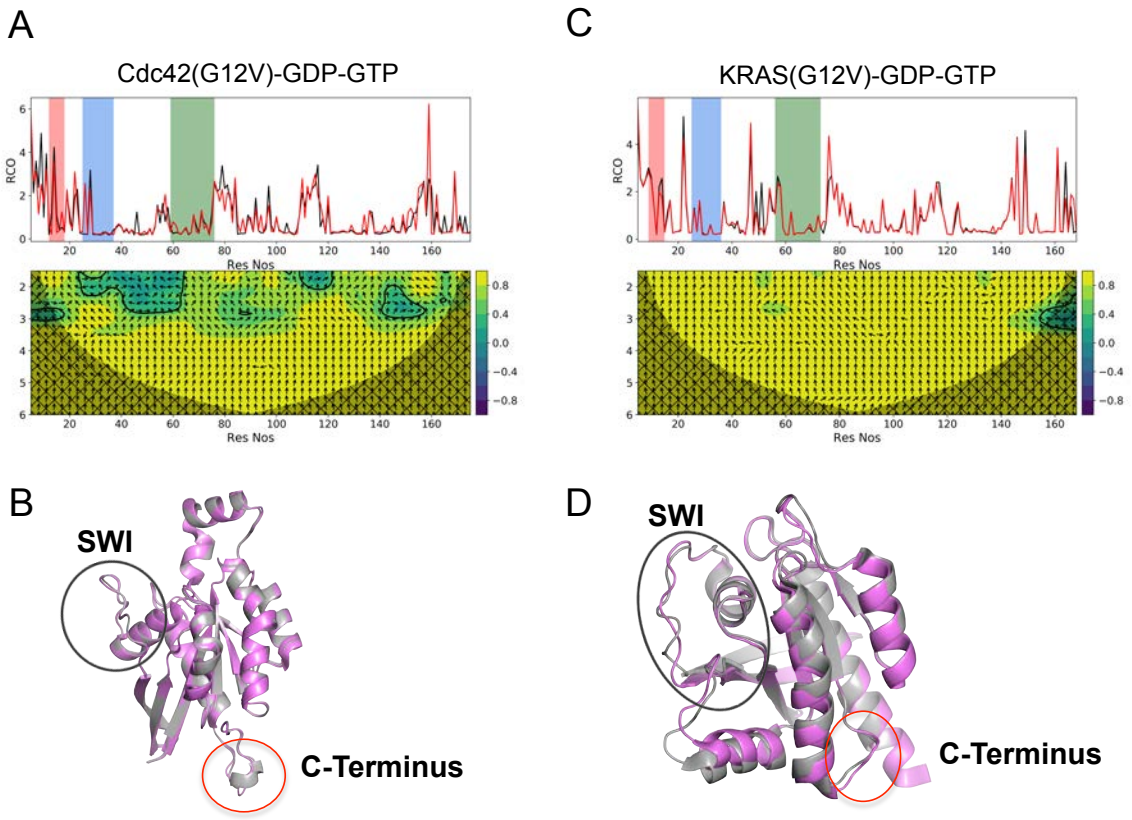

Fig. 3. Comparison of RCO and WC of KRas and Cdc42 G12V mutants with respect to their GTP bound *wild type* structures.
(A) RCO and WC plots of *wild type* Cdc42 GTP bound (PDB CODE: 2QRZ) and GDP bound mutant G12V (PDB CODE: 1A4R) structures and *(B)* their superimposition. *(C)* RCO and WC plots of *wild type* KRas GTP bound (PDB CODE: 6GOD) and GDP bound mutant G12V (PDB CODE: 4TQ9) structures and *(D)* their superimposition. The SWI region and C-terminus region showing significant conformational deviation are highlighted in black and red circles, respectively.

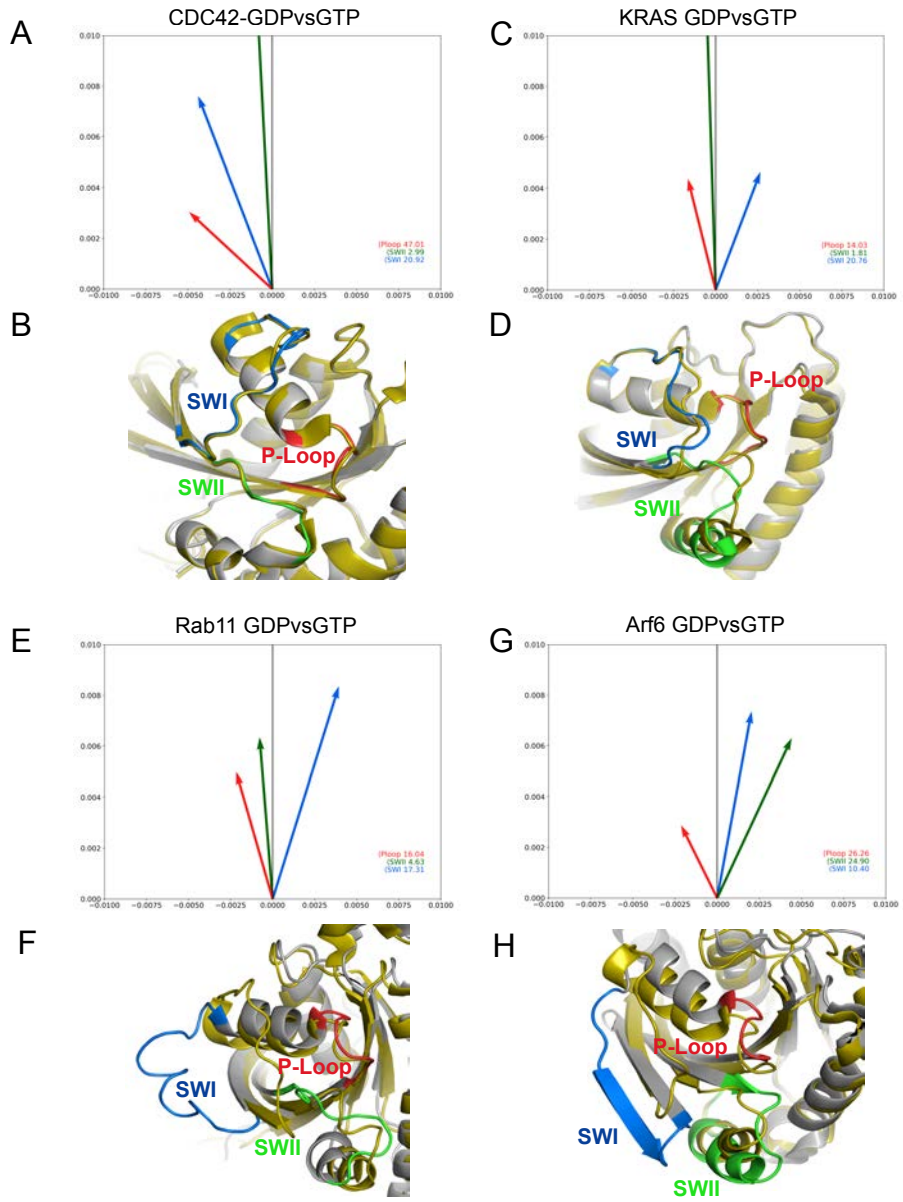

Fig. 4. Wavelet coherence phase (WCP) vector plots of Ras superfamily members
(A) WCP Vector plot for *wild type* Cdc42 GTP& GDP bound structures and *(B)* their structure superimposition. *(C)* WCP Vector plot for *wild type* KRas GTP& GDP bound structures and *(D)* their structure superimposition. *(E)* WCP Vector plot for *wild type* Rab11 GTP& GDP bound structures and *(F)* their structure superimposition. *(G)* WCP Vector plot for *wild type* Arf6 GTP& GDP bound structures and *(H)* their structure superimposition. The vector corresponding to the P-loop, SWI and SWII regions are shown in red, blue and green, respectively. The same color scheme is employed to highlight these regions in the structure superimposition. In the structure superimposition figures GTP bound structures are shown in olive and the GDP bound structure is shown in grey. The PDB codes of the structures shown are as mentioned in Fig. 2 & 3.

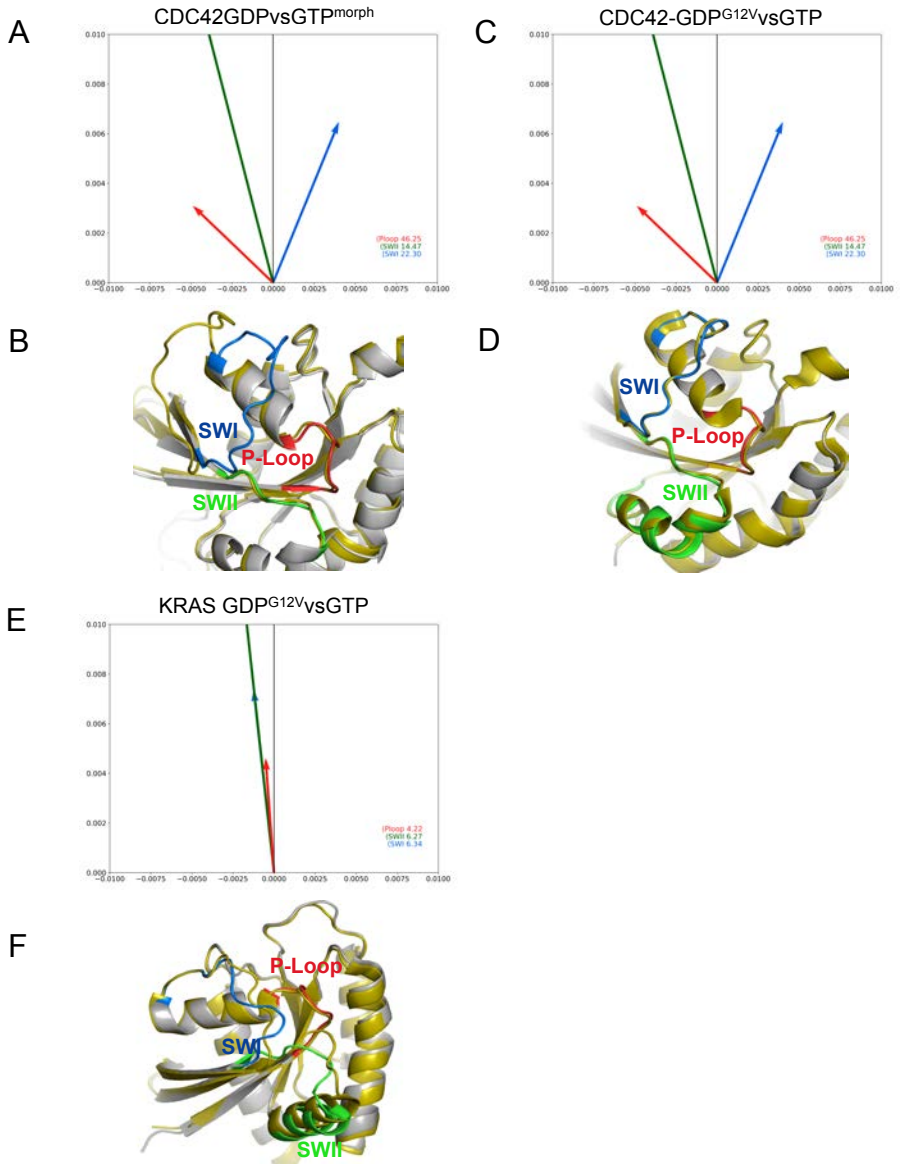

Fig. 5. Wavelet coherence phase (WCP) vector plots of GTPases with aberrations.
(A) WCP Vector plot for *wild type* Cdc42 morphed GTP& GDP bound structures and *(B)* their structure superimposition. *(C)* WCP Vector plot for *wild type* Cdc42 GTP & G12V mutant GDP bound structures and *(D)* their structure superimposition. *(E)* WCP Vector plot for *wild type* KRas GTP & G12V mutant GDP bound structures and *(F)* their structure superimposition. The vector corresponding to the P-loop, SWI and SWII regions are shown in red, blue and green, respectively. The same color scheme is employed to highlight these regions in the structure superimposition. In the structure superimposition figures GTP bound structures are shown in olive and the GDP bound structure is shown in grey. The PDB codes of the structures shown are as mentioned in Fig. 2 & 3.

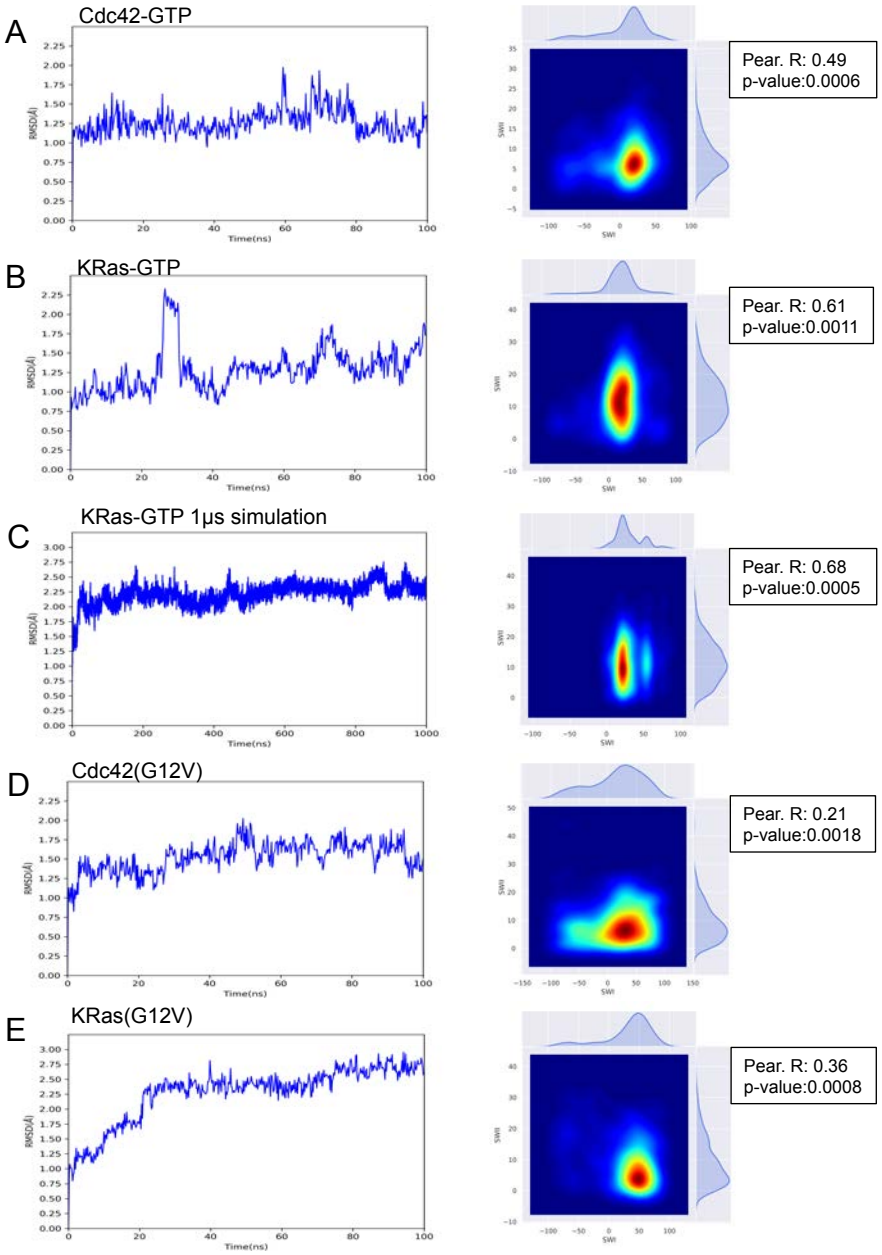

Fig. 6. RMSD of MD trajectories of Cdc42 & KRas and bivariate distribution of SWI & SWII Wavelet coherence phase (WCP) angles. *(A)* (left) RMSD of MD trajectories of GTP bound Cdc42 and (right) bivariate distribution SWI & SWII WC phase angles. *(B)* (left) RMSD of MD trajectories of GTP bound KRas and (right) bivariate distribution SWI & SWII WC. *(C)* (left) RMSD of 1μs MD trajectories of GTP bound KRas structure and (right) bivariate distribution SWI & SWII WC phase angles. *(D)* (left) RMSD of MD trajectories of GTP bound Cdc42 and (right) bivariate distribution SWI & SWII WC phase angles. *(E)* (left) RMSD of MD trajectories of GTP bound KRas and (right) bivariate distribution SWI & SWII WC phase angles. The PDB codes of the structures shown are as mentioned in Fig. 2 & 3. Perason R and p-value for each of the bivariate distribution are mentioned in the box.

Table 1. Conformational deviations between the two different states of GTPases in terms of angles made by the resultant phase vectors corresponding to P-loop, SWI and SWII regions with y-axis

	P-Loop	SWI	SWII
*wt*Cdc42 (GTP)- *wt*Cdc42(GDP)	47.01	20.92	2.99
morphed Cdc42 (GTP)- *wt*Cdc42(GDP)	46.25	22.3	14.47
*wt*Cdc42 (GTP)- G12V mutant Cdc42(GDP)	38.49	33.92	5.5
*wt*KRas (GTP)- *wt*KRas(GDP)	14.03	20.76	1.81
*wt*KRas(GTP)- G12V mutant KRas(GDP)	4.22	6.34	6.27
*wt*Rab11 (GTP)- *wt*Rab11(GDP)	16.04	17.31	4.63
*wt*Arf6(GTP)- *wt*Arf6(GDP)	26.26	10.4	24.9

Supplemental Information

Wavelet coherence phases decode the universal switching mechanism of Ras GTPase superfamily

Zenia Motiwala[†$‡], Anand S. Sandholu[†$‡], Durba Sengupta[¶$], Kiran Kulkarni[†$*]

[†]Division of Biochemical Sciences, CSIR-National Chemical Laboratory, Pune-411008, India.

[$]Academy of Scientific and Innovative Research (AcSIR), CSIR-National Chemical Laboratory, Pune-411008, India.

[¶]Division of Physical and Material Chemistry, CSIR-National Chemical Laboratory, Pune-411008, India.

* **Corresponding author**: Dr. Kiran Kulkarni

E-mail: ka.kulkarni@ncl.res.in

ORCID: 0000-0002-7857-8932

‡ These authors contributed equally

Transparent Methods

Unsupervised clustering of GTPase structures

Python implementation of KReas image clustering was employed to cluster the wavelet transforms of the GTPase structures (1). The COI of Wavelet transformation spectrums of the structures were masked to avoid influence of the artifacts in sorting.

MD simulation methods

Molecular dynamics (MD) simulations were carried out using Desmond v3.6 (2–4), implemented in Schrodinger-maestro 2020 (5). Simulation system was prepared by placing the GTPase in an orthogonal box with 10.0Å buffered distance in all the directions, after assigning bonds. The hydrogen atoms were added to the protein keeping pH 7.0 of the model. The system was solvated with TIP3P (6) solvent atoms and neutralized by the addition of Na+/Cl- ions. Subsequently, the system was subjected to energy minimization followed by 2000 steps of conjugate gradient algorithm with 120.0 kcal/mol/Å convergence threshold energy under NPT ensemble. Finally, simulation of the system was performed for 100 ns. Trajectory

recorded in each 200 ps. For KRas GTP, a longer simulation of 1 μs was performed and the trajectories were sampled at every 1ns. All the simulations, except the 1 μs KRas simulation, were performed in triplicates.

Table S1: List of PDBs used

Sl. No	Name of the GTPase	PDB Code	Remark	Organism	Reference
1	RHOA	1A2B	GTP Analogue bound	Homo sapiens	(17)
2	CDC 42 (G12V mutant)	1A4R	GDP bound	Homo sapiens	(24)
3	CDC 42	1AN0	GDP bound	Homo sapiens	To be published
4	RHOA (Complex with RHOGDI)	1CC0	GDP bound	Homo sapiens	(35)
5	CDC 42 (Complex with RHOGDI)	1DOA	GTP bound	Homo sapiens	(36)
6	RHOA	1DPF	GDP bound	Homo sapiens	(37)
7	RAC2 (Complex with RHOGDI)	1DS6	GDP bound	Homo sapiens	(38)
8	ARF6	1E0S	GDP bound	Homo sapiens	(39)
9	RAC1	1FOE	Nucleotide free	Homo sapiens	(7)
10	RHOA	1FTN	GDP bound	Homo sapiens	(8)
11	RAC1 (complex with Salmonella effector SptP)	1G4U	GDP bound	Homo sapiens	(9)
12	CDC 42 (complex with Cdc42GAP)	1GRN	GDP bound	Homo sapiens	(10)
13	CDC 42 (complex with Bacterial Toxin Sope)	1GZS	GTP bound	Homo sapiens	(11)
14	RAC1(complex with Pseudomonas Aeruginosa Exos Toxin)	1HE1	GDP bound	Homo sapiens	(12)
15	RAC1 (complex with RHOGDI)	1HH4	GDP bound	Homo sapiens	(13)
16	CDC 42(complex with Dbl exchange factors)	1KI1	Nucleotide free	Homo sapiens	(14)
17	RHOA (Q63L mutant)	1KMQ	GTP analogue	Homo sapiens	(15)

			bound		
18	CDC 42(complex with GEF Dbs)	1KZ7	Nucleotide free	Homo sapiens	(16)
19	CDC 42 (Complex with GEF dbs)	1KZG	Nucleotide free	Homo sapiens	(16)
20	RAC1	1MH1	GTP analogue bound	Homo sapiens	(18)
21	RHOA (complex with RHOGAP)	1OW3	GDP bound	Homo sapiens	(19)
22	RHOA (complex with the DH/PH fragment of PDZRhoGEF)	1XCG	Nucleotide free	Homo sapiens	(20)
23	CDC 42 (F28L mutant)	2ASE	Nucleotide free	Homo sapiens	(21)
24	CDC 42 (complex with GEF)	2DFK	Nucleotide free	Homo sapiens	(22)
25	RAC3	2G0N	GDP bound	Homo sapiens	To be published
26	RAC1 (complex with Protein kinase ypkA)	2H7V	GDP bound	Homo sapiens	(23)
27	Rab 28A	2HXS	GDP-3'P-Bound	Homo sapiens	To be published
28	RAC3	2IC5	GTP analogue bound	Homo sapiens	To be published
29	Arf6	2J5X	GTP bound	Homo sapiens	(25)
30	CDC 42 (T35A mutant)	2KB0	Nucleotide free	Homo sapiens	To be published
31	RAC1(complex with Trio)	2NZ8	Nucleotide free	Homo sapiens	(26)
32	CDC42	2QRZ	GTP bound	Homo sapiens	(27)
33	RAC1(complex with GEF-Vav1)	2VRW	Nucleotide free	Homo sapiens	(28)
34	RAC2 (G12V mutant)	2W2T	GDP bound	Homo sapiens	(29)
35	CDC 42 (complex with Dock9)	2WM9	Nucleotide free	Homo sapiens	(30)
36	CDC 42 (complex with Dock9)	2WMN	GDP bound	Homo sapiens	(30)
37	CDC 42 (complex with Dock9)	2WMO	GTP bound	Homo sapiens	(30)
38	RAC1 (complex with DOCK2)	2YIN	Nucleotide free	Homo sapiens	(31)
39	RAC1 (T17N mutant)	3B13	Nucleotide free	Homo sapiens	(32)
40	Rab 28	3E5H	GTP	Homo	(33)

			bound	sapiens	
41	RHOA (complex with GEF)	3KZ1	GTP bound	Homo sapiens	(34)
42	RHOA (complex with RhoGAP)	3MSX	GDP bound	Homo sapiens	To be published
43	RAC1 (P29S mutant)	3SBD	GTP bound	Homo Sapiens	(40)
44	RAC1(complex with Plexin-B1)	3SU8	GTP bound	Homo sapiens	(48)
45	RAC1	3TH5	GTP bound	Homo sapiens	(40)
46	RHOA	3TVD	GTP bound	Rattus norvegicus	(55)
47	CDC42 (T17N mutant)	3VHL	Nucleotide free	Homo sapiense	(56)
48	RHOA (Complex with RhoGEF)	4D0N	GDP bound	Homo sapiense	(57)
49	RHOA (complex with RhoGDI)	4F38	GTP bound	Mus musculus	(58)
50	RAC1 F28L mutant	4GZM	GTP bound	Homo sapiens	(59)
51	KRAS	4OBE	GDP bound	Homo sapiense	(60)
52	KRAS G12V mutant	4TQ9	GDP bound	Homo sapiens	(41)
53	RAC1 (complex with P-Rex1)	4YON	Nucleotide free	Homo sapiense	(42)
54	CDC42 (complex with RACGAP)	5C2J	GDP bound	Mus musculus	To be published
55	RHOA(complex with RACGAP)	5C2K	GDP bound	Homo sapiens	To be published
56	CDC 42(complex with GAP)	5CJP	GTP bound	Homo sapiens	(43)
57	RAC1 (complex with P-Rex1)	5FI0	Nucleotide free	Homo sapiense	(44)
58	CDC42 (complex with P-Rex1)	5FI1	Nucleotide free	Homo sapiense	(44)
59	RHOA (complex with RHOGDI)	5FR1	GDP bound	Homo sapiense	(45)
60	RHOA (complex with RhoGDI)	5FR2	GDP bound	Homo sapiense	(46)
61	RHOA (complex with RhoGAP)	5HPY	GDP bound	Homo sapiense	(47)
62	RHOA (complex with ARHGEF)	5JHG	Nucleotide free	Homo sapiense	To be published
63	RHOA (complex with ARHGEF)	5JHH	Inhibitor bound	Homo sapiens	To be published
64	RAC1	5N6O	GDP bound	Homo sapiense	(49)

65	RAC1	6AGP	GDP bound	Homo sapiens	(50)
66	CDC 42(complex with DOCK7)	6AJ4	Nucleotide free	Homo sapiense	(51)
67	RHOA (complex with RhoGEF)	6BC0	GTP analouge bound	Homo sapiense	(52)
68	RAC1 (complex with RhoGEF)	6BC1	GTP analouge bound	Homo sapiense	(52)
69	RHOA (complex with RhoGEF)	6BCA	GTP analouge bound	Homo sapiense	(53)
70	RHOA (complex with RhoGEF)	6BCB	GTP analouge bound	Homo sapiens	(53)
71	KRAS	6GOD	GTP bound	Homo sapiens	(54)

Table S2: Sorting of GTP and GDP bound structures by Kmeans clustering

	Prediction Efficiency (%)	**False Positive (%)**
GTP	85.7%	28.5%
GDP	83.3%	8.93%

References

1. F. Nelli, F. Nelli, "Deep Learning with TensorFlow" in *Python Data Analytics*, (2018) https:/doi.org/10.1007/978-1-4842-3913-1_9.

2. D. Shivakumar, *et al.*, Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the opls force field. *J. Chem. Theory Comput.* (2010) https:/doi.org/10.1021/ct900587b.

3. Z. Guo, *et al.*, Probing theα-helical structural stability of stapled p53 peptides: Molecular dynamics simulations and analysis: Research article. *Chem. Biol. Drug Des.* (2010) https:/doi.org/10.1111/j.1747-0285.2010.00951.x.

4. K. J. Bowers, *et al.*, Scalable algorithms for molecular dynamics simulations on commodity clusters in *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, SC'06*, (2006) https:/doi.org/10.1145/1188455.1188544.

5. Schrödinger Release, Desmond Molecular Dynamics System. *Schrödinger LLC* (2019).

6. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

7. D. K. Worthylake, K. L. Rossman, J. Sondek, Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1. *Nature* (2000) https:/doi.org/10.1038/35047014.

8. Y. Wei, *et al.*, Crystal structure of RhoA-GDP and its functional implications. *Nat. Struct. Biol.* (1997) https:/doi.org/10.1038/nsb0997-699.

9. C. E. Stebbins, J. E. Galán, Modulation of host signaling by a bacterial mimic: Structure of the Salmonella effector SptP bound to Rac1. *Mol. Cell* (2000) https:/doi.org/10.1016/S1097-2765(00)00141-6.

10. N. Nassar, G. R. Hoffman, D. Manor, J. C. Clardy, R. A. Cerione, Structures of Cdc42 bound to the active and catalytically compromised forms of Cdc42GAP. *Nat. Struct. Biol.* (1998) https:/doi.org/10.1038/4156.

11. G. Buchwald, *et al.*, Structural basis for the reversible activation of a Rho protein by the bacterial toxin SopE. *EMBO J.* (2002) https:/doi.org/10.1093/emboj/cdf329.

12. M. Würtele, *et al.*, How the Pseudomonas aeruginosa ExoS toxin downregulates Rac. *Nat. Struct. Biol.* (2001) https:/doi.org/10.1038/83007.

13. S. Grizot, *et al.*, Crystal structure of the Rac1 - RhoGDI complex involved in NADPH oxidase activation. *Biochemistry* (2001) https:/doi.org/10.1021/bi010288k.

14. J. T. Snyder, *et al.*, Structural basis for the selective activation of rho gtpases by dbl exchange factors. *Nat. Struct. Biol.* (2002) https:/doi.org/10.1038/nsb796.

15. K. Longenecker, *et al.*, Structure of a constitutively activated RhoA mutant (Q63L) at 1.55 Å resolution. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* (2003) https:/doi.org/10.1107/S0907444903005390.

16. K. L. Rossman, *et al.*, A crystallographic view of interactions between Dbs and Cdc42: PH domain-assisted guanine nucleotide exchange. *EMBO J.* (2002) https:/doi.org/10.1093/emboj/21.6.1315.

17. K. Ihara, *et al.*, Crystal structure of human RhoA in a dominantly active form complexed with a GTP analogue. *J. Biol. Chem.* (1998) https:/doi.org/10.1074/jbc.273.16.9656.

18. M. Hirshberg, R. W. Stockley, G. Dodson, M. R. Webb, The crystal structure of human rac1, a member of the rho-family complexed with a GTP analogue. *Nat. Struct. Biol.* (1997) https:/doi.org/10.1038/nsb0297-147.

19. D. L. Graham, *et al.*, MgF3- as a transition state analog of phosphoryl transfer. *Chem. Biol.* **9**, 375–381 (2002).

20. U. Derewenda, *et al.*, The crystal structure of RhoA in complex with the DH/PH fragment of PDZRhoGEF, an activator of the Ca2+ sensitization pathway in smooth muscle. *Structure* **12**, 1955–1965 (2004).

21. P. D. Adams, R. E. Oswald, Solution structure of an oncogenic mutant of Cdc42Hs. *Biochemistry* (2006) https:/doi.org/10.1021/bi051686g.

22. S. Xiang, *et al.*, The Crystal Structure of Cdc42 in Complex with Collybistin II, a Gephyrin-interacting Guanine Nucleotide Exchange Factor. *J. Mol. Biol.* (2006) https:/doi.org/10.1016/j.jmb.2006.03.019.

23. G. Prehna, M. I. Ivanov, J. B. Bliska, C. E. Stebbins, Yersinia Virulence Depends on Mimicry of Host Rho-Family Nucleotide Dissociation Inhibitors. *Cell* (2006) https:/doi.org/10.1016/j.cell.2006.06.056.

24. M. G. Rudolph, A. Wittinghofer, I. R. Vetter, Nucleotide binding to the G12V-mutant of Cdc42 investigated by X-ray diffraction and fluorescence spectroscopy: Two different nucleotide states in one crystal. *Protein Sci.* (2008) https:/doi.org/10.1110/ps.8.4.778.

25. S. Pasqualato, J. Ménétrey, M. Franco, J. Cherfils, The structural GDP/GTP cycle of human Arf6. *EMBO Rep.* (2001) https:/doi.org/10.1093/embo-reports/kve043.

26. M. K. Chhatriwala, L. Betts, D. K. Worthylake, J. Sondek, The DH and PH Domains of Trio Coordinately Engage Rho GTPases for their Efficient Activation. *J. Mol. Biol.* (2007) https:/doi.org/10.1016/j.jmb.2007.02.060.

27. M. J. Phillips, G. Calero, B. Chan, S. Ramachandran, R. A. Cerione, Effector proteins exert an important influence on the signaling-active state of the small GTPase Cdc42. *J. Biol. Chem.* (2008) https:/doi.org/10.1074/jbc.M706271200.

28. J. Rapley, V. L. J. Tybulewicz, K. Rittinger, Crucial structural role for the PH and C1 domains of the Vav1 exchange factor. *EMBO Rep.* (2008) https:/doi.org/10.1038/embor.2008.80.

29. T. D. Bunney, *et al.*, Structural Insights into Formation of an Active Signaling Complex between Rac and Phospholipase C Gamma 2. *Mol. Cell* (2009) https:/doi.org/10.1016/j.molcel.2009.02.023.

30. J. Yang, Z. Zhang, S. Mark Roe, C. J. Marshall, D. Barford, Activation of Rho GTPases by DOCK exchange factors is mediated by a nucleotide sensor. *Science (80-.).* (2009) https:/doi.org/10.1126/science.1174468.

31. K. Kulkarni, J. Yang, Z. Zhang, D. Barford, Multiple factors confer specific Cdc42 and Rac protein activation by dedicator of cytokinesis (DOCK) nucleotide exchange factors. *J. Biol. Chem.* (2011) https:/doi.org/10.1074/jbc.M111.236455.

32. K. Hanawa-suetsugu, M. Kukimoto-niino, C. Mishima-tsumagari, R. Akasaka, N. Ohsawa, Structural basis for mutual relief of the Rac guanine nucleotide exchange

factor DOCK2 and its partner ELMO1 from their autoinhibited forms. **109**, 3305–3310 (2012).

33. S. H. Lee, K. Baek, R. Dominguez, Large nucleotide-dependent conformational change in Rab28. *FEBS Lett.* (2008) https:/doi.org/10.1016/j.febslet.2008.11.008.

34. Z. Chen, *et al.*, Activated RhoA binds to the Pleckstrin Homology (PH) domain of PDZ-RhoGEF, a potential site for autoregulation. *J. Biol. Chem.* (2010) https:/doi.org/10.1074/jbc.M110.122549.

35. K. Longenecker, *et al.*, How RhoGDI binds Rho. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (1999) https:/doi.org/10.1107/S090744499900801X.

36. G. R. Hoffman, N. Nassar, R. A. Cerione, Structure of the Rho family GTP-binding protein Cdc42 in complex with the multifunctional regulator RhoGDI. *Cell* (2000) https:/doi.org/10.1016/S0092-8674(00)80670-4.

37. T. Shimizu, *et al.*, An Open Conformation of Switch I Revealed by the Crystal Structure of a Mg2+-free Form of RHOA Complexed with GDP IMPLICATIONS FOR THE GDP/GTP EXCHANGE MECHANISM. *J. Biol. Chem.* **275**, 18311–18317 (2000).

38. K. Scheffzek, I. Stephan, O. N. Jensen, D. Illenberger, P. Gierschik, The Rac-RhoGDI complex and the structural basis for the regulation of Rho proteins by RhoGDI. *Nat. Struct. Biol.* (2000) https:/doi.org/10.1038/72392.

39. J. Ménétrey, E. Macia, S. Pasqualato, M. Franco, J. Cherfils, Structure of Arf6-GDP suggests a basis for guanine nucleotide exchange factors specificity. *Nat. Struct. Biol.* (2000) https:/doi.org/10.1038/75863.

40. M. Krauthammer, *et al.*, Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. *Nat. Genet.* (2012) https:/doi.org/10.1038/ng.2359.

41. J. C. Hunter, *et al.*, Biochemical and structural analysis of common cancer-associated KRAS mutations. *Mol. Cancer Res.* (2015) https:/doi.org/10.1158/1541-7786.MCR-15-0203.

42. C. M. Lucato, *et al.*, The phosphatidylinositol (3,4,5)-trisphosphate-dependent Rac exchanger 1·Ras-related C3 botulinum toxin substrate 1 (P-Rex1·Rac1) complex reveals the basis of Rac1 activation in breast cancer cells. *J. Biol. Chem.* (2015) https:/doi.org/10.1074/jbc.M115.660456.

43. L. LeCour, *et al.*, The Structural Basis for Cdc42-Induced Dimerization of IQGAPs. *Structure* (2016) https:/doi.org/10.1016/j.str.2016.06.016.

44. J. N. Cash, E. M. Davis, J. J. G. Tesmer, Structural and Biochemical Characterization of the Catalytic Core of the Metastatic Factor P-Rex1 and Its Regulation by PtdIns(3,4,5)P3. *Structure* (2016) https:/doi.org/10.1016/j.str.2016.02.022.

45. N. Kuhlmann, S. Wroblowski, L. Scislowski, M. Lammers, RhoGDIα Acetylation at K127 and K141 Affects Binding toward Nonprenylated RhoA. *Biochemistry* (2016) https:/doi.org/10.1021/acs.biochem.5b01242.

46. N. Kuhlmann, *et al.*, Structural and mechanistic insights into the regulation of the fundamental Rho regulator RhoGDIα by lysine acetylation. *J. Biol. Chem.* (2016) https:/doi.org/10.1074/jbc.M115.707091.

47. F. Yi, *et al.*, Noncanonical Myo9b-RhoGAP Accelerates RhoA GTP Hydrolysis by a Dual-Arginine-Finger Mechanism. *J. Mol. Biol.* (2016) https:/doi.org/10.1016/j.jmb.2016.06.014.

48. C. H. Bell, A. R. Aricescu, E. Y. Jones, C. Siebold, A dual binding mode for RhoGTPases in plexin signalling. *PLoS Biol.* (2011) https:/doi.org/10.1371/journal.pbio.1001134.

49. Y. Ferrandez, *et al.*, Allosteric inhibition of the guanine nucleotide exchange factor DOCK5 by a small molecule. *Sci. Rep.* (2017) https:/doi.org/10.1038/s41598-017-13619-2.

50. Y. Toyama, K. Kontani, T. Katada, I. Shimada, Conformational landscape alternations promote oncogenic activities of Ras-related C3 botulinum toxin substrate 1 as revealed by NMR. *Sci. Adv.* (2019) https:/doi.org/10.1126/sciadv.aav8945.

51. M. Kukimoto-Niino, *et al.*, Structural Basis for the Dual Substrate Specificity of DOCK7 Guanine Nucleotide Exchange Factor. *Structure* (2019) https:/doi.org/10.1016/j.str.2019.02.001.

52. O. Dada, S. Gutowski, C. A. Brautigam, Z. Chen, P. C. Sternweis, Direct regulation of p190RhoGEF by activated Rho and Rac GTPases. *J. Struct. Biol.* (2018) https:/doi.org/10.1016/j.jsb.2017.11.014.

53. Z. Chen, S. Gutowski, P. C. Sternweis, Crystal structures of the PH domains from Lbc family of RhoGEFs bound to activated RhoA GTPase. *Data Br.* (2018) https:/doi.org/10.1016/j.dib.2018.01.024.

54. A. Cruz-Migoni, *et al.*, Structure-based development of new RAS-effector inhibitors from a combination of active and inactive RAS-binding compounds. *Proc. Natl. Acad. Sci. U. S. A.* (2019) https:/doi.org/10.1073/pnas.1811360116.

55. C. Jobichen, K. Pal, K. Swaminathan, Crystal structure of mouse RhoA:GTPcS complex in a centered lattice. *J. Struct. Funct. Genomics* (2012) https:/doi.org/10.1007/s10969-012-9143-5.

56. Y. Harada, *et al.*, DOCK8 is a Cdc42 activator critical for interstitial dendritic cell migration during immune responses. *Blood* (2012) https:/doi.org/10.1182/blood-2012-01-407098.

57. K. R. A. Azeez, S. Knapp, J. M. P. Fernandes, E. Klussmann, J. M. Elkins, The crystal structure of the RhoA-AKAP-Lbc DH-PH domain complex. *Biochem. J.* (2014) https:/doi.org/10.1042/BJ20140606.

58. Z. Tnimov, *et al.*, Quantitative analysis of prenylated RhoA interaction with its chaperone, RhoGDI. *J. Biol. Chem.* (2012) https:/doi.org/10.1074/jbc.M112.371294.

59. M. J. Davis, *et al.*, RAC1P29S is a spontaneously activating cancer-associated GTPase. *Proc. Natl. Acad. Sci. U. S. A.* (2013) https:/doi.org/10.1073/pnas.1220895110.

60. J. C. Hunter, *et al.*, In situ selectivity profiling and crystal structure of SML-8-73-1, an active site inhibitor of oncogenic K-Ras G12C. *Proc. Natl. Acad. Sci. U. S. A.* (2014) https:/doi.org/10.1073/pnas.1404639111.

Supplementary Figures

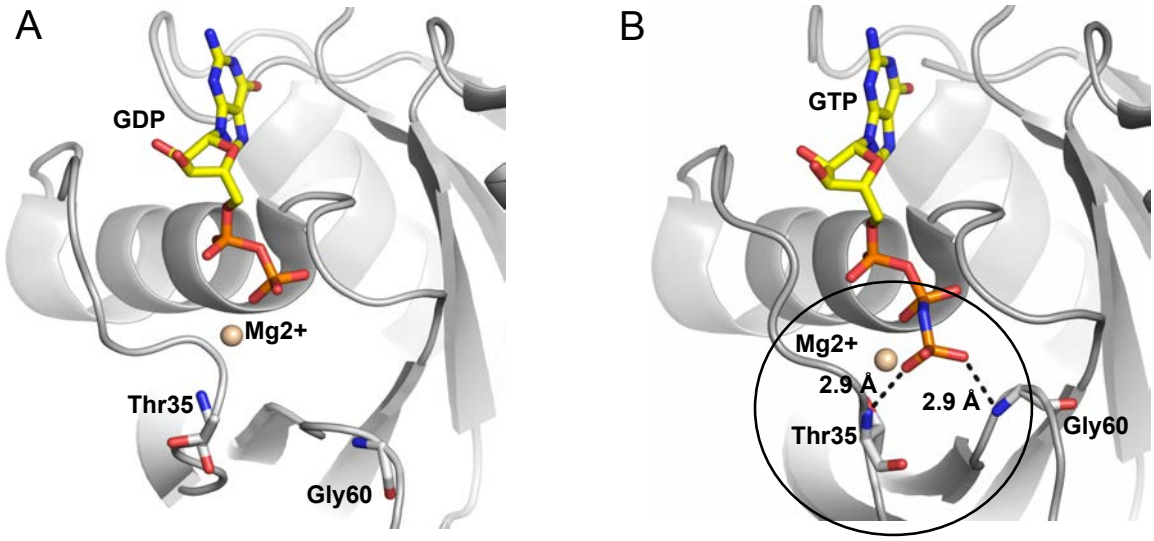

Fig. S1. *(A)* GDP bound K Ras (PDB CODE: 4OBE) *(B)* GTP bound KRas (PDB CODE: 6GOD). The hydrogen bonding between γ-phosphate & Thr35 and γ-phosphate & Gly60 are shown with black dots and highlighted in a circle. GTP and GDP are indicated with yellow sticks and Mg^{2+} is indicated as a wheat colour sphere.

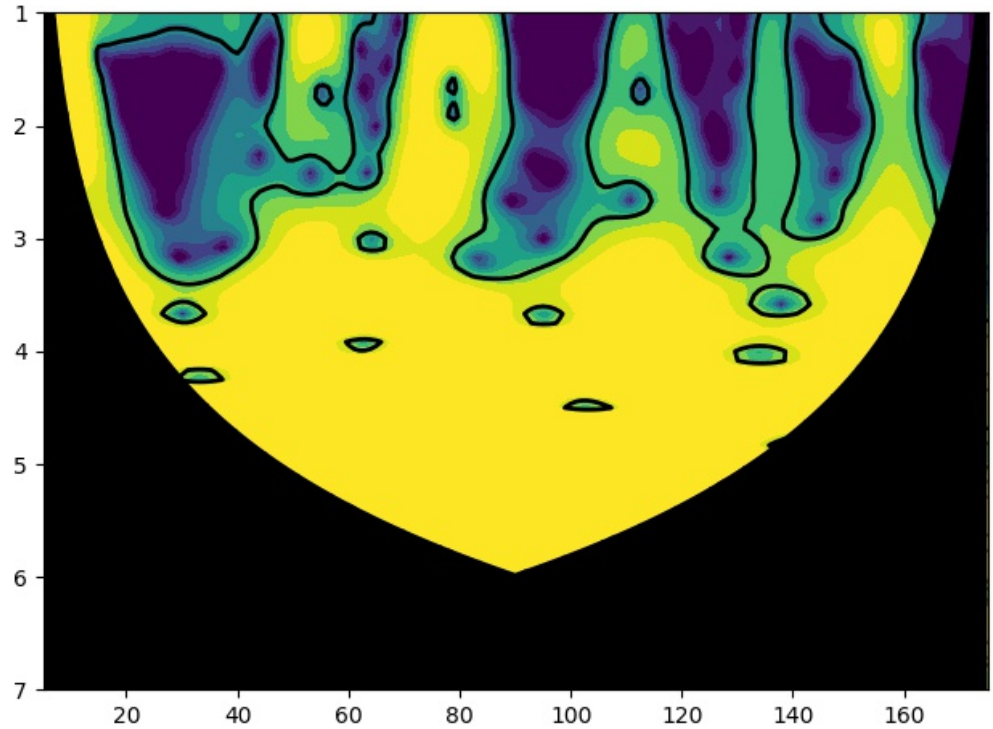

Fig. S2. CWT of CDC42 GDP bound (PDB CODE: 1AN0). The COI is masked in black.

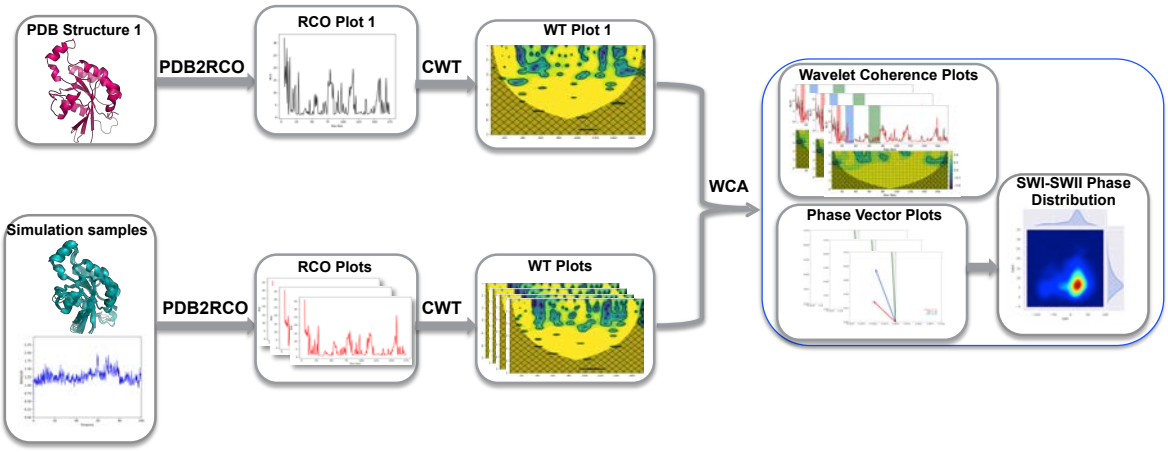

Fig. S3. Flowchart showing the algorithm for calculating the bivariate distribution of SWI & SWII Wavelet coherence phase (WCP) angles, employing MD simulation for GTP bound structure.