Intrinsic K-Ras dynamics:

A novel molecular dynamics data analysis method

shows causality between residue pairs

Sezen Vatansever[1,2,3], Zeynep H. Gümüş[2, 3*], Burak Erman[1*]

[1]Department of Chemical and Biological Engineering, College of Engineering, Koç University, Rumelifeneri Yolu, 34450, Sarıyer, Istanbul, Turkey

[2]Department of Genetics and Genomics, [3]Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY 10029

*Correspondence: berman@ku.edu.tr (B.E.) and zeynep.gumus@mssm.edu (Z.H.G.)

Summary

While mutant K-Ras is an important therapeutic target for human cancers, there are still no drugs that directly target it. Recent promising studies emphasize the significance of dynamics data to selectively target its active/inactive states. However, despite tremendous information on K-Ras, the direction of information flow in the allosteric regulation of its dynamics has not yet been elucidated. Here, we present a novel approach that identifies causality in correlated motions of proteins and apply it to K-Ras dynamics. Specifically, we analyze molecular dynamics simulations data and comprehensively investigate nucleotide-dependent intrinsic K-Ras activity. We show that GTP binding leads to characteristic residue correlations with relatively long decay times by stabilizing K-Ras motions. Furthermore, we identify for the first time driver-follower relationships of correlated motions in the regulation of K-Ras activity. Our results can be utilized for directly targeting mutant K-Ras in future studies.

Introduction

K-Ras is a small GTP-binding protein pivotal in cellular signaling. Somatic K-Ras mutations are among the most common activating cancer lesions, driving 71% of pancreas, 35% of colon and 17% of lung cancers (Slebos et al., 1990, Stephen et al., 2014, Forbes et al., 2015). Signaling through K-Ras is dependent on the bound nucleotide, where the GTP-bound state is active (on) while the GDP-bound state is inactive (off). In GTP-bound K-Ras, P-loop (residues 10-17), switch I (SI, residues 25-40) and switch II (SII, residues 60-74) regions make up the active site whose well-ordered conformations allow effector protein binding for protein signaling (Figure 1). However, oncogenic gain-of-function mutations impair GTP hydrolysis and freeze K-Ras in its active state (Vetter and Wittinghofer, 2001), causing uncontrollable cellular growth and evasion of apoptotic signals (Downward, 2003, Chen et al., 2013). Tumors driven by oncogenic K-Ras are often resistant to standard therapies and result in poor outcomes; they are also excluded from treatment with other targeted therapies, making mutant K-Ras an extremely high priority target in cancer treatment (Pao et al., 2005, Lievre et al., 2006). However, there are still no drugs in the clinic today that directly target mutant K-Ras.

Part of the challenge in oncogenic K-Ras inhibitor design has been due to structure analyses that suggest a lack of well-defined druggable sites on its surface (Zhang and Cheong, 2016). However, studies that have utilized protein dynamics data such as NMR and mass spectrometry have identified binding pockets on specific K-Ras oncogenic mutants and have attempted to stabilize their on or off conformational states (Taveras et al., 1997, Ostrem et al., 2013, Lim et al., 2014, Lito et al., 2016). Accumulating studies suggest that K-Ras proteins are in dynamic and flexible states and their distinct characteristics cannot be identified by structural studies alone (Patricelli et al., 2016, Ostrem et al., 2013, Lim et al., 2014, Singh et al., 2015, Grant et al., 2011b, Leshchiner et al., 2015, Lito et al., 2016). K-Ras dynamic activities in different conformational states, which can also change due to allosteric interactions between protein residues, also need to be quantified (Motlagh et al., 2014).

However, we still do not clearly understand the intra-molecular allosteric networks between distant sites on K-Ras (Marcus and Mattos, 2015). While such allosteric interaction sites have recently been discovered in its catalytic domain (Buhrman et al., 2010, Abankwa et al., 2010, Kearney et al., 2014), they remain largely understudied (Marcus and Mattos, 2015). Understanding allosteric interactions can present novel opportunities for small molecule targeting of mutant K-Ras conformations while sparing those of wild-type K-Ras, which first requires a deeper understanding of the intrinsic K-Ras dynamics.

In allosteric regulation of protein dynamics, correlated motions between protein residues are essential (Goodey and Benkovic, 2008, Kern and Zuiderweg, 2003, Wand, 2001). These motions enable the transfer of fluctuation information through the allosteric network (Kamberaj and van der Vaart, 2009), which inherently involves directionality, or “causality” of events (Guarnera and Berezovsky, 2016). If the motions of two residues are correlated, it would be extremely valuable to identify whether the motions of one residue drive the motions of the other. However, while correlation calculations indicate interaction (which is necessary for allosteric transitions), they are symmetric and do not reveal the direction of information flow. To understand the directionality (or causality) between the motions of two residues, we need to incorporate time-delayed correlation estimations in our analyses (Schreiber, 2000, Guarnera and Berezovsky, 2016).

Here, we introduce a novel method that enables identification of causality between all residue pairs of a protein using time-delayed correlation information. For this purpose, we first record residue fluctuations calculated at every time step of a molecular dynamics (MD) simulation as a time series, and then calculate the time-delayed correlation of a residue pair as the correlation between two time series where one is shifted in time relative to the other. We conclude that the fluctuations of a residue control and modify the fluctuations of the delayed one if we observe a significant correlation between the fluctuations of two residues with a positive time lag. This is a completely novel approach to the analysis of structure-function relations, as well as to understanding K-Ras dynamics. We demonstrate the

77 simplicity of computing time-correlation functions in studying protein dynamics by applying
78 the method to study K-Ras. We specifically focus on K-Ras because of its clinical importance
79 and the availability of a large body of data which enable the validation of our theoretical
80 predictions. Note that while correlations between the fluctuations of residue pairs have
81 already been shown in several Ras protein studies (Grant et al., 2009b, Lukman et al., 2010,
82 Kapoor and Travesset, 2015), despite the significantly large body of literature on K-Ras,
83 there has been surprisingly little attention on the role of causality (or directionality) in
84 correlation dynamics of K-Ras conformations.

85 In summary, we present a comprehensive study of intrinsic K-Ras dynamics, including
86 detailed analyses of causality between the motions of its residues. We first provide detailed,
87 quantitative descriptions of both active and inactive K-Ras from extensive MD simulations.
88 We use a statistical thermodynamics interpretation of fluctuation correlations to quantify K-
89 Ras 'stiffening' upon activation. Using stiffness calculations jointly with measurements of
90 reduced relative fluctuations, we define protein stability and show that K-Ras is more stable
91 in active conformation. To investigate time-dependent characteristics of correlated motions,
92 we map and compare correlated motion patterns of active and inactive K-Ras, then discuss
93 the differences of decay times of correlations between the two K-Ras forms in detail. Our
94 results show that inactive K-Ras is marked by a pronounced decrease in correlated motions
95 of residues for shorter periods, while active K-Ras correlations have longer decay times. We
96 analyze the ensuing events at the atomic scale. Finally, to enable a deeper understanding of
97 K-Ras dynamics, we introduce the first causality calculations for K-Ras and identify specific
98 driver and follower residues during protein simulations.

99 Results and Discussion

100 Comparison of stiffness changes in active and inactive K-Ras

101 **GTP binding increases K-Ras stiffness.** To understand how nucleotide binding effects K-
102 Ras dynamics, we quantified changes in its 'stiffness' – a metric that inversely correlates with

residue pair fluctuations - upon GTP vs. GDP binding. For this purpose, we represented the interaction between two fluctuating residue pairs (*i* and *j)* as a spring with a constant k_{ij}, based on a statistical thermodynamics interpretation we previously developed (Erman, 2015). Plotting this spring constant for every residue pair in both GTP- (Figure 2A) and GDP-bound (Figure 2B) K-Ras, we observe strong coordination in the fluctuations of GTP phosphate groups with those of K-Ras (Figure 2A).

To zoom in on and directly compare the effects of nucleotide binding on K-Ras stiffness, we simply calculated the differences in spring constant values between GTP- and GDP- bound K-Ras. In Figure 2C, red dots indicate that the differences are largely due to the stiffening effects of GTP-binding on residue pair fluctuations. Notice that Regions 1-3 in Figure 2C that correspond to secondary structures show significant increase in k_{ij} when GTP-bound. Furthermore, Region 1 corresponds to strong coordination of β2 and β3 motions, while Regions 2 and 3 correspond to increased stiffness of β4-α3 and α4.

Nucleotide binding affects spring constant of α2 (SII). We next investigated the effects of nucleotide binding on the spring constant of α2 (SII), because previous studies have shown that stiffness increases when SII refolds into an α-helical conformation through GTP binding (Noe et al., 2005). We calculated the spring constants of the two terminal residues of α2 (A66 and T74), which were 0.10 kcal/mol·A^2 (69.91 pN/nm) for active and 0.04 kcal/mol·A^2 (27.78 pN/nm) for inactive K-Ras. Previous studies have utilized various experimental methods that have all led to spring constants within ~0.09-1.15 kcal/mol·A^2 (60–80 pN/nm) for helices (Adamovic et al., 2008, Howard, 2001). Our results for both forms of K-Ras are in the same order of magnitude. Note that for active K-Ras the α2 spring constant is equalent to the characteristic spring constant of α-helices, while it is lower in inactive form. Hence, our results validate and quantify earlier, qualitative observations of Noe *et. Al.* (Noe et al., 2005) that the α2 spring constant reaches to the level of an α-helix spring constant during GTP binding.

Overall spring constant is higher in active complex. To estimate global changes in stiffness in response to nucleotide binding, we calculated the overall spring constants $k_{overall}$ *(details in Methods)* of nucleotide-K-Ras complexes, which were 0.70 kcal/mol·A^2 (481.75 pN/nm) for GTP-bound, and 0.55 kcal/mol·A^2 (385.12 pN/nm) for GDP-bound K-Ras. Both are of the same order of magnitude with an experimental study for another protein, myoglobin, which has an overall spring constant of ~300 pN/m (Rico et al., 2013, Zaccai, 2000) pointing to an order of magnitude agreement of overall stiffnesses of proteins in general. In conclusion, GTP-binding makes K-Ras more rigid.

Secondary structure motions show the strongest coordination with the rest of the protein. Quantifying the spring constant based on fluctuations allows for analyzing how, analogous to a virtual spring, the fluctuations of a specific residue are coupled with fluctuations of rest of the protein. To discover residues whose fluctuations are in strong coordination with K-Ras fluctuations and how they change between the two states, we compared the mean spring constant $\bar{k}_i$ of each residue *i*, for both active and inactive K-Ras (Figure 2D) as described in *Methods*. A large $\bar{k}_i$ value indicates that the motions of residue *i* are stiffly coupled with the motions of the protein; while a small $\bar{k}_i$ value indicates that the motions of the *i*th residue and the protein are flexibly coupled. For simplicity, we categorized the significant mean spring constant $\bar{k}_i$ values as highest, high and smallest (For details please see Table S1). In both states, the highest $\bar{k}_i$ values are of β-strand residues β4, β5 and β6, showing the strongest coordination of their motions with K-Ras motions. Next, high $\bar{k}_i$ values of β1, P-loop and α5 residues indicate that their fluctuations are also strongly coupled with those of the protein. On the other hand, the smallest $\bar{k}_i$ values belong to SII region in active and SI and SII regions in inactive K-Ras which show that their residue fluctuations are not correlated with the rest of the protein (Table S2). Since we have defined the stiffness metric as a signifier of a decrease in residue fluctuations, we provide a second line of proof that increased stiffness stabilizes dynamic fluctuations in both forms of K-Ras by using Root Mean Square Fluctuation (RMSF) graph (Figure S1). Clearly, the residues with

the smallest mean spring constant $\bar{k}_i$ values from Figure 2D have the highest RMSF values in Figure S1 and vice versa.

As indicated in previous studies where NMR and Atomic Force Microscopy were used, protein stiffness depends on secondary structure (Zaccai, 2000, Rico et al., 2013), where loops contribute to structural flexibility and show large fluctuations, while β-strands and α-helices provide mechanical stability and show small fluctuations (Rico et al., 2013). Our K-Ras results are consistent with these general observations. In addition, we observe stiff coupling of the fluctuations of the P-loop and the protein. This observation is important since P-loop is the phosphate binding site of K-Ras and connects β1 and α1 (Figure 1). Although loops are often flexible regions of proteins and show higher fluctuations, in K-Ras, motions of P-loop residues are stiffly coupled to those of the protein, especially in active state ($\bar{k}_i$ =1.08 kcal/mol·A^2 for K-Ras-GTP, $\bar{k}_i$=0.85 kcal/mol·A^2 for K-Ras-GDP).

The mean spring constant values of residues in β2, β3, α3 and switch regions – especially SI- are higher in active K-Ras than in inactive K-Ras. Finally, we calculated mean spring constant differences between active and inactive K-Ras, $\Delta\bar{k}_i$ ($\bar{k}_{i\ \text{K-Ras-GTP}}$ - $\bar{k}_{i\ \text{K-Ras-GDP}}$). Figure 2E shows that the fluctuations of β2 and β3 terminal (D38 and D57) and α3 center (D92-I93) residues are in stronger coordination with those of active K-Ras (vs. inactive K-Ras) ($\Delta\bar{k}_i$ > 0.43 kcal/mol·A^2). Our results also indicate that although residues of switch regions have the smallest $\bar{k}_i$ values in both forms, some of their $\bar{k}_i$ values increase significantly in active form. In Figure 2E, $\Delta\bar{k}_i$ ranges between 0.20-0.36 kcal/mol·A^2 for residues in SI (D30-R41) and 0.02-0.19 kcal/mol·A^2 for residues in SII (G60-T74). These $\Delta\bar{k}_i$ values show stiffer coupling of the motions of GTP-K-Ras with the motions of switch residues, especially SI (vs GDP-K-Ras). This result is important as SI includes the binding site to effector proteins which only bind to GTP-bound K-Ras when SI flexibilty is reduced (Spoerner et al., 2010). Earlier studies that used NMR spectra and RMSF calculation also support our results that GTP binding reduces the flexibility of both SI and SII, especially SI

(Shima et al., 2010, Kapoor and Travesset, 2015). Our results improve on this information by showing that fluctuations of switch regions –notably SI- are more stiffly coupled with K-Ras-GTP fluctuations (Figure 2E).

Comparison of residue pair correlations for active and inactive K-Ras

To identify if the fluctuations of one residue are related to fluctuations of another residue, we calculated the correlations of all residue-residue pairs in both GTP- vs GDP bound K-Ras complexes. As expected, cross-correlation coefficient maps of K-Ras-GTP (Figure 3A) and K-Ras GDP (Figure 3B) exhibit different correlation characteristics. The most remarkable differences between Figure 3A and 3B belong to two parts: (i) the correlation of α1-S1 with L10-α5 and (ii) the correlations between β2 and β3. Positive correlation patterns within these two parts are evident in K-Ras-GTP simulations, but absent in K-Ras GDP simulations. To provide comprehensive information on nucleotide-dependent K-Ras dynamics, we present these two remarkable results from correlation analyses (Figure 3) as well as sources of correlated motions (i.e. H-bonds) together in the following sections.

The correlation of α1-S1 with L10-α5 in active K-Ras motions is due to three specific H-bonds. MD simulations show that the correlation between α1-S1and L10-α5 in the active form results from GTP binding to active site residues, which also form specific H bonds with other K-Ras residues and water. Based on the average number of H-bonds each residue forms throughout the simulation, we estimated that the nucleotides remain bound to active site residues S17, D30, D119 and K147, and that GTP-binding (vs. GDP) is more stable for S17 and D30 (Table S3). Furthermore, correlated motions of α1-S1 and L10-α5 in GTP-bound K-Ras originate specifically from three H-bonds: (i) A146-Q22, (ii) D30-GTP, (iii) D30-a water molecule. We observed a sustained H-bond between A146-Q22 during active but not in inactive complex simulation. This suggests that A146-Q22 interaction causes a strong relationship between L10α5 (A146-D154) and α1 (L19-I24) in active K-Ras (Figure 3A) with a correlation coefficient of 0.75, and a weak correlation coefficient of 0.28 for inactive K-Ras. At

the same time, the active site residue D30 forms an H-bond with the nucleotide in both active and inactive K-Ras, while it also binds to a water molecule only in the active form. However, the H-bond in the active form between D30(O)-GTP(O2A) is more permanent than the H-bond in the inactive form between D30(O)-GDP(O2').

Since H-bond of D30-GTP is permanent, D30-GTP distance is invariant and their fluctuation correlations have longer decay times during K-Ras-GTP simulation. We next combined cross-correlation results with the distance distribution of D30 and nucleotides and quantified the decay times of their correlations during MD simulations. In addition to more permanent binding of D30(O)-GTP(O2A), nucleotide-D30 distance distribution pattern is close to the normal distribution curve with a mean of a smaller value in active K-Ras (Figure 4A), with a correlation coefficient of 0.97. To quantify decay time of this correlation in both complexes, we first defined two "connectivity vectors", $\Delta R_{30\text{-}GTP}$ and $\Delta R_{30\text{-}GDP}$, between D30(O) and nucleotides. As illustrated in Figure 4B, $\Delta R_{30\text{-}GTP}$ connects the starting point of fluctuation vector of $\Delta R_{D30(O)}$ to end point of negative $\Delta R_{GTP(O2A)}$; $\Delta R_{30\text{-}GDP}$ starts from $\Delta R_{D30(O)}$ to negative $\Delta R_{GDP(O2')}$. We then calculated time-delayed autocorrelations of each connectivity vector throughout the MD simulations. The autocorrelation plot in Figure 4B summarizes the correlation of connectivity vectors at various time delays, where vector correlation coefficients are plotted with 1 ns delays at a time; slow decay of correlations in active K-Ras is clearly observed. Correlations decay to 1/e in about 3 ns for K-Ras-GDP (red line), vs. to ~10 ns for K-Ras-GTP (black line). One reason for this slow correlation decay is the H-bond, which binds D30 to a water molecule in active K-Ras. The O atom of D30 establishes an H-bond with the nearest water during 28% of the trajectory while it does not make any contact with waters when K-Ras is inactive.

A continuous H-bond stabilizes β2-β3 distance and promotes longer decay times for β2-β3 correlations during K-Ras-GTP simulation. β2 and β3 are two parallel β strands located between SI and SII regions (Figure 5A). Due to the presence of a persistent H-bond between R41(β2)-D54(β3) in K-Ras-GTP simulation, the peak value of $R_{41\text{-}54}$ distribution

decreases (Figure 5B) and fluctuations of β2 and β3 become correlated (Figure 3A). The time-delayed autocorrelations of the vector $\Delta R_{38\text{-}57}$ between their terminal residues D38 and D57 are presented in Figure 5C showing that $\Delta R_{38\text{-}57}$ correlation decays much more slowly in active K-Ras.

Causality of Correlated Motions

Correlated motions of proteins often have a direction or causal relationship (Kamberaj and van der Vaart, 2009). Correlations in the fluctuations of two residues indicate interaction, which is necessary for allosteric transitions. However, this is not sufficient for understanding the dynamic phenomenon completely since these symmetric correlations do not contain information on driver and follower relationships. To deduce causality, time-delayed correlations need to be analyzed. Our trajectory analysis shows that the symmetry assumption on time-delayed correlations does not hold (Callen, 1985) for several residue pairs i.e., time delayed correlations of fluctuations of two atoms are not the same in forward or backwards in time. Our observation is supported by recent work (Kamberaj and van der Vaart, 2009) that identified causality in correlated motions from MD simulations using an information theory measure of transfer entropy. This work, in turn, was built on a study by Schreiber, who introduced the entropy transfer concept for fluctuating environments (Schreiber, 2000). We follow up on these ideas and introduce a new method to dissect time-dependent correlations of all residue pairs of a protein to identify driver and follower residues. For this purpose, we evaluate strong time-delayed (τ=5ns) correlations between residue pairs (defined as < -0.3 or >0.6). The strongest causal relations are as follows (Figure 6):

SII motions drive SI in active K-Ras. SI-SII relationship is better understood by examining residues that drive their motions throughout the trajectory. Our causality calculations show that SI is driven by SII (Figure 6A and 7). We present time-delayed correlation plots of R68(SII) with V29(SI) (Figure 7A) and with P34(SI) (Figure 7B) for active K-Ras. Red curve shows that the fluctuations of R68 at time t affect the fluctuations of V29 at time $t+\tau$.

Fluctuation decay of K-Ras residues is in the order of 1ns. The red curve persists for time periods that are an order of magnitude longer. The reverse does not show a significant correlation: V29 does not correlate with later fluctuations of R68. Previously, a study reported that SI loop at residues 29–34 swings into the water using V29 and P34 as hinges during Ras inactivation (Noe et al., 2005). We improved on this information by calculating time-delayed correlations and identified that SII residues - especially R68 and D69- sustain active state conformation of SI by driving the motions of hinge residues V29 and P34. Another study also assessed the conformational transition of Ras from inactive to active state (Gorfe et al., 2008), where displacement of SII triggers the active state transition and SI follows SII after a lag time of multiple nanoseconds. Dominance of SII region motions was also observed in several studies (Clausen et al., 2015, Grant et al., 2009a). The nucleotide-bound form behavior is regulated by the relative arrangement of the two switches, rather than their individual conformations. We quantified this by verifying that SI fluctuations follow SII fluctuations in K-Ras-GTP. Since from an information theoretic point of view correlations are regarded as information sources, we conclude that information flows from SII to SI. The directionality originates from the differences in the characteristic decay times. The problem is therefore a problem of dynamics within few nanosecond time periods. Disruption of this flow is expected to interfere with the switch mechanism function, which is the basis of K-Ras activity.

α3 and Loop 7 (L7) motions drive switch region (SI & SII) motions in active K-Ras (Figure 6B). Fluctuations of the helical dimer interface residues of α3, E98 and R102 (Muratcioglu et al., 2015) drive fluctuations of A66 (α2; SII), as shown in Figures 8C and 8D. Additionally, helical dimer interface residue S106 (L7) drives the motion of Y71 (α2; SII) (Figure 8E). On the other hand, fluctuations of R102 (α3) and S106 (L7) drive SI residues N26, D30, Y32 (Figures 8F to 8H).

Correlated motions of α2 and α3-L7 have been described in other studies, which also emphasized the necessity of understanding their effect on protein function (Grant et al.,

2009b, Clausen et al., 2015, Grant et al., 2011b). We contribute to this knowledge by identifying their cause and effect relations. Furthermore, in previous studies, starting from the allosteric interaction between α2 and α3-L7, a novel ligand binding pocket, termed p3, which includes residues of L7 was defined and targeted for lead generation (Grant et al., 2011b, Spoerner et al., 2005, Rosnizeck et al., 2010). It was reported that ligand binding to p3 pocket weakens effector protein binding by allosterically stabilizing Ras effector binding site (SI). Another proposed allosteric mechanism is that ligand binding to p3 pocket changes the switch region conformation. Our results implicate the role of allosteric modulation of ligand binding, which may freeze the fluctuations of L7 and stabilize SI motions. This is based on our finding that motions of effector binding site (D30-Y32) are driven by S106 (L7).

β2-β3 are both drivers and followers in active K-Ras. Causality calculations suggest the following information flow in fluctuations: ILE21-GLN22 (α1) drives β2-β3 (Figure 6 and S2); which drives Y157 (α5), Q61 (SII) and T74 (SII) (Figure 6 and S3, with details in Table S4). Specifically, the differences in the characteristic decay times in Figures S3A-B demonstrate that information flows from β2-β3 to Y157 (α5). These findings improve on the previous observations of Abankwa et al. where they defined β2–β3 and α5 as a novel conformational switch (Abankwa et al., 2008). Most importantly, we showed that Q61 (SII) motions follow E49 motions (β2-β3) (Figure S3D). Abankwa et al. also observed that mutations in D47-E49 cause hyperactive Ras. Our findings support this too by showing that fluctuations of E49 of the wild type cause fluctuations of the catalytic residue Q61 within SII, whose proper positioning is essential for effective catalysis (Ito et al., 1997). Based on these results, we suggest that mutations in D47-E49 region may alter E49 fluctuations that cause improper Q61 fluctuations. Therefore, GTP catalysis is disrupted which results in constituently active K-Ras.

Conclusions

We present a novel approach that combines several distinct analysis methods to quantify in detail dynamics of GTP and GDP bound K-Ras, for which a significant amount of experimental and theoretical data already exists in the literature to test our predictions. Oncogenic K-Ras is an extremely high priority drug target in cancer treatment. In order to develop new direct inhibitors that selectively bind to mutant K-Ras conformations while sparing those of WT K-Ras, it is necessary to first understand the dynamic activity of the WT protein in detail. To evaluate the nucleotide binding dependent changes in K-Ras stability, we used stiffness and RMSF calculations and proved that GTP binding rigidifies and hence stabilizes K-Ras motions. These results are in agreement with previous experimental and computational K-Ras studies (Kapoor and Travesset, 2015, Raimondi et al., 2011, Diaz et al., 1995).

Our calculations that use stiffness, RMSF and correlation graphs (Figure 1-2, S1) confirm that GTP binding increases K-Ras stiffness and thereby decreases fluctuation amplitudes, leading to distinct correlation patterns. These striking changes in GTP bound-K-Ras dynamics enable its GTPase activity. Note that this nucleotide exchange is the first step in active to inactive transition (Prakash et al., 2012, Prakash et al., 2015, Zhang et al., 2012, Buhrman et al., 2011, Prakash and Gorfe, 2013, Grant et al., 2011a, Edreira et al., 2009). Overall, our results support the well-established allosteric nature of K-Ras activation (Prakash and Gorfe, 2013), which has been suggested to play an important role in GTPase activity (Grant et al., 2010). Although correlated fluctuations are necessary for allosteric information flow, their longer correlation decay times are also of crucial importance for complete allosteric transition. We calculated time-dependent autocorrelations of fluctuation vectors between residue pairs and discovered that correlations of K-Ras-GTP are stronger and persist for longer correlation times during simulations. Their persistency may allow complete allosteric information flow in K-Ras-GTP.

We broadened our analysis to quantify causality in allosteric regulation of K-Ras function. The most important results from our study are on causality. We applied a simple but powerful

method that we defined as *time-delayed correlation* into protein dynamics. To understand K-Ras dynamics, we investigated whether fluctuations of any residue caused fluctuations of another. Our results revealed the information flow in K-Ras switch mechanism and that SII fluctuations drive SI fluctuations. This prediction is an essential validation of our approach, since the dominance of SII motions over SI motions was observed in previous experimental and computational studies (Clausen et al., 2015, Grant et al., 2009a). Surprisingly, in addition to the canonical switch mechanism, our algorithm also revealed causality relations in the novel switch mechanism that includes β2–β3 and α5, where β2–β3 motions drive α5. Moreover, fluctuations of α3-L7 drive fluctuations of SI and SII. Interestingly, previous studies reported that Ras effector binding site (SI) is allosterically stabilized by ligand binding into a novel pocket that includes L7 (Grant et al., 2011b, Spoerner et al., 2005, Rosnizeck et al., 2010). Our results explain the allosteric effect of ligand binding on SI motions by showing the information flow from L7 to SI.

Note that functionally, the identified driver and follower sites do not show an enrichment trend in oncogenic mutations observed in human cancers (from 2266 missense K-Ras mutations observed in all cancers within cBioPortal www.cbioportal.org on July 28, 2016). However, the motions of residue Q61 (SII region), which is the second most frequently mutated residue in cancer patients (113/2266 missense mutations) are driven by those of E49 (β3). While there are no known oncogenic mutations of residue E49, experiments have shown that mutating E49 leads to hyperactive Ras (Abankwa et al., 2008), consistent with what we would expect from a driver of Q61 motions, which also causes the same effect.

Our ongoing research on mutated K-Ras suggests deviations in dynamics from that of WT K-Ras. Any such investigation would first necessitate a deep understanding of the causal relationships in intrinsic K-Ras motions. Understanding intrinsic WT dynamics is imperative and serves as a necessary reference. Our objective behind this detailed analysis is to provide such a reference for any future mutant K-Ras studies. It would be of interest to identify how these findings change when there is a mutation on the protein. Only after

identifying these differences can we discover molecules that can eliminate the unfavorable changes caused by the mutations.

The computational tools we introduce in the present work are easily applicable to the analysis of simulation data from different proteins to understand causality in their allosteric regulations which can then be utilized in drug discovery. From this perspective, our approach sets a novel paradigm for drug design that directs attention to changes in protein dynamics. The latter is in close relation to changes in protein function whose restoration to normal is the target of all drug design activities.

Experimental Procedures

MD Simulations

We performed all-atom MD simulations for both Mg^{+2}GDP- and Mg^{+2}GTP-bound K-Ras. As the initial point of all simulations, we used the crystal structure corresponding to PDB ID 4OBE (K-Ras-GDP). For K-Ras-GTP structure, we simply substituted the GDP of K-Ras-GDP structure with GTP using Discovery Studio 4.5 software (BIOVIA, 2015). We solvated each protein in a water box (TIP3 water) where buffering distance between box edges and protein was 12 Å. We applied periodic boundary conditions and added counter-ions to neutralize the system. We used a 2 fs time-step with a 12Å cutoff for Van der Waals interactions and full particle-mesh Ewald electrostatics. We carried out all computations in dynamics procedure for an N, P (1 atm), T (310K) ensemble. We used NAMD 2.10 with AMBER ff99SB and general amber force fields (GAFF). We obtained parameters of GTP and GDP (see Supplemental Experimental Procedures). The initial system energy was first minimized for 10,000 steps, followed by 10,000 steps for equilibration. After equilibration, we performed 300 ns simulations. Atomic coordinates $\hat{R}$ of all atoms were saved every 1 ps. To eliminate all rotational and translational motions, we aligned the trajectories to the initial structure by using VMD software 1.9.2 (Humphrey et al., 1996). We visualized trajectories using VMD.

393 **Stiffness**

394 We quantified nucleotide-bound K-Ras stiffness using a statistical thermodynamics
395 interpretation of fluctuation correlations (Erman, 2015). We assumed that the interaction
396 between two fluctuating residues *i* and *j* can be represented by a spring, where the spring
397 constant follows from the Gaussian Network Model (GNM) (Haliloglu et al., 1997): $k_{ij} =$
398 $\frac{k_\beta T}{\langle(\Delta R_i)^2\rangle - 2\langle(\Delta R_i \Delta R_j)\rangle + \langle(\Delta R_j)^2\rangle}$ where ΔR_i is the instantaneous fluctuation of one end of the rod, ΔR_j
399 is the fluctuation of the other end (Details in Supplementary), k_B is the Boltzmann constant
400 and *T* is the absolute temperature. The spring constant has dimensions of force/length. In
401 GNM spring definition, each residue *i* is attached to *N-1* other residues via *N-1* springs
402 (Haliloglu et al., 1997). Thus, how stiffly a residue *i* is attached to a protein can be quantified
403 by $\bar{k}_i = \sum_j k_{ij}/N - 1$ where $\bar{k}_i$ is the mean spring constant for a residue *i*. For stiffness
404 estimates of the complete complexes, we define an overall stiffness parameter $k_{overall}$ by the
405 expression $k_{overall} = \sum_{j>i} k_{ij}$. To estimate the stiffness differences in active versus inactive K-
406 Ras, we calculated $\bar{k}_i$ for each residue and $k_{overall}$ for the protein for both states.

407 **Stability**

408 We defined the stability of an interacting system of residues as the joint state of reduced
409 RMSF and increased interaction stiffness. RMSF relates to the magnitude of fluctuations of
410 individual residues and stiffness relates to the distance between two residues and therefore
411 they are two independent quantities. (For details see Supplemental Experimental
412 Procedures). A small RMSF and a high stiffness denote increased stability.

413 **Distance distributions between residue pairs**

414 We calculated the distance between two residues (*i*,*j*) as $R_{ij} = \sqrt{(R_i(t) - R_j(t))^2}$. Residue
415 pair distance distributions $W(R_{ij})$ were calculated by dividing the maximum distance between

the pair into small bins and counting the number of observed distances in each bin. All distributions were normalized.

Time independent correlations (cross-correlation coefficient map): Correlations intrinsic to K-Ras structure are defined by the cross-correlation coefficient map, $C(\Delta R_i, \Delta R_j)$:

$$C_{ij} = C(\Delta Ri, \Delta Rj) = \frac{\langle \Delta R_i(t) \cdot \Delta R_j(t) \rangle}{\langle (\Delta R_i(t))^2 \rangle^{1/2} \langle (\Delta R_j(t))^2 \rangle^{1/2}} = \frac{\sum_{t=1}^{N_t-\tau} \Delta R_i(t) \cdot \Delta R_j(t)}{\left[\sum_{t=1}^{N_t-\tau} (\Delta {R_i}^2(t))\right]^{1/2} \left[\sum_{t=1}^{N_t-\tau} (\Delta {R_j}^2(t))\right]^{1/2}}$$

where $\cdot$ denotes the dot product. Correlation varies between -1 and 1. If motions of two atoms are independent, then $\langle \Delta R_i(t) \cdot \Delta R_j(t) \rangle = 0$ and C_{ij}= 0. If the atoms always move in parallel in the same direction, then they are perfectly positively correlated, and C_{ij} = 1. If they always move in parallel in opposite directions, they are perfectly negatively correlated, and C_{ij} = -1. Cross-correlation coefficients lie in the range of -1 ≤ C_{ij} ≤1.

Time delayed correlations, mobility and causality

Time-delayed correlation of two fluctuations is defined by:

$$C_{ij}(\tau) = \frac{\dfrac{\langle \Delta R_i(t) \cdot \Delta R_j(t+\tau) \rangle}{\langle (\Delta R_i(t))^2 \rangle^{1/2} \langle (\Delta R_j(t+\tau))^2 \rangle^{1/2}}}{C_{ij}} = \frac{\dfrac{\sum_{t=1}^{N_t-\tau} \Delta R_i(t) \cdot \Delta R_j(t+\tau)}{\left[\sum_{t=1}^{N_t-\tau} (\Delta {R_i}^2(t))\right]^{1/2} \left[\sum_{t=1}^{N_t-\tau} (\Delta {R_j}^2(t+\tau))\right]^{1/2}}}{C_{ij}}$$

where $C_{ij}(\tau)$ represents the correlation of ΔR_j at time $t + \tau$ and earlier values of ΔR_i at time t. Similarly, if indices are exchanged, then $C_{ji}(\tau)$ represents the correlations of ΔR_i at time $t + \tau$ with earlier values of ΔR_j at time t. If the fluctuations of residue i drive the fluctuations of residue j, then $C_{ij}(\tau) > C_{ji}(\tau)$. If $C_{ji}(\tau) > C_{ij}(\tau)$, residue j drives residue i because the fluctuation ΔR_j at time t is correlated with future fluctuations of ΔR_i. However, at τ=0, the equality $C_{ij}(0)$=$C_{ji}(0)$ holds.

Note that time-delayed autocorrelation $C_{ii}(\tau)$ is the correlation of the trajectory with its own past and future coordinates. If autocorrelation is large, this could correspond to a specific

form of “persistence”, a tendency for a system to remain in the same state from one observation to the next.

Acknowledgements

SV acknowledges 2214-International Doctoral Research Fellowship funding from The Scientific and Technological Research Council of Turkey (TUBITAK). ZHG and SV were supported by the LUNGevity Foundation and the start-up funds to ZHG from Icahn School of Medicine at Mount Sinai.

Figure Legends

Figure 1. Crystal structure of wild-type K-Ras protein in GTP-bound state (PDB: 4OBE). A) K-Ras structure ribbon representation with secondary structures in blue for α-helices and green for β-sheets. B) Schematic of K-Ras sequences (residues 1-169). Functional regions are in same color used in K-Ras structure in A.

Figure 2. Stiffness results for GTP- and GDP-bound K-Ras and their difference. In panels A and B, both axes marks 1-169 represent the residue Cα atoms of K-Ras and marks 170-on represent GDP and GTP nucleotide heavy atoms, respectively with k_{ij} >1.5 kcal/mol·A^2. (A) k_{ij} for K-Ras-GTP. Atoms 170-181 are the γ, β, α-phosphate groups and 182-201 are the guanine atoms of GTP. (B) k_{ij} for K-Ras-GDP. Atoms 170-178 are the β and α phosphate groups and 178-197 are the guanine atoms of GDP. (C) Difference between active and inactive K-Ras k_{ij} values. Red regions are stiffer in K-Ras-GTP (k_{ij} values of K-Ras-GTP are greater than K-Ras-GDP at least 0.75 kcal/mol·A^2) and blue regions are stiffer in K-Ras-GDP (k_{ij} values of K-Ras-GDP are greater than K-Ras-GTP at least 0.75 kcal/mol·A^2). (D) Mean spring constants $\bar{k}_i$ for the GTP and GDP bound states. (E) Mean spring constant differences $\Delta\bar{k}_i$ for the GTP and GDP bound states. Positive values correspond to larger mean stiffness in K-Ras-GTP.

Figure 3. Cross-correlation coefficient maps for GTP and GDP bound states. Red dots show positive correlations ($1 \geq C(\Delta R_i, \Delta R_j) \geq 0.6$) and blue dots show negative correlations ($-0.45 \geq C(\Delta R_i, \Delta R_j) \geq -1$). Residues indexed 1-169 belong to K-Ras. (A) Correlated fluctuations of K-Ras-GTP. Indices between 170-201 refer to GTP heavy atoms (182-201 are guanine atoms). (B) Correlated fluctuations of K-Ras-GDP. Indices between 170-197 refer to GDP heavy atoms (178-197 are guanine atoms).

Figure 4. D30-GTP distance is more stable than that of D30-GDP. Fluctuations of D30(O) to GTP(O2A) "connecting vector" are persistently correlated. (A) Distance distribution between D30 and connecting O atoms of GTP (black) and GDP (red) (B)Time delayed autocorrelations for the vector connecting Oxygen atom of D30 to O2A of GTP (black curve) and O2' of GDP (red curve). X-axis is the time delay (τ) and Y axis is the time delayed autocorrelation of the vector for τ.

Figure 5. Correlation of β2 and β3 fluctuations is persistent in active K-Ras. (A) Locations of R41 (β2) and D54 (β3) relative to SI & SII. (B) Distance distribution between Cα atoms of R41 and D54 in K-Ras-GTP (black) and K-Ras-GDP (red). Distance values between R41 (β2) and D54 (β3) populate at 3.90 Å during GTP binding whereas these distance values populate at 5.46 Å for GDP-bound K-Ras. (C) Time delayed autocorrelations for the fluctuations of the vector from D38 (Cα) to D57 (Cα).

Figure 6. Causality relations in active K-Ras motions. Directionality in causal relationships is illustrated with arrows. Arrows start from driver residues and end at follower residues. Both residue types are represented with yellow spheres and marked with their residue numbers. The secondary structures they belong to are in turquoise. A) R68 (SII) drives V29 and P34(SI). B) E98 and R102 (α3) drive A66 (α2; SII). S106 (L7) drives Y71 (α2; SII). R102 (α3) drives N26 and Y32(SI). S106 (L7) drives D30. C) ILE21-GLN22 (α1) drives β2-β3. D) I46 and D47 (**β2-β3)** drive Y157 (α5).

Figure 7. SII fluctuations drive SI fluctuations; α3-L7 motions drive switch region (SI & SII) motions in K-Ras-GTP. Red curves for $\langle \Delta R_i(t) \Delta R_j(t+\tau) \rangle$ show that the fluctuations of residue *i* at time *t* affect the fluctuations of residue *j* at a later time *t*+τ. (A) R68 (SII) drives V29(SI). (B) R68 drives P34(SI). (C) E98(α3) drives A66 (α2; SII). (D) R102 (α3) drives A66. (E) S106 (L7) drives Y71 (α2; SII). (F) R102 (α3) drives N26(SI). (G) R102 drives Y32(SI). (H) S106 (L7) drives D30(SI).

Reference List

ABANKWA, D., GORFE, A. A., INDER, K. & HANCOCK, J. F. 2010. Ras membrane orientation and nanodomain localization generate isoform diversity. *Proc Natl Acad Sci U S A,* 107**,** 1130-5.

ABANKWA, D., HANZAL-BAYER, M., ARIOTTI, N., PLOWMAN, S. J., GORFE, A. A., PARTON, R. G., MCCAMMON, J. A. & HANCOCK, J. F. 2008. A novel switch region regulates H-ras membrane orientation and signal output. *Embo Journal,* 27**,** 727-735.

ADAMOVIC, I., MIJAILOVICH, S. M. & KARPLUS, M. 2008. The elastic properties of the structurally characterized myosin II S2 subdomain: a molecular dynamics and normal mode analysis. *Biophys J,* 94**,** 3779-89.

BIOVIA, D. S. 2015. Discovery Studio Modeling Environment. Release 4.5 ed. San Diego: Dassault Systèmes.

BUHRMAN, G., HOLZAPFEL, G., FETICS, S. & MATTOS, C. 2010. Allosteric modulation of Ras positions Q61 for a direct role in catalysis. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 4931-4936.

BUHRMAN, G., O'CONNOR, C., ZERBE, B., KEARNEY, B. M., NAPOLEON, R., KOVRIGINA, E. A., VAJDA, S., KOZAKOV, D., KOVRIGIN, E. L. & MATTOS, C. 2011. Analysis of Binding Site Hot Spots on the Surface of Ras GTPase. *Journal of molecular biology,* 413**,** 773-789.

CALLEN, H. B. 1985. *Thermodynamics and an introduction to thermostatistics,* New York ; Chichester, Wiley.

CHEN, C. C., ER, T. K., LIU, Y. Y., HWANG, J. K., BARRIO, M. J., RODRIGO, M., GARCIA-TORO, E. & HERREROS-VILLANUEVA, M. 2013. Computational analysis of KRAS mutations: implications for different effects on the KRAS p.G12D and p.G13D mutations. *PLoS One,* 8**,** e55793.

CLAUSEN, R., MA, B. Y., NUSSINOV, R. & SHEHU, A. 2015. Mapping the Conformation Space of Wildtype and Mutant H-Ras with a Memetic, Cellular, and Multiscale Evolutionary Algorithm. *Plos Computational Biology,* 11.

DIAZ, J. F., WROBLOWSKI, B. & ENGELBORGHS, Y. 1995. MOLECULAR-DYNAMICS SIMULATION OF THE SOLUTION STRUCTURES OF HA-RAS-P21 GDP AND GTP COMPLEXES - FLEXIBILITY, POSSIBLE HINGES, AND LEVERS OF THE CONFORMATIONAL TRANSITION. *Biochemistry,* 34**,** 12038-12047.

DOWNWARD, J. 2003. Targeting ras signalling pathways in cancer therapy. *Nature Reviews Cancer,* 3**,** 11-22.

EDREIRA, M. M., LI, S., HOCHBAUM, D., WONG, S., GORFE, A. A., RIBEIRO-NETO, F., WOODS, V. L. & ALTSCHULER, D. L. 2009. Phosphorylation-induced Conformational Changes in Rap1b: ALLOSTERIC EFFECTS ON SWITCH DOMAINS AND EFFECTOR LOOP. *The Journal of Biological Chemistry,* 284**,** 27480-27486.

ERMAN, B. 2015. Effects of ligand binding upon flexibility of proteins. *Proteins,* 83**,** 805-8.

FORBES, S. A., BEARE, D., GUNASEKARAN, P., LEUNG, K., BINDAL, N., BOUTSELAKIS, H., DING, M., BAMFORD, S., COLE, C., WARD, S., KOK, C. Y., JIA, M., DE, T., TEAGUE, J. W., STRATTON, M. R., MCDERMOTT, U. & CAMPBELL, P. J. 2015. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res,* 43**,** D805-11.

GOODEY, N. M. & BENKOVIC, S. J. 2008. Allosteric regulation and catalysis emerge via a common route. *Nat Chem Biol,* 4**,** 474-82.

GORFE, A. A., GRANT, B. J. & MCCAMMON, J. A. 2008. Mapping the nucleotide and isoform-dependent structural and dynamical features of ras proteins. *Structure,* 16**,** 885-896.

GRANT, B. J., GORFE, A. A. & MCCAMMON, J. A. 2009a. Ras Conformational Switching: Simulating Nucleotide-Dependent Conformational Transitions with Accelerated Molecular Dynamics. *Plos Computational Biology,* 5.

GRANT, B. J., GORFE, A. A. & MCCAMMON, J. A. 2009b. Ras conformational switching: simulating nucleotide-dependent conformational transitions with accelerated molecular dynamics. *PLoS Comput Biol,* 5**,** e1000325.

GRANT, B. J., LUKMAN, S., HOCKER, H. J., SAYYAH, J., BROWN, J. H., MCCAMMON, J. A. & GORFE, A. A. 2011a. Novel Allosteric Sites on Ras for Lead Generation. *PLoS ONE,* 6**,** e25711.

GRANT, B. J., LUKMAN, S., HOCKER, H. J., SAYYAH, J., BROWN, J. H., MCCAMMON, J. A. & GORFE, A. A. 2011b. Novel Allosteric Sites on Ras for Lead Generation. *Plos One,* 6.

GRANT, B. J., MCCAMMON, J. A. & GORFE, A. A. 2010. Conformational Selection in G-Proteins Lessons from Ras and Rho. *Biophysical Journal,* 99**,** L87-L89.

GUARNERA, E. & BEREZOVSKY, I. N. 2016. Structure-Based Statistical Mechanical Model Accounts for the Causality and Energetics of Allosteric Communication. *PLoS Comput Biol,* 12**,** e1004678.

HALILOGLU, T., BAHAR, I. & ERMAN, B. 1997. Gaussian dynamics of folded proteins. *Physical Review Letters,* 79**,** 3090-3093.

HOWARD, J. 2001. *Mechanics of Motor Proteins and the Cytoskeleton*, Sinauer Associates, INC.

HUMPHREY, W., DALKE, A. & SCHULTEN, K. 1996. VMD: Visual molecular dynamics. *Journal of Molecular Graphics & Modelling,* 14**,** 33-38.

ITO, Y., YAMASAKI, K., IWAHARA, J., TERADA, T., KAMIYA, A., SHIROUZU, M., MUTO, Y., KAWAI, G., YOKOYAMA, S., LAUE, E. D., WALCHLI, M., SHIBATA, T., NISHIMURA, S. & MIYAZAWA, T. 1997. Regional polysterism in the GTP-bound form of the human c-Ha-Ras protein. *Biochemistry,* 36**,** 9109-19.

KAMBERAJ, H. & VAN DER VAART, A. 2009. Extracting the Causality of Correlated Motions from Molecular Dynamics Simulations. *Biophysical Journal,* 97**,** 1747-1755.

KAPOOR, A. & TRAVESSET, A. 2015. Differential dynamics of RAS isoforms in GDP- and GTP-bound states. *Proteins,* 83**,** 1091-106.

KEARNEY, B. N., JOHNSON, C. W., ROBERTS, D. M., SWARTZ, P. & MATTOS, C. 2014. DRoP: A Water Analysis Program Identifies Ras-GTP-Specific Pathway of Communication between Membrane-Interacting Regions and the Active Site. *Journal of Molecular Biology,* 426**,** 611-629.

KERN, D. & ZUIDERWEG, E. R. 2003. The role of dynamics in allosteric regulation. *Curr Opin Struct Biol,* 13**,** 748-57.

LESHCHINER, E. S., PARKHITKO, A., BIRD, G. H., LUCCARELLI, J., BELLAIRS, J. A., ESCUDERO, S., OPOKU-NSIAH, K., GODES, M., PERRIMON, N. & WALENSKY, L. D. 2015. Direct inhibition of oncogenic KRAS by hydrocarbon-stapled SOS1 helices. *Proceedings of the National Academy of Sciences of the United States of America,* 112**,** 1761-1766.

LIEVRE, A., BACHET, J., BOIGE, V., LANDI, B., EMILE, J., COTE, J., TOMASIC, G., ROUGIER, P., PENAULT-LLORCA, F. & LAURENT-PUIG, P. 2006. KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer. *Annals of Oncology,* 17**,** 42-42.

LIM, S. M., WESTOVER, K. D., FICARRO, S. B., HARRISON, R. A., CHOI, H. G., PACOLD, M. E., CARRASCO, M., HUNTER, J., KIM, N. D., XIE, T., SIM, T., JANNE, P. A., MEYERSON, M., MARTO, J. A., ENGEN, J. R. & GRAY, N. S. 2014. Therapeutic targeting of oncogenic K-Ras by a covalent catalytic site inhibitor. *Angew Chem Int Ed Engl,* 53**,** 199-204.

LITO, P., SOLOMON, M., LI, L. S., HANSEN, R. & ROSEN, N. 2016. Allele-specific inhibitors inactivate mutant KRAS G12C by a trapping mechanism. *Science,* 351**,** 604-608.

LUKMAN, S., GRANT, B. J., GORFE, A. A., GRANT, G. H. & MCCAMMON, J. A. 2010. The Distinct Conformational Dynamics of K-Ras and H-Ras A59G. *Plos Computational Biology,* 6.

MARCUS, K. & MATTOS, C. 2015. Direct Attack on RAS: Intramolecular Communication and Mutation-Specific Effects. *Clinical Cancer Research,* 21**,** 1810-1818.

MOTLAGH, H. N., WRABL, J. O., LI, J. & HILSER, V. J. 2014. The ensemble nature of allostery. *Nature,* 508**,** 331-339.

MURATCIOGLU, S., CHAVAN, T. S., FREED, B. C., JANG, H., KHAVRUTSKII, L., FREED, R. N., DYBA, M. A., STEFANISKO, K., TARASOV, S. G., GURSOY, A., KESKIN, O.,

589 TARASOVA, N. I., GAPONENKO, V. & NUSSINOV, R. 2015. GTP-Dependent K-Ras
590 Dimerization. *Structure,* 23, 1325-1335.
591 NOE, F., ILLE, F., SMITH, J. C. & FISCHER, S. 2005. Automated computation of low-energy
592 pathways for complex Rearrangements in proteins: Application to the conformational switch of
593 ras p21. *Proteins-Structure Function and Bioinformatics,* 59, 534-544.
594 OSTREM, J. M., PETERS, U., SOS, M. L., WELLS, J. A. & SHOKAT, K. M. 2013. K-Ras(G12C)
595 inhibitors allosterically control GTP affinity and effector interactions. *Nature,* 503, 548-+.
596 PAO, W., WANG, T. Y., RIELY, G. J., MILLER, V. A., PAN, Q. L., LADANYI, M., ZAKOWSKI, M.
597 F., HEELAN, R. T., KRIS, M. G. & VARMUS, H. E. 2005. KRAS mutations and primary
598 resistance of lung adenocarcinomas to gefitinib or erlotinib. *Plos Medicine,* 2, 57-61.
599 PATRICELLI, M. P., JANES, M. R., LI, L. S., HANSEN, R., PETERS, U., KESSLER, L. V., CHEN, Y.,
600 KUCHARSKI, J. M., FENG, J., ELY, T., CHEN, J. H., FIRDAUS, S. J., BABBAR, A., REN, P.
601 & LIU, Y. 2016. Selective Inhibition of Oncogenic KRAS Output with Small Molecules Targeting
602 the Inactive State. *Cancer Discov,* 6, 316-29.
603 PRAKASH, P. & GORFE, A. A. 2013. Lessons from computer simulations of Ras proteins in solution
604 and in membrane. *Biochimica et biophysica acta,* 1830, 10.1016/j.bbagen.2013.07.024.
605 PRAKASH, P., HANCOCK, J. F. & GORFE, A. A. 2015. Binding hotspots on K-Ras: consensus ligand
606 binding sites and other reactive regions from probe-based molecular dynamics analysis. *Proteins,*
607 83, 898-909.
608 PRAKASH, P., SAYYED-AHMAD, A. & GORFE, A. A. 2012. The Role of Conserved Waters in
609 Conformational Transitions of Q61H K-ras. *PLoS Computational Biology,* 8, e1002394.
610 RAIMONDI, F., PORTELLA, G., OROZCO, M. & FANELLI, F. 2011. Nucleotide Binding Switches
611 the Information Flow in Ras GTPases. *Plos Computational Biology,* 7.
612 RICO, F., RIGATO, A., PICAS, L. & SCHEURING, S. 2013. Mechanics of proteins with a focus on
613 atomic force microscopy. *J Nanobiotechnology,* 11 Suppl 1, S3.
614 ROSNIZECK, I. C., GRAF, T., SPOERNER, M., TRANKLE, J., FILCHTINSKI, D., HERRMANN,
615 C., GREMER, L., VETTER, I. R., WITTINGHOFER, A., KONIG, B. & KALBITZER, H. R.
616 2010. Stabilizing a Weak Binding State for Effectors in the Human Ras Protein by Cyclen
617 Complexes. *Angewandte Chemie-International Edition,* 49, 3830-3833.
618 SCHREIBER, T. 2000. Measuring information transfer. *Physical Review Letters,* 85, 461-464.
619 SHIMA, F., IJIRI, Y., MURAOKA, S., LIAO, J. L., YE, M., ARAKI, M., MATSUMOTO, K.,
620 YAMAMOTO, N., SUGIMOTO, T., YOSHIKAWA, Y., KUMASAKA, T., YAMAMOTO, M.,
621 TAMURA, A. & KATAOKA, T. 2010. Structural Basis for Conformational Dynamics of GTP-
622 bound Ras Protein. *Journal of Biological Chemistry,* 285, 22696-22705.
623 SINGH, H., LONGO, D. L. & CHABNER, B. A. 2015. Improving Prospects for Targeting RAS. *Journal*
624 *of Clinical Oncology,* 33, 3650-+.
625 SLEBOS, R. J. C., KIBBELAAR, R. E., DALESIO, O., KOOISTRA, A., STAM, J., MEIJER, C. J. L. M.,
626 WAGENAAR, S. S., VANDERSCHUEREN, R. G. J. R. A., VANZANDWIJK, N., MOOI, W.
627 J., BOS, J. L. & RODENHUIS, S. 1990. K-Ras Oncogene Activation as a Prognostic Marker in
628 Adenocarcinoma of the Lung. *New England Journal of Medicine,* 323, 561-565.
629 SPOERNER, M., GRAF, T., KONIG, B. & KALBITZER, H. R. 2005. A novel mechanism for the
630 modulation of the Ras-effector interaction by small molecules. *Biochemical and Biophysical Research*
631 *Communications,* 334, 709-713.
632 SPOERNER, M., HOZSA, C., POETZL, J. A., REISS, K., GANSER, P., GEYER, M. & KALBITZER,
633 H. R. 2010. Conformational States of Human Rat Sarcoma (Ras) Protein Complexed with Its
634 Natural Ligand GTP and Their Role for Effector Interaction and GTP Hydrolysis. *Journal of*
635 *Biological Chemistry,* 285, 39768-39778.
636 STEPHEN, A. G., ESPOSITO, D., BAGNI, R. K. & MCCORMICK, F. 2014. Dragging ras back in the
637 ring. *Cancer Cell,* 25, 272-81.
638 TAVERAS, A. G., REMISZEWSKI, S. W., DOLL, R. J., CESARZ, D., HUANG, E. C.,
639 KIRSCHMEIER, P., PRAMANIK, B. N., SNOW, M. E., WANG, Y. S., DELROSARIO, J. D.,
640 VIBULBHAN, B., BAUER, B. B., BROWN, J. E., CARR, D., CATINO, J., EVANS, C. A.,
641 GIRIJAVALLABHAN, V., HEIMARK, L., JAMES, L., LIBERLES, S., NASH, C., PERKINS,
642 L., SENIOR, M. M., TSARBOPOULOS, A., GANGULY, A. K., AUST, R., BROWN, E.,
643 DELISLE, D., FUHRMAN, S., HENDRICKSON, T., KISSINGER, C., LOVE, R., SISSON,
644 W., VILLAFRANCA, E. & WEBBER, S. E. 1997. Pas oncoprotein inhibitors: The discovery of

potent, ras nucleotide exchange inhibitors and the structural determination of a drug-protein complex. *Bioorganic & Medicinal Chemistry,* 5, 125-133.

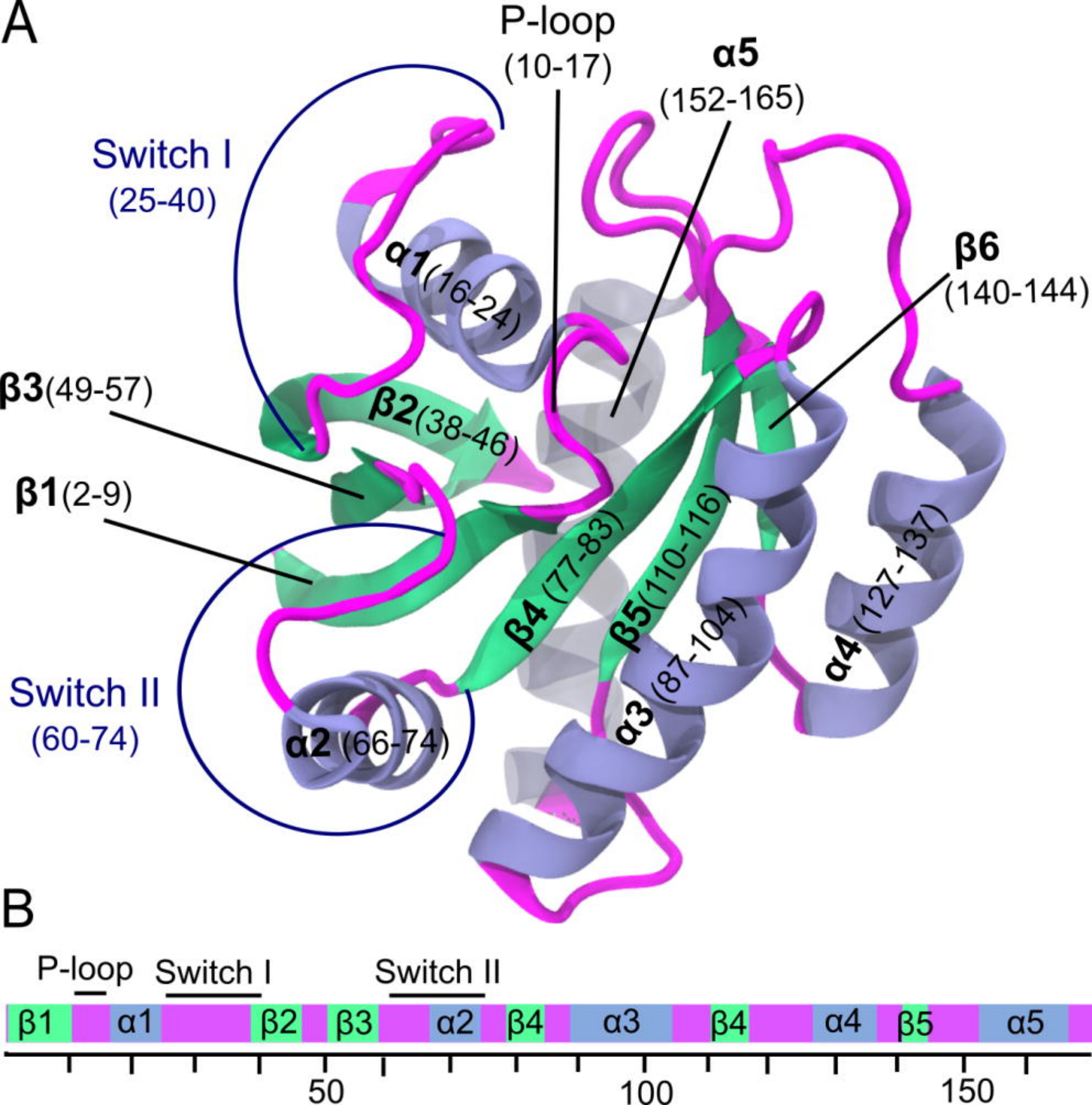
VETTER, I. R. & WITTINGHOFER, A. 2001. Signal transduction - The guanine nucleotide-binding switch in three dimensions. *Science,* 294, 1299-1304.

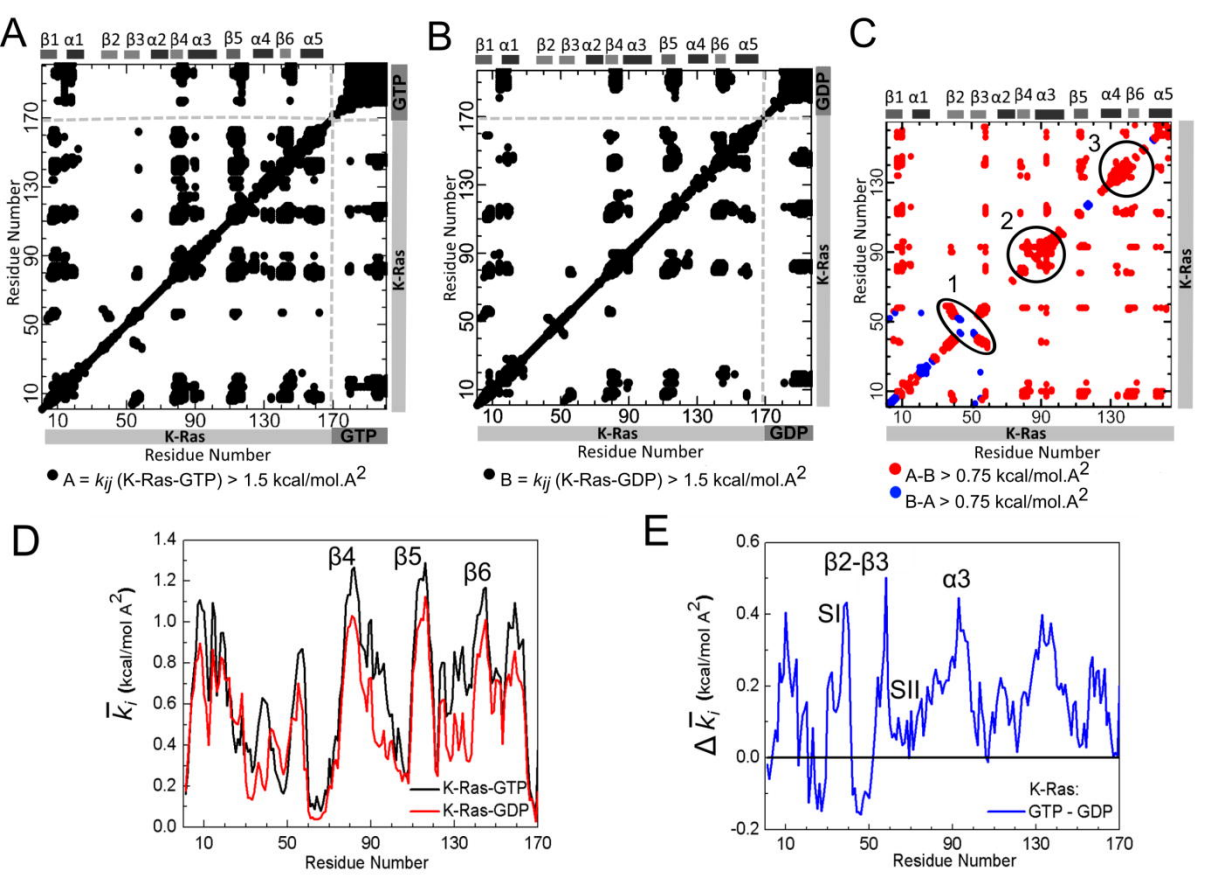
WAND, A. J. 2001. On the dynamic origins of allosteric activation. *Science,* 293, 1395.

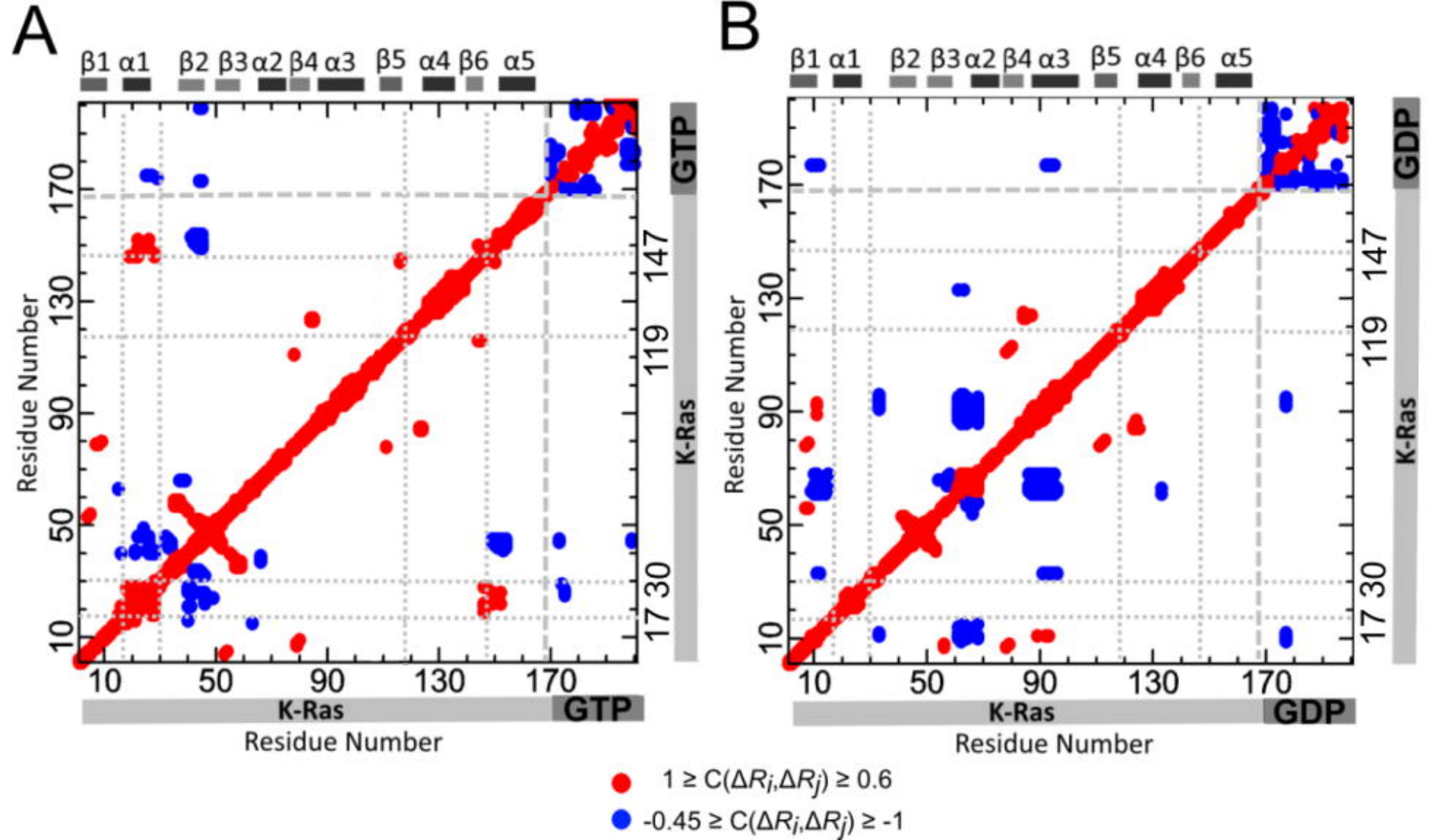
ZACCAI, G. 2000. How soft is a protein? A protein dynamics force constant measured by neutron scattering. *Science,* 288, 1604-7.

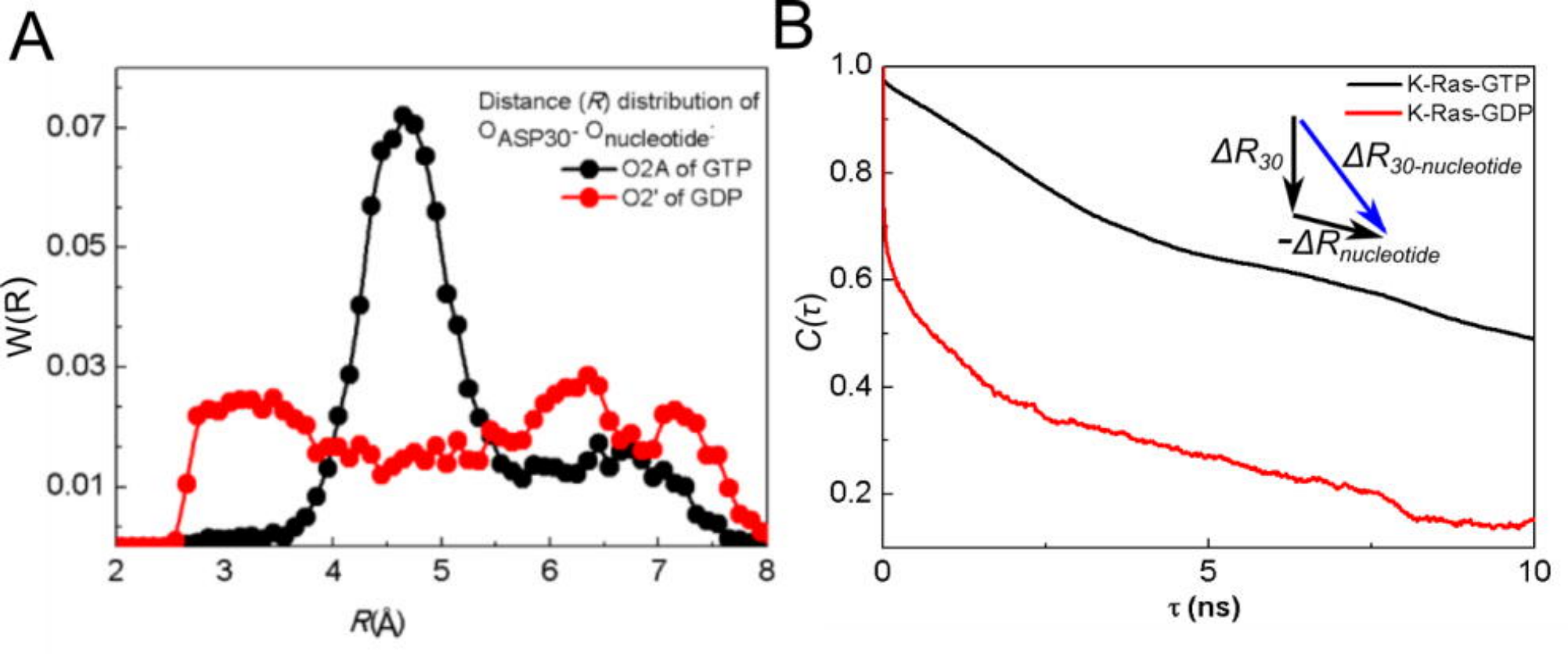
ZHANG, F. & CHEONG, J. K. 2016. The renewed battle against RAS-mutant cancers. *Cell Mol Life Sci,* 73, 1845-58.

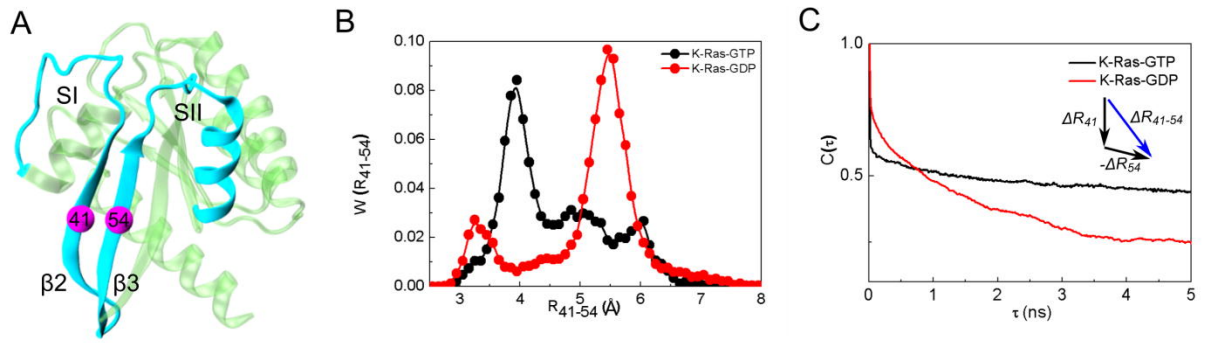
ZHANG, L., BOUGUET-BONNET, S. & BUCK, M. 2012. Combining NMR and Molecular Dynamics Studies for Insights into the Allostery of Small GTPase–Protein Interactions. *Methods in molecular biology (Clifton, N.J.),* 796, 235-259.

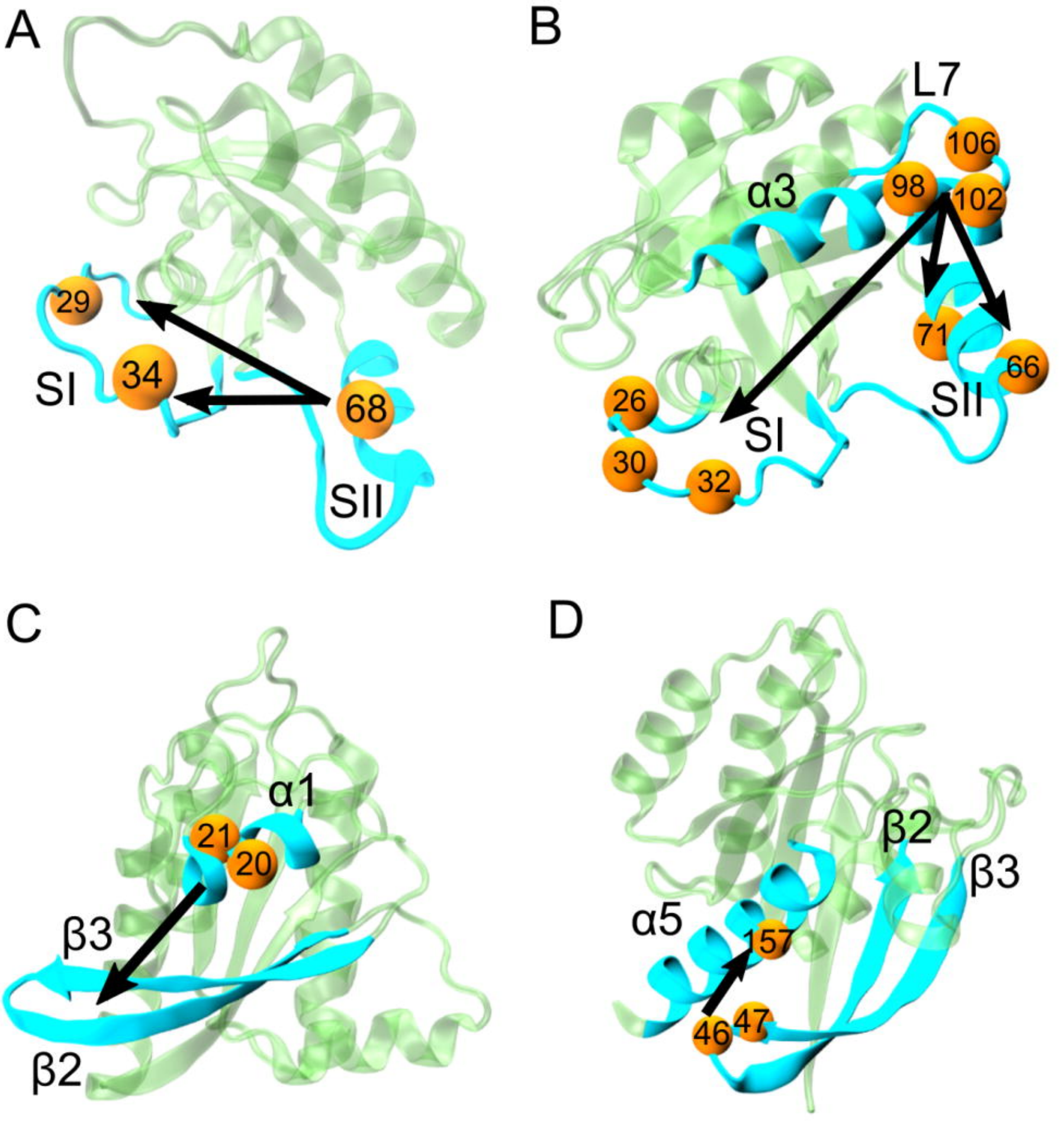

A
P-loop
(10-17)
α5
(152-165)
Switch I
(25-40)
α1(16-24)
β6
(140-144)
β3(49-57)
β2(38-46)
β1(2-9)
β4(77-83)
β5(110-116)
α3(87-104)
α4(127-137)
Switch II
(60-74)
α2(66-74)
B
P-loop
Switch I
Switch II
β1
α1
β2
β3
α2
β4
α3
β4
α4
β5
α5
50
100
150

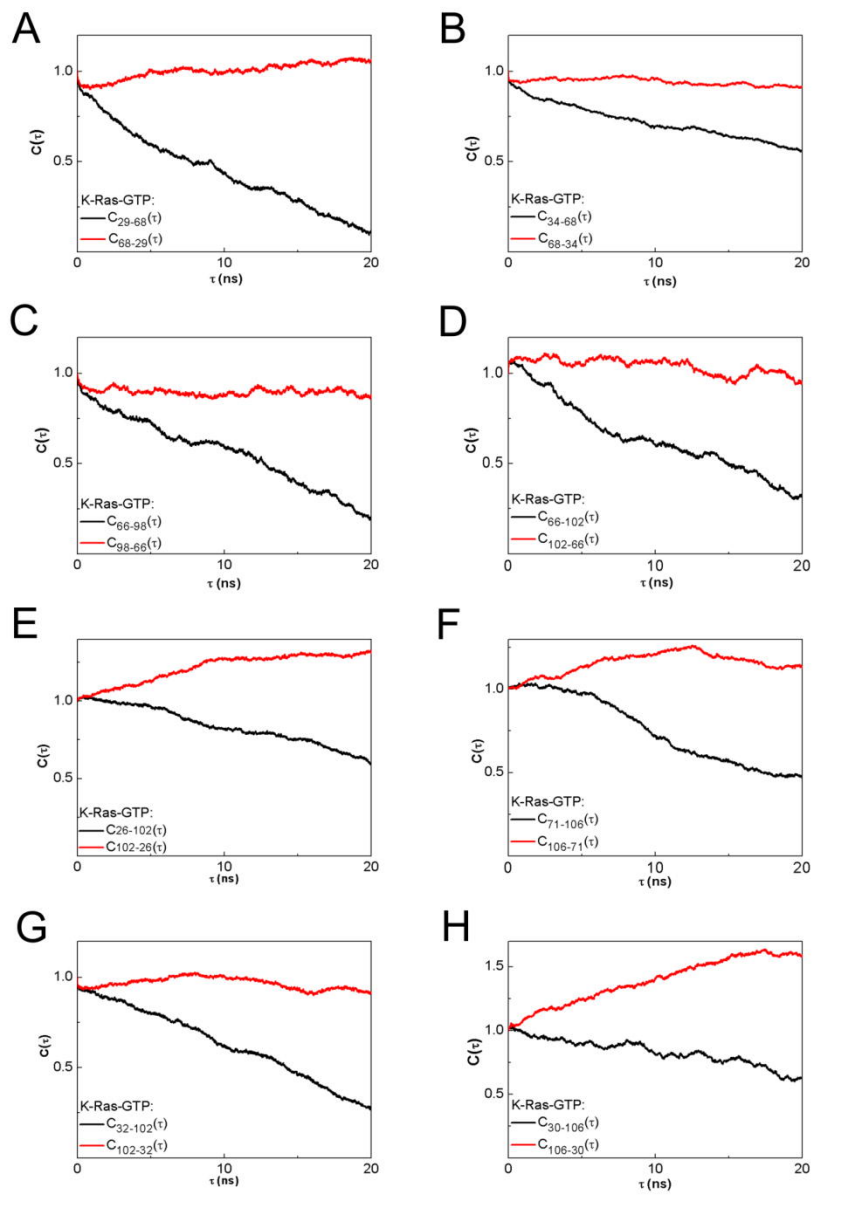

A
β1 α1 β2 β3 α2 β4 α3 β5 α4 β6 α5
GTP
K-Ras
Residue Number
10 50 90 130 170
K-Ras GTP
Residue Number
● A = kij (K-Ras-GTP) > 1.5 kcal/mol.A2
B
β1 α1 β2 β3 α2 β4 α3 β5 α4 β6 α5
GDP
K-Ras
Residue Number
10 50 90 130 170
K-Ras GDP
Residue Number
● B = kij (K-Ras-GDP) > 1.5 kcal/mol.A2
C
β1 α1 β2 β3 α2 β4 α3 β5 α4 β6 α5
1
2
3
K-Ras
Residue Number
10 50 90 130
K-Ras
Residue Number
● A-B > 0.75 kcal/mol.A2
● B-A > 0.75 kcal/mol.A2
D
β4
β5
β6
$\bar{k}_i$ (kcal/mol A2)
K-Ras-GTP
K-Ras-GDP
Residue Number
E
β2-β3
α3
SI
SII
$\Delta \bar{k}_i$ (kcal/mol A2)
K-Ras:
GTP - GDP
Residue Number


A
B
β1 α1 β2 β3 α2 β4 α3 β5 α4 β6 α5
Residue Number
10 50 90 130 170
17 30 119 147
K-Ras
GTP
GDP
1 ≥ C(ΔRi,ΔRj) ≥ 0.6
-0.45 ≥ C(ΔRi,ΔRj) ≥ -1


A
Distance (R) distribution of
O ASP30- O nucleotide
O2A of GTP
O2' of GDP
W(R)
0.07
0.05
0.03
0.01
2
3
4
5
6
7
8
R(Å)
B
K-Ras-GTP
K-Ras-GDP
ΔR30
ΔR30-nucleotide
-ΔRnucleotide
C(τ)
1.0
0.8
0.6
0.4
0.2
0
5
10
τ (ns)


A
SI
SII
41
54
β2
β3
B
K-Ras-GTP
K-Ras-GDP
W (R41-54)
R41-54 (Å)
0.10
0.08
0.06
0.04
0.02
0.00
3
4
5
6
7
8
C
K-Ras-GTP
K-Ras-GDP
ΔR41
ΔR41-54
-ΔR54
C(τ)
τ (ns)
1.0
0.5
0
1
2
3
4
5


A
29
34
SI
68
SII
B
L7
106
98
102
α3
71
66
SII
26
30
32
SI
C
α1
21
20
β3
β2
D
β2
β3
α5
157
46
47


A
B
C
D
E
F
G
H
C(τ)
τ (ns)
K-Ras-GTP:
C29-68(τ)
C68-29(τ)
C34-68(τ)
C68-34(τ)
C66-98(τ)
C98-66(τ)
C66-102(τ)
C102-66(τ)
C26-102(τ)
C102-26(τ)
C71-106(τ)
C106-71(τ)
C32-102(τ)
C102-32(τ)
C30-106(τ)
C106-30(τ)