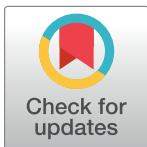


RESEARCH ARTICLE

Comparative MD simulations and advanced analytics based studies on wild-type and hot-spot mutant A59G HRas

Neeru Sharma, Uddhavesh Sonavane*, Rajendra Joshi^{1D}*

HPC–Medical and Bioinformatics Applications Group, Centre for Development of Advanced Computing, Pune, India

* uddhaveshs@cdac.in (US); rajendra@cdac.in (RJ)

Abstract

The Ras family of proteins is known to play an important role in cellular signal transduction. The oncoprotein Ras is also found to be mutated in ~90% of the pancreatic cancers, of which G12V, G13V, A59G and Q61L are the known hot-spot mutants. These ubiquitous proteins fall in the family of G-proteins, and hence switches between active GTP bound and inactive GDP bound states, which is hindered in most of its oncogenic mutant counterparts. Moreover, Ras being a GTPase has an intrinsic property to hydrolyze GTP to GDP, which is obstructed due to mutations and lends the mutants stuck in constitutively active state leading to oncogenic behavior. In this regard, the present study aims to understand the dynamics involved in the hot-spot mutant A59G-Ras using long 10 μ s classical MD simulations (5 μ s for each of the wild-type and mutant systems) and comparing the same with its wild-type counterpart. Advanced analytics using Markov State Model (MSM) based approach has been deployed to comparatively understand the transition path for the wild-type and mutant systems. Roles of crucial residues like Tyr32, Gln61 and Tyr64 have also been established using multivariate PCA analyses. Furthermore, this multivariate PCA analysis also provides crucial features which may be used as reaction coordinates for biased simulations for further studies. The absence of formation of pre-hydrolysis network is also reported for the mutant conformation, using the distance-based analyses (between crucial residues) of the conserved regions. The implications of this study strengthen the hypothesis that the disruption of the pre-hydrolysis network in the mutant A59G ensemble might lead to permanently active oncogenic conformation in the mutant conformers.

OPEN ACCESS

Citation: Sharma N, Sonavane U, Joshi R (2020) Comparative MD simulations and advanced analytics based studies on wild-type and hot-spot mutant A59G HRas. PLoS ONE 15(10): e0234836. <https://doi.org/10.1371/journal.pone.0234836>

Editor: Paul A. Randazzo, National Cancer Institute, UNITED STATES

Received: June 1, 2020

Accepted: October 5, 2020

Published: October 16, 2020

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0234836>

Copyright: © 2020 Sharma et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Introduction

The RAS (RAt Sarcoma) is a protein encoded by a crucial proto-oncogene, which is found to be mutated in ~25% of human cancers. The Ras protein has three isoforms namely: N-Ras, H-Ras and K-Ras where the catalytic domains have 90% sequence identity and the major difference is confined only in the so-called hypervariable region (HVR) at the C-terminal [1]. The distinct HVR of the C-terminal region is responsible for different interactions with the

Funding: The Authors received funding from National Supercomputing Mission (NSM), Govt. of India

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: RAS, RA_t Sarcoma; HRas, Harvey Rat sarcoma; GTP, Guanosine Tri-Phosphate; MD, Molecular Dynamics; SwI, Switch-1; SwII, Switch-2; P-loop, Phosphate-binding loop; GBR, GEF-Binding Region.

membrane and differential accessibility to the activator proteins as well. The differences in the C-terminal region amongst these isoforms lead to distinct post translational modifications which ultimately lead to characteristic biological functions of the different isoforms [2]. In normal conditions, Ras is known to be associated with a range of cellular processes like cell-proliferation, cell-differentiation, apoptosis and plays a regulatory role in cytoplasmic signaling networks [3]. It is a well established fact that Ras functions as a nodal point in signal transduction to control cell-proliferation, migration and neuronal activity [4–6]. The Ras protein also plays a pivotal role in downstream signalling of the associated effector molecules and membrane localization [7]. The best characterized and most studied Ras effectors are Raf proteins, which belong to the family of Serine/Threonine kinases [8]. The Ras protein is known to adopt distinct conformational states, including multiple transient states, where majority of the local changes are notably present in GEF-binding region (GBR), switch 1 (SwI) and switch 2 (SwII) [6, 9–12]. Moreover being a GTPase, the Ras family of proteins possess an intrinsic property to hydrolyze the associated GTP (Guanosine Tri-Phosphate) molecule to GDP (Guanosine Di-Phosphate) molecule [13]. The Ras proteins basically are G-proteins, which inherently function as a binary signalling switch with “ON” and “OFF” states [14, 15]. The GTP bound active conformation of the Ras protein is referred as the “ON” state while the GDP bound inactive conformation is referred as the “OFF” state. Moreover, Ras isoforms have a considerably low energetic barrier between these active “ON” and inactive “OFF” states [16]. The continuously programmed switching between the active and inactive states is assisted by two classes of proteins namely: GEF (Guanine-nucleotide Exchange Factor) and GAP (GTPase Activating Protein) [17, 18]. On one hand, GEF specifically helps in Ras activation by replacing GDP with the GTP, while GAP catalyzes the process of GTP hydrolysis by hydrolyzing the associated GTP molecules [19, 20]. The role of GAP in catalyzing the GTP hydrolysis process by forming a “two-water model” network is also well established [21–23].

The loss of intrinsic GTP hydrolysis activity or becoming insensitive towards external GAP proteins is a prominent feature of the oncogenic gain-of-function mutations [24]. Any single amino acid mutation at the sequence level may lead to subtle structural variation but pronounced functional differences in the mutant structures as compared to the wild-type structures [25]. Gain-of-function Ras mutations are found to be present in almost 25% of human cancers and due to different biological functions served by different Ras isoforms RAS-isoform/effector and RAS-mutation-specific therapeutic approach must be explored to achieve improved therapeutic effects [3]. Efforts are, hence, going on to design isoform and mutant specific Ras inhibitors which could efficiently increase the therapeutic window [26–28]. Also, the mutational variants of the Ras protein are prone to get stuck in the active state as the respective activation energies are increased in mutant conformations [29, 30]. Hence, exploring the dynamics of the mutants and comparing the same with the wild-type is required to address the mutational effects on the protein structure, as a single residue substitution might also have dramatic effects on the protein dynamics. Hence, many studies have focussed on targeting this RAS pathway, understanding the underlying mutants for cancer therapy and designing better therapeutics to treat such disorders [31, 32]. The studies have aimed at understanding the mechanism of Ras signalling and how they are related to human disorders including cancer. Furthermore, correlating the structural, functional and associated dynamics is quintessential for related future investigation including drug repurposing as well drug development based studies [1, 24]. After the availability of the crystal structure of Ras protein, initial simulation studies were focussed only on the individual nucleotide states of the wild-type and mutant conformations [33, 34]. Later on, specific studies on the conserved RAS regions were reported where conformational states of Sw I region were studied in detail, where residues Val29 and Ile36 were specifically mutated to glycines [35]. Research encompassing classical

and advanced MD simulation studies on various hot-spot Ras mutants like G12V, G13V, A59G, Q61L, Q61H have gained acceleration since the last decade.

Earlier, many classical MD, QM/MM simulation studies were carried out on wild-type and mutant Ras systems to understand the associated dynamics. Such studies explore the structural and dynamical aspect of the crucial regions of the Ras protein including switch regions (SwI and SwII) and few loop regions as well [22, 36]. The studies could also explain the distinct conformational states of conserved regions SwI and SwII in GDP and GTP bound states [37, 38]. A study demonstrating the milliseconds time-scale activation phenomenon of the wild-type Ras activation was also reported recently [39]. Role of conserved crystal waters was also reported in this study. Similarly, another study demonstrating the role of conserved water molecules in Q61H K-Ras mutants has also been reported [40]. A restrained and free MD simulation spanning 1.76 μ s simulation length has been reported in the presence and absence of selected crystallographic water molecules. The study also demonstrated that the presence and absence of these waters made the protein to sample distinct conformations and such water molecules act as allosteric ligands to induce a population shift. A similar MD simulation study has reported the dynamic behavior of A59G H-Ras mutants as compared to the wild-type H-Ras [25]. The study also reported increased flexibility in the conserved switch regions along with loop 3, helix 3 and loop 7 regions and a lower energetic barrier between GTP and GDP bound conformations in the A59G H-Ras. Another crucial residue of the Ras protein is Q61, which plays an important role in GAP assisted GTP hydrolysis. Also, the role of Gln61 along with Glu63 and water molecules has also been reported for the GTP hydrolysis process in Ras conformers [41]. Specific hydrogen bond interactions of residues Thr35, Gly60 and Lys16 have also been reported.

Mutations in the Ras proteins are most frequently found at G12, G13 and Q61 positions. In this regard, a comprehensive 6.4 μ s of MD simulation study demonstrated the mechanism underlying the GAP mediated GTP hydrolysis process for the mutants at G12, G13 and Q61 positions [1]. In the reported study, G12V, G12C, G12D, G13D and Q61H mutants of K-Ras were investigated for their respective mechanisms and differential oncogenicity. Differently impaired GAP mediated hydrolysis activity was reported in this study for the mutants and previously distinct intrinsic GTP hydrolysis active was also reported in NMR-based functional profiling based report for Ras mutations [42]. This suggests an overall impairment of the GTP hydrolysis (intrinsic as well as GAP mediated) in the hot-spot missense Ras mutants. Recently, a QM/MM and QM/MM-MD simulation study on the G12V and G13V Ras mutations was also reported [43]. Gln61 residue and conserved crystal waters were found to be displaced in G12V mutants while G13V mutants showed enhanced flexibility in the pre-hydrolysis network formed by Gln61, GTP, water molecules and Arg789-GAP. Both the mutants ultimately encountered a slower hydrolysis reaction and suggest that residue Gln61 is directly involved in the chemical transformations involved in GTP hydrolysis. The pivotal role of Gln61 and its interaction with GAP protein is also established, though QM/MM based *in silico* studies [44, 45]. Role of crucial residues (Thr35, Asp57, Gly60 and Gln61) of the conserved regions from SwI and SwII, especially in maintaining the pre-hydrolysis state has been reported in previous studies [39, 46].

Another hot-spot gain-of-function mutation in the conserved switch region (SwII) is A59G, which plays a crucial role in local conformational changes of the Ras protein [47]. A 20-ns unbiased MD simulation demonstrated a GTP-to-GDP conformational transition by removing γ phosphate of the bound GTP from A59G HRas [25]. The study also depicted that A59G HRas is intrinsically more dynamic than wild-type HRas and mutant HRas conformation has a lower energy barrier between GTP and GDP bound states, as compared to its wild-type counterpart. More recently, a 15 μ s MD simulation study on the wild-type and mutant

A59G HRas was recently reported [48]. The study emphasized on the role of Gln61 residue and the absence of pre-hydrolysis state in mutant A59G HRas systems. The study suggested a decreased rate of intrinsic hydrolysis as well as GAP-mediated hydrolysis by making the movement of SwII region highly restricted. In this regard, the current study presents a detailed insight into the dynamics involved in the mutant A59G HRas conformations. Advanced analysis techniques like MSM based analytics, multivariate PCA analysis and MM-GBSA energy calculations were deployed to draw comparative conclusions for the wild-type and mutant simulation trajectories. In addition, interactions between relevant crucial residues from the P-loop (GBR), SwI and SwII regions were also calculated.

The crystal structure of wild-type HRas with PDB id 1QRA [21] and for mutant A59G HRas with PDB id 1LF0 [47] were available in RCSB Protein Data Bank [49]. The two structures were taken for simulation studies with GNP replaced with GTP in mutant structure and a cumulative $\sim 10 \mu\text{s}$ MD simulations ($5 \mu\text{s}$ each for wild-type and mutant) were carried out. The analysis of interactions of the crucial residues of the pre-hydrolysis network of the wild-type Ras suggested important role of Tyr32, Thr35, Gln61 and Tyr64 residues, hence these features were chosen to perform feature based PCA analysis, along with the energy component.

Materials and methods

System preparation for MD simulations

The crystal structures of the wild-type (GTP-bound) and mutant A59G-HRas (GNP-bound) were downloaded from RCSB-PDB database with PDB ids: 1QRA and 1LF0, respectively. To maintain the uniformity in the respective systems and the simulation protocols, GNP from the mutant A59G-H-Ras structure was replaced with the GTP molecule. The start structure for these classical MD simulation trajectories were captured from the wild-type and mutant A59G ensembles from the well-tempered metadynamics simulation from the previous study [48].

MD simulations

All-atom classical MD simulations were carried out for these GTP-bound systems using GRO-MACS [50] with AMBERFF99SB force-field [51]. A time step of 2 fs was used for production MD runs with short-range van der Waals cut-off kept at 14 Å. The water model chosen was Simple Point Charge (SPC) and coordinates were saved at every 2 ps. FFT optimization and PME algorithms were used to calculate electrostatic terms. An energy minimization of 50,000 steps was performed initially and further, NVT (volume) and NPT (pressure) equilibrations were carried out for 1ns each. Modified Berendsen thermostat for maintaining the simulation temperature and Parrinello-Rahman coupling for the pressure were used. The production runs for wild-type WT-H-Ras-GTP-Mg²⁺ and mutant A59G-H-Ras-GTP-Mg²⁺ complexes were then performed. As GTP being a non-standard molecule and its parameters not available in gromacs, its parameters were calculated from the Prodrgr server [52]. For the present study, $\sim 10 \mu\text{s}$ of classical MD simulations were performed for the wild-type and mutant A59G H-Ras systems namely *simulationW* and *simulationM*, respectively. The MD simulations were performed using in-house PARAM-BRAF supercomputing cluster at C-DAC.

Visualization and analysis

For structure visualization, molecular graphics and analyses, tools like UCSF Chimera [53] and VMD (Visual Molecular Dynamics) [54] were used. Xmgrace [55] was used to plot the two-dimensional plots of the relevant analysis carried out. Detailed analyses were carried out for these simulations, namely interactions of the residues belonging to conserved switch

regions, multivariate PCA (using selected features) using RStudio package, MMGBSA based energy calculations, Markov State Model (MSM) based analysis of molecular kinetics and thermodynamic models from MD data using PyEMMA [56].

MSM based analyses

With respect to MSM, the creation of states and the transition matrix depends on the choice of collective variable (CVs) or features and the lag time [57–60]. In order to build the MSM model for the MD simulation data in this study, following steps were carried out: (a) VAMP Score Calculation: This was done for different features, which is based on the variational principle [61]. The variational principle helps to check the robustness of MSM. VAMP scores were calculated for following features viz. dihedral angles, residue minimum distance between 2 groups of residues (mindist) and the distances between groups with respect to their C-alpha atoms (Cadist). (b) Feature Selection (CV): The Cadist feature was then used as the feature for further MSM calculations. (c) Time Independent Component Analysis (TICA): Using the selected feature i.e. Cadist, TICA calculation was performed in the second step. TICA helps to reduce the dimensionality of the trajectory data. (d) Clustering: The MD simulation data for all the trajectories was then clustered using K-means clustering algorithm for both the wild-type and mutant systems. (e) Macrostates Creation: In the next step, Principal Components & Classification Analysis (PCCA) calculation was performed and macrostates were created for the respective MD simulation clusters (f) Transition Matrix: Finally, the transition matrix was then built using the chosen lag time in such a way that the transition matrix follows markovian behaviour. The eigen decomposition of the transition matrix results in eigenvectors which represents slow motions of the systems. The fluxes and mean passage time calculations were performed by computing the transition pathways between the states using the discrete transition path theory [62]. The choice of the lag time is related to eigenvalue which gives us a physically measurable timescale known as implied timescale (t_i) using Eq 1:

$$t_i = -\tau / (\ln \lambda_i) \quad (\text{Eq1})$$

where t_i is implied timescale, τ is the lag time and λ_i is an eigenvalue.

The validation of the MSM was done using Chapman Kolmogorov equation [57] given by the Eq 2:

$$T(n\tau) = T(\tau)^n \quad (\text{Eq2})$$

where n is an integer representing the number of steps, $T(\tau)$ is a transition matrix at lag time τ .

The above relationships help in validating the chosen MSM model, which ultimately plays a crucial role in determining the transient intermediate states as well the involved kinetics amongst them. All these steps of MSM model preparation and analysis were carried out for each of the systems under study.

Results

Basic analyses: Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuation (RMSF)

RMSD. Basic RMSD analysis was carried out for *simulationW* and *simulationM* of the wild-type and mutant simulation trajectories, respectively. The RMSD calculation was done for the protein backbone and respective conserved regions (P-loop, SwI and SwII) with respect to the respective first frame. RMSD plots for wild-type Ras showed more deviation in case of wild-type trajectories as compared to the mutant A59G system, as shown in Fig 1. Fig 1

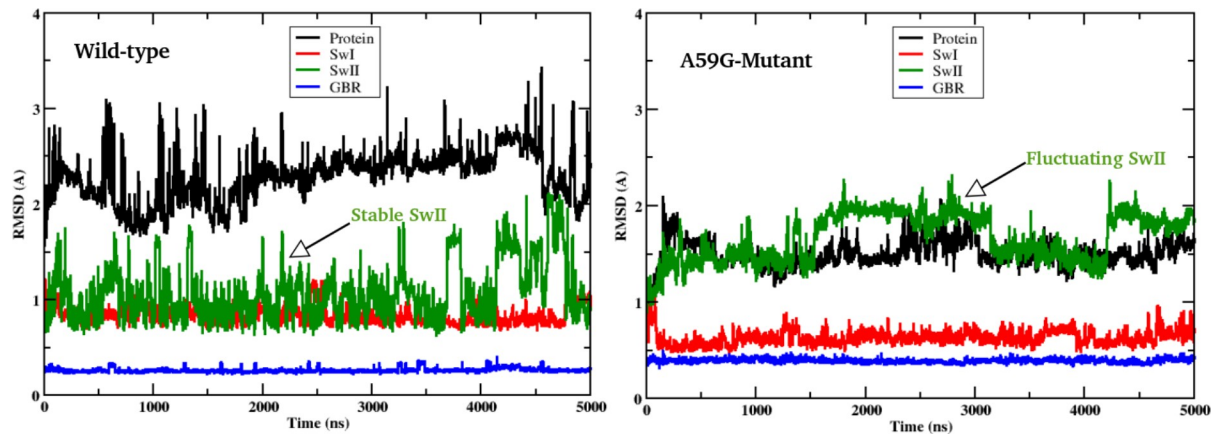


Fig 1. 2D-plot showing RMSD trend for the wild-type and mutant *simulationW* and *simulationM*, respectively. The mutant system showed a fluctuating SwII region as compared to its wild-type counterpart.

<https://doi.org/10.1371/journal.pone.0234836.g001>

showed a deviation for protein backbone of 2.5–3 Å for wild-type and 1.5–2 Å for the mutant, respectively. The GBR and SwI region did not show major deviation, with respective RMSD values of 0.25 Å and 0.5 Å for GBR and 0.75 Å and 0.5 Å for SwI (for wild-type and mutant systems, respectively). Fig 1 also showed an increased deviation in case of mutant simulation for SwII region (especially the loop region, L4) where the RMSD value laid between 1.5–1.75 Å, as compared to 0.75 to 1 Å for the wild-type simulation. The RMSD suggests a notable change in SwII region as compared to the wild-type counterpart, which is also reflected in the complete protein backbone's RMSD values.

RMSF. RMSF was also calculated for the two systems. Fig 2 shows the RMSF plot *simulationW* and *simulationM*. For the residues flanking the SwI region (residues 25–29 and residues

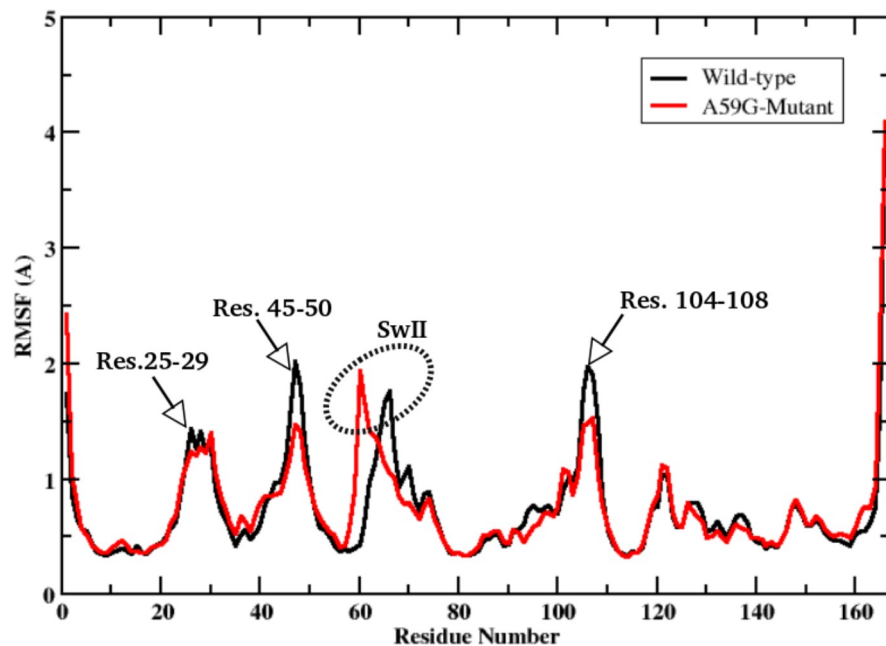


Fig 2. 2D-plot showing RMSF trend for the wild-type and mutant *simulationW* (black line) and *simulationM* (red line), respectively. The regions showing changes are labelled accordingly.

<https://doi.org/10.1371/journal.pone.0234836.g002>

45–50), a difference of 0.5 Å was observed for the wild-type and mutant trajectories. Also, the RMSF trend of the SwII region, especially the loop region (residues 59 to 65) showed distinct trends for the wild-type and mutant simulations. Interestingly, the region between residues 45–50 and 104–108 also showed a decreased RMSF variation (of ~0.5 Å) in mutant simulation as compared to its wild-type counterpart. As these deviations are small in quantity, structural analysis of the respective trajectories was also performed to understand these changes in more detail. As depicted in Fig 3, the highlighted labels show the localised variations for above three mentioned regions. The relative thicknesses of the respective regions show the distinct variation throughout the 5 μ s of the respective simulation trajectories. The above mentioned residue stretches (25–29, 45–50, 59–65 and 104–108) showed clear distinction for the wild-type and mutant systems, as depicted by the corresponding ribbon thickness.

Analyses for conserved GBR, SwI and SwII regions

Interactions of crucial residues of the conserved regions. The basic RMSD and RMSF analyses suggested distinct differences in GBR, SwI and SwII regions. Hence, crucial residues from these regions were then checked for their interactions with other residues as follows: Lys16, Ser17 (GBR), Tyr32, Thr35 (SwI), Asp57, Gly60, Gln61 and Tyr64 (SwII). Figs 4–6 show the plots for these interactions as the number of conformers (for 5 μ s of the respective trajectories) versus distance between the respective pairs. Residue Gln61 showed distinct interactions with Lys16, Ser17, Thr35 and Tyr64 for the two trajectories. For this, the simulation length is represented as the number of conformers and the respective distance bin for each residue pair.

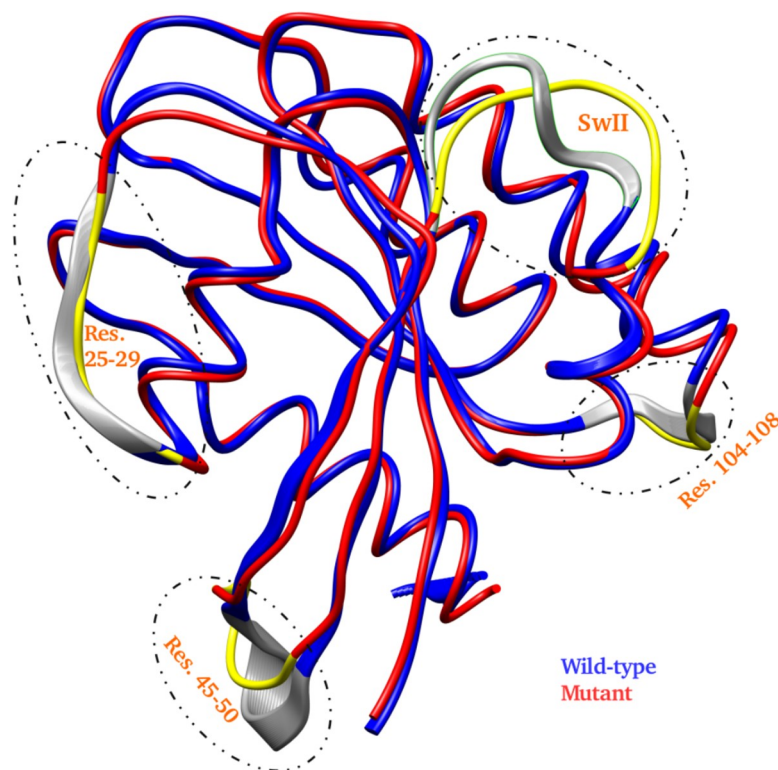


Fig 3. Overlaps of the wild-type and mutant simulation trajectories to depict the deviation in the selected regions in dotted ovals (gray for wild-type and yellow for A59G mutant). The wild-type and mutant trajectories are shown in blue and red cartoon representations, respectively.

<https://doi.org/10.1371/journal.pone.0234836.g003>

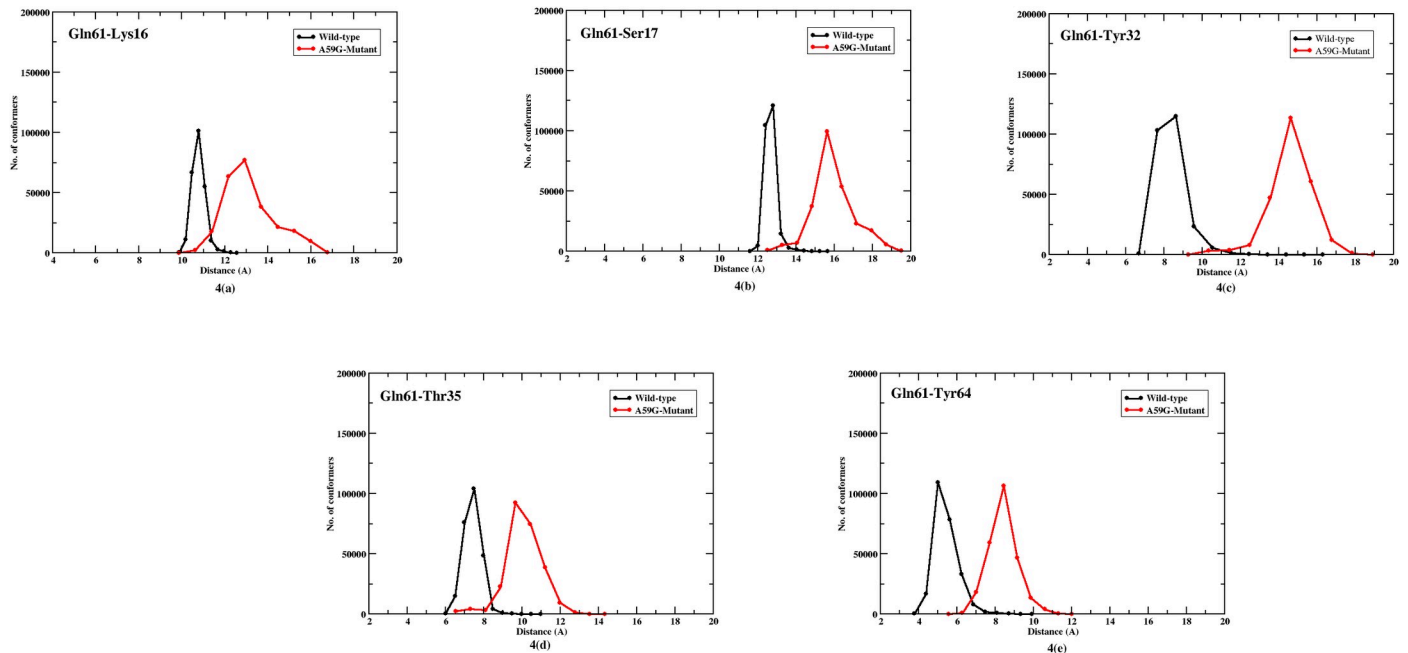


Fig 4. 2-D plot showing number of conformers for Gln61 residue's interaction with Lys16, Ser17, Thr35, Tyr32 and Tyr64. Black and red lines represent trends for the wild-type and mutant systems, respectively.

<https://doi.org/10.1371/journal.pone.0234836.g004>

Firstly in the wild-type *simulation W*, for the Gln61-Lys16 pair, the interaction laid in range 10–12 Å with maximum conformation concentrated at 11 Å and mutant *simulation M* had this interaction in range of 12–17 Å with an average 13 Å, as shown in Fig 4(A). Similarly for the Gln61-Ser17 pair, Fig 4(B) showed that the *simulation W* showed maximum conformations in 12–13.5 Å and the mutant *simulation M* had a spread out conformations in 14–19 Å range. The third interaction for the wild-type simulation showed a distance of 6–8 Å and mutant simulation at 8–13 Å for the Gln61-Thr35 residue pair, as depicted in Fig 4(C). The residues Tyr32 and Tyr64 showed distinct interaction trends with Gln61. Fig 4(D) demonstrated that the pair namely: Gln61-Tyr32, showed significant difference in the distance range for the wild-type and mutant systems. For the wild-type simulation, the interaction fell in the 7–11 Å range and for mutant simulation the residues were more distant at a distance of 12–18 Å. The most crucial difference of Gln61 interaction was found with Tyr64 (as shown in Fig 4(E)), where for the wild-type system the residues were facing each other (as visualized in the structural

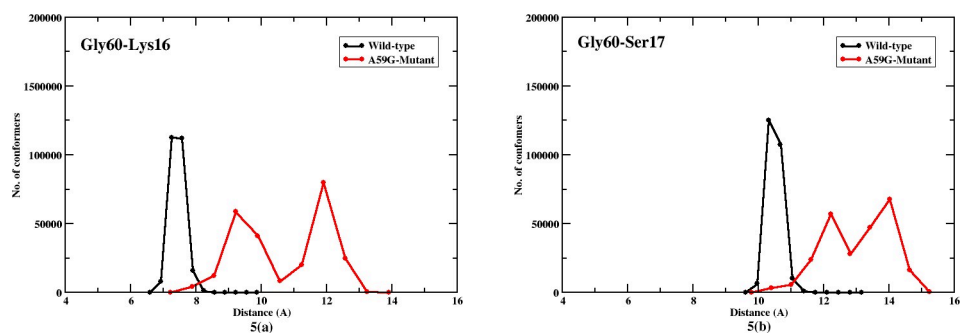


Fig 5. 2-D plot showing number of conformers for Gly60 residue's interaction with Lys16, and Ser17. Black and red lines represent trends for the wild-type and mutant systems, respectively.

<https://doi.org/10.1371/journal.pone.0234836.g005>

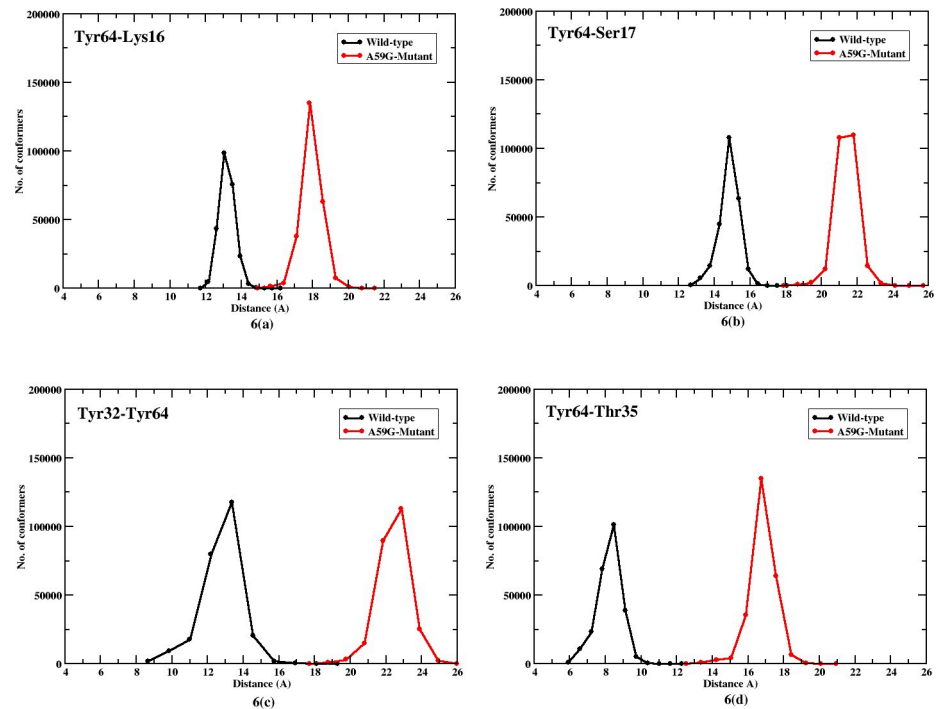


Fig 6. 2-D plot showing number of conformers for Tyr64 residue's interaction with Lys16, Ser17, Thr35 and Tyr32. Black and red lines represent trends for the wild-type and mutant systems, respectively.

<https://doi.org/10.1371/journal.pone.0234836.g006>

analysis) within range 4–7 Å, while for the mutant system the residues were an increased distance of 6–11 Å, due to a change in Gln61 orientation.

Further, residue Gly60 was also analyzed for its interaction with other residues of the conserved regions. As residue Gly60 directly plays a crucial role in GTP hydrolysis for Ras systems, its interaction was analyzed for the residues belonging to Lys16 and Ser17 (the conserved residues of GBR). Majorly Lys16 and Ser17 showed distinct interaction in wild-type and mutant simulations. Fig 5(A) and 5(B) showed plots for Gly60-Lys16 and Gly60-Ser17 respectively. For the wild-type simulation, Gly60-Lys16 mostly fall in 6.5–8 Å and 8–13 Å for the mutant simulation. A similar trend was observed for Gly60-Ser17 interaction as well, where the wild-type simulation had interaction in range 10–11 Å and between 11–15 Å for the mutant counterpart. The residue Gly60 showed fluctuating trends with its interaction with GTP, due to absence of conserved water molecules needed for maintaining pre-hydrolysis state (plots not shown). Interestingly, two maxima were observed for these interactions in the mutant A59G simulations owing to the reason that Gly60 showed increased fluctuations in mutant simulationM.

Similarly, residue Tyr64 was also checked for its interaction with Lys16, Ser17, Tyr32 and Thr35, of which Lys16 and Ser17 (GBR) showed a similar trend while Tyr32 and Thr35 (SwI) share a typical trend of their respective interactions with Tyr64. Fig 6(A) and 6(B) show Tyr64-Lys16 and Tyr64-Ser17 interactions. The wild-type simulation showed a range of 12–14 Å and 13–17 Å and mutant simulations with 16–20 Å and 12–19 Å, for Tyr64-Lys16 and Tyr64-Ser17 interactions respectively. Interestingly, interactions with Tyr32 and Thr35 segregated the wild-type and mutant conformation in two separated clusters. Fig 6(C) depicts that the Tyr64-Tyr32 interaction in wild-type was in the range 9–15 Å and 20–25 Å for the mutant system. Similarly for the wild-type system, the Tyr64-Thr35 interaction was in range 6–10 Å

and for mutant it was 12–19 Å, as shown in Fig 6(D). The comparative trends of these crucial interactions suggest a significant effect on the orientation and distinct behavior of the SwII residues, especially Gly60, Gln61 and Tyr64.

Dihedral analyses of Tyr64 residue. As interaction of Tyr64 with residues of SwI region (Tyr32 and Thr35) showed distinguishable characteristics for wild-type and mutant trajectories, a visual inspection of the Tyr64 side chain was carried out for both the trajectories. The orientation of Tyr64 (along with neighboring Gln61 residue) differed in the wild-type and mutant trajectories. To quantify this change, dihedral angle analysis was carried out for Tyr64 residue as shown in Fig 7. For Φ dihedral, both the systems laid in -30° to -150° . On the other hand, the ψ dihedral showed clear distinction between wild-type and mutant simulations. For the wild-type system, the ψ angle laid in -60° to 60° while for the mutant counterpart, the ψ dihedral had a different orientation between 60° to 120° .

On the basis of above analyses, it is established that the residues Gln61 and Tyr64 majorly differentiates between the wild-type and A59G mutant ensembles.

Multivariate / Feature-based principal component analysis. Collating these residues and the associated features like distances, dihedral and energy component (S1 File), a multivariate PCA analysis was performed. The preliminary analyses helped in searching the available search space for all the potential features and then finalizing the most informative ones for the multivariate PCA. The major rationale behind performing PCA is to determine low-dimensional variates, in this case amongst many interactions, angles, dihedrals etc like features/factors. For this, scaling of all the input features was performed first which might have different variances for distinct individual features. After this standardization of the variables involved, the selected feature-based PCA was performed using the R statistical package. The features (variables) for this PCA analysis were chosen on the basis of previous analyses for the two

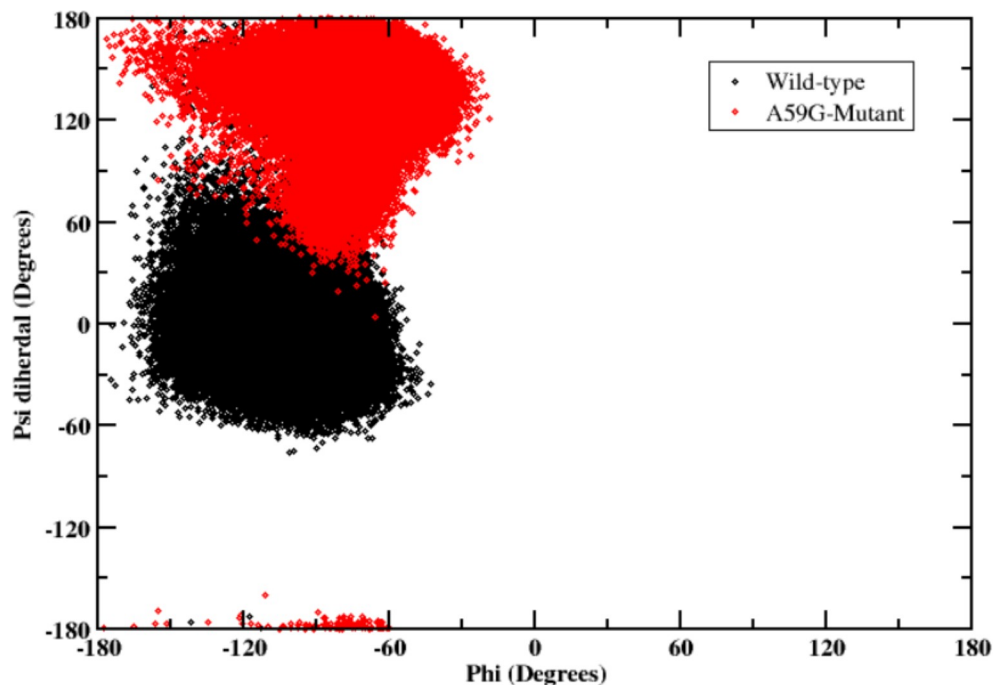


Fig 7. 2-D plot showing the segregation of wild-type and mutant simulation trajectories, on the basis of $\phi(\Phi)$ and $\psi(\Psi)$ dihedrals values. Black color depicts the wild-type and red color represents the A59G-mutant cluster.

<https://doi.org/10.1371/journal.pone.0234836.g007>

trajectories as listed: Φ and ψ dihedrals of Gln61 and Tyr64 residues, crucial interactions between Tyr64-Tyr32, Tyr64-Thr35 and Tyr64-Gln61 and the total energy component as calculated from the MMGBSA energy analysis (S1 File). Fig 8 shows the projections of these features in the two-dimensional matrix, where the wild-type system is represented in green clusters and the mutant system in red clusters. The plot shows that the ψ dihedral of Gln61 and Tyr64 along with the distances and energy component were able to distinguish the wild-type and mutant ensembles. Fig 9 shows the projections of PC1 on PC2 for the wild-type (cyan) and A59G mutant (orange) trajectories based on the above mentioned multiple features and ~90% of the population is covered along these two PCs. The wild-type and mutant clusters showed distinct clusters especially along the PC1, as shown in Fig 9. The plot also shows that dihedrals of the residue Gln61 are negatively correlated with all the other features chosen for this multivariate PCA. Hence, these features can also prove to be an important reaction coordinate/collective variable (CV) for performing advanced biased metadynamics simulations.

MSM based analyses of MD simulation trajectories

After performing the detailed analyses of the crucial residues of the conserved regions and the multivariate PCA, detailed insight is gained for the two systems. To further understand the origin of these distinct behaviors and to comprehensively analyze the two trajectories, Markov State Model (MSM) based analysis of the simulation trajectories was performed using PyEMMA. In principle, MSM models are the simplified coarse grained kinetic models which are constructed from the available simulation data. The MSM based analysis provided us with the distinct metastable states for the wild-type and mutant systems. The feature chosen in this MSM based kinetic models generation was “Cadist”, which is the distance between

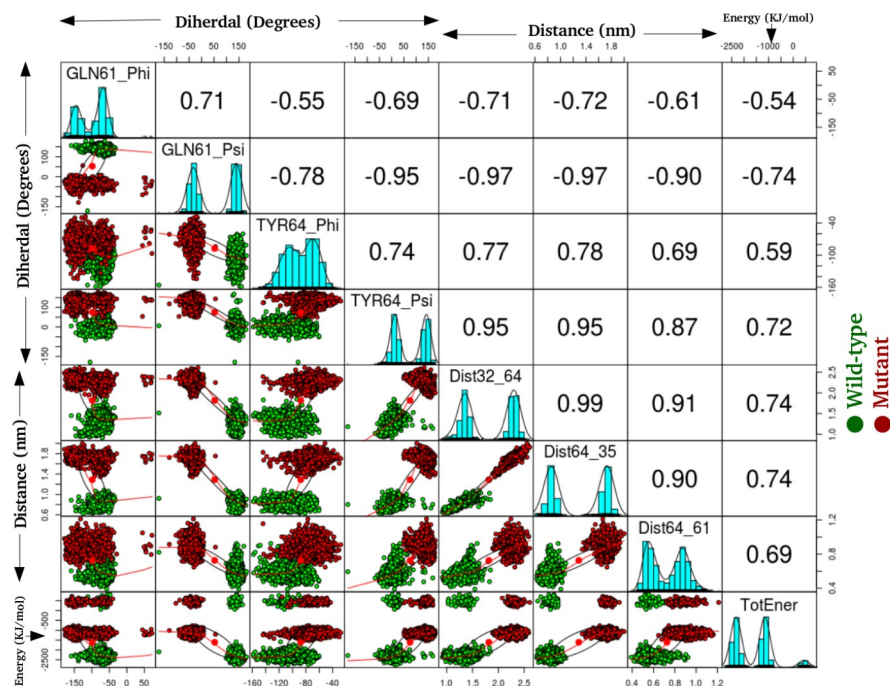


Fig 8. The 2-D plot showing the projections of wild-type (green) and mutant (red) trajectories, using multivariate PCA approach. A combination of dihedral, interactions and energy were used as features, as labelled on the X and Y axis.

<https://doi.org/10.1371/journal.pone.0234836.g008>

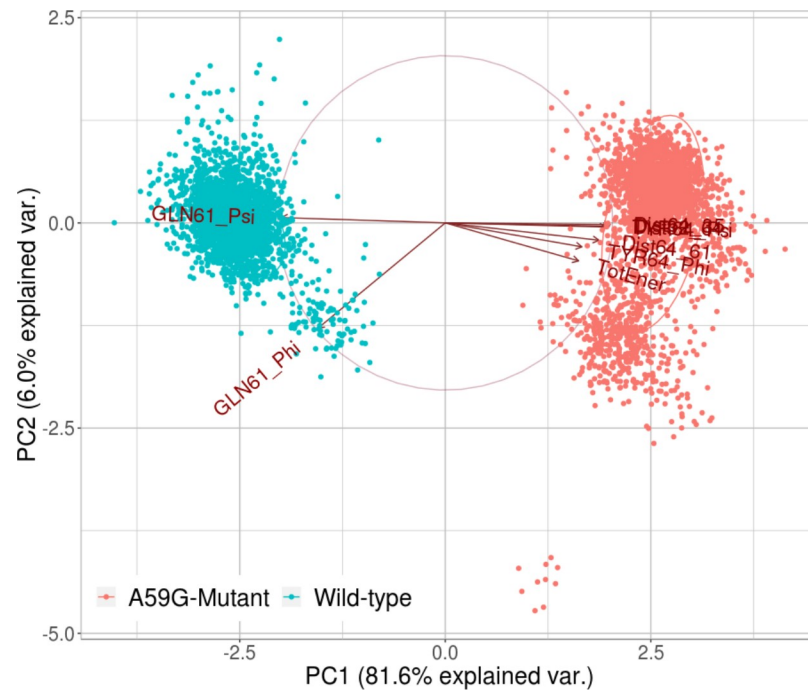
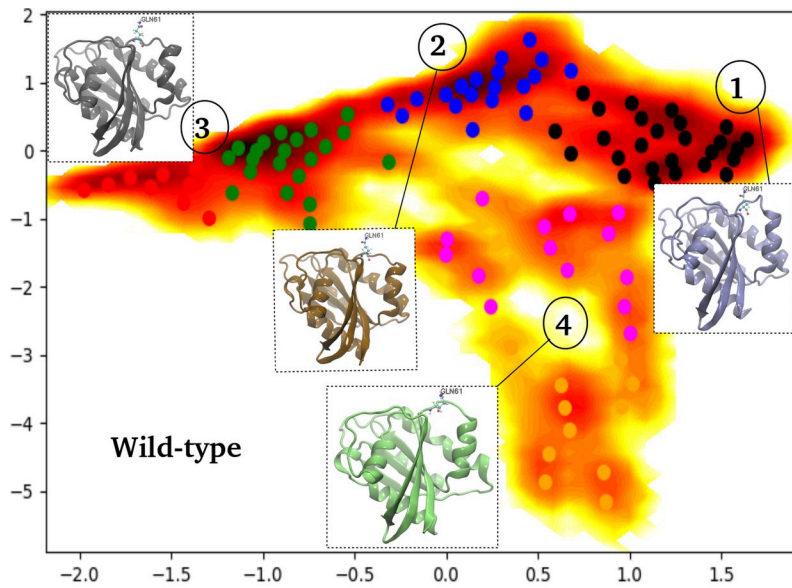


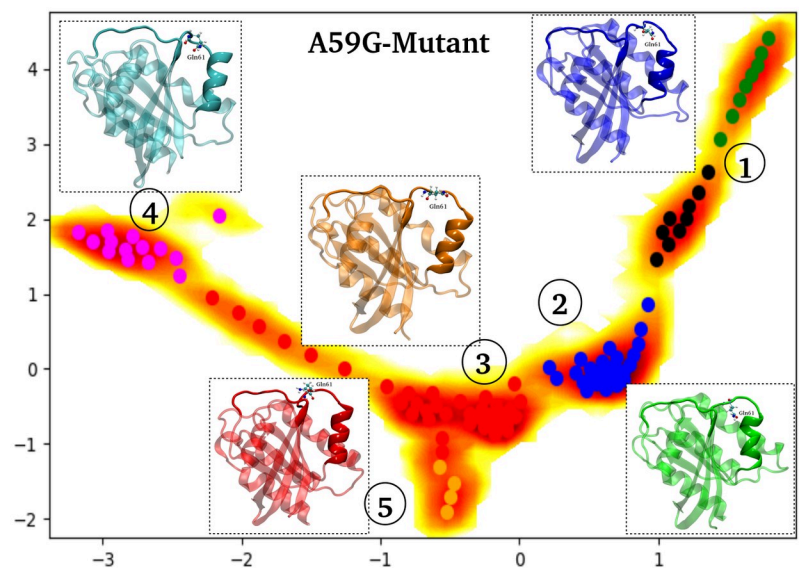
Fig 9. Projections of PC1 on PC2, of the features chosen for multivariate PCA. Cyan and orange color clusters are for wild-type and mutant trajectories, respectively. The individual arrows represent the eight features used for this analysis.

<https://doi.org/10.1371/journal.pone.0234836.g009>

Ca-coordinates of the Ras protein. A lag time of 125 ns was chosen for both simulation trajectories. Fig 10(A) and 10(B) shows the schematic representation of the distinct states ensembles for the wild-type and mutant simulation trajectories using the MSM analysis. The MSM analysis of the wild-type trajectory identified four metastable states, while for the mutant trajectory five metastable states were identified, as shown in Fig 10. Fig 11(A) and 11(B) show the pathways from the rightmost to the leftmost state, for the wild-type and mutant trajectories respectively. The leftmost state is assumed to be a rare-event, while the right one is the frequently visited. The thick arrows represent the most probable and stable pathway and other splits and thin arrows demonstrate other intermediate pathways encountered for the given system. For the wild-type simulation, states 1, 2 and 3 are mostly visited, while state 4 is sparsely visited. Upon structural analysis of these states, the wild-type system was found to have similar structural characteristics. On the other hand, the mutant structure showed five metastable states, where the metastable state 5 had altogether different characteristics than the rest of the trajectories. Metastable states 1–4 had similar structural features. Of these five states, state 5 bears most importance as it shows structural features similar to that of the wild-type counterpart. Moreover, when the transition pathways were computed for these MSM formulations, the transition time between cluster 1 (highly populated) and sparsely populated cluster 5 for the mutant system showed significant difference. It suggests that it approximately takes 3.36 millisecond to reach to the transient state in cluster 5 (wild-type like pre-hydrolysis state conformation) for the mutant, while the reverse would take 2.25 milliseconds to reach back to cluster 1. The transition values of the 5 μ s of the wild ensemble are not discussed here, as other analyses show that the system did not differ in terms of active/inactive states, where most of them possess similar structural characteristics.



10(a)



10(b)

Fig 10. Depiction of metastable clusters for the wild-type and mutant trajectories, represented by numbers, as captured using PyEMMA while performing MSM based analysis. The feature used was Cadist (distances between the C-alpha atoms of the respective trajectories).

<https://doi.org/10.1371/journal.pone.0234836.g010>

Furthermore, Fig 12 shows an overlap of the pre-hydrolysis network region of a mutant structure from the metastable state cluster 5 (yellow CPK representation) and the most stable conformations, each from wild-type simulation and mutant trajectories (cluster 1). This pre-hydrolysis state is difficult to encounter in the mutant simulations and in this 5 μ s mutant simulation, the structure was observed for a short time as demonstrated by the sparsely populated cluster 5.

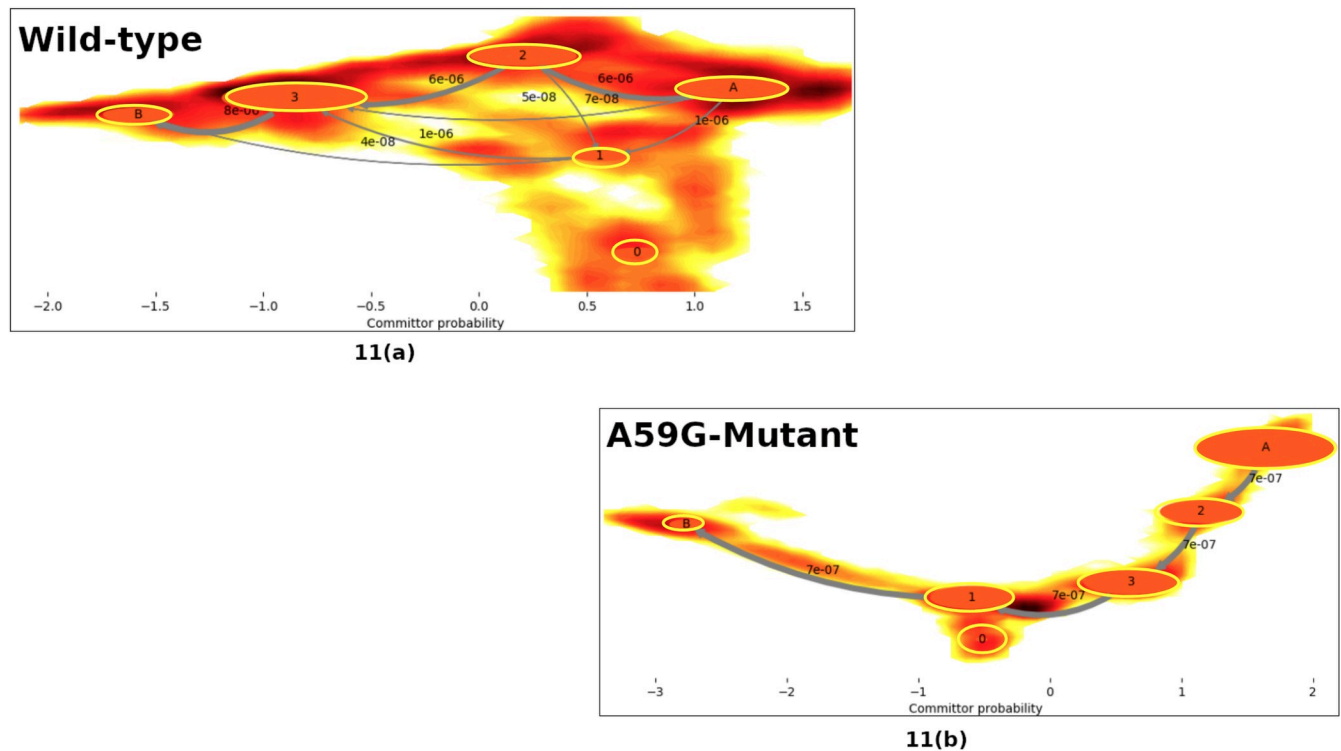


Fig 11. The 2D-plot demonstrating the probable pathways for the wild-type and mutant simulations, calculated using PyEMMA. The most probable pathway is denoted by a bold gray arrow (rightmost to left most cluster). The committor probability of each state transition is written along the respective arrow.

<https://doi.org/10.1371/journal.pone.0234836.g011>

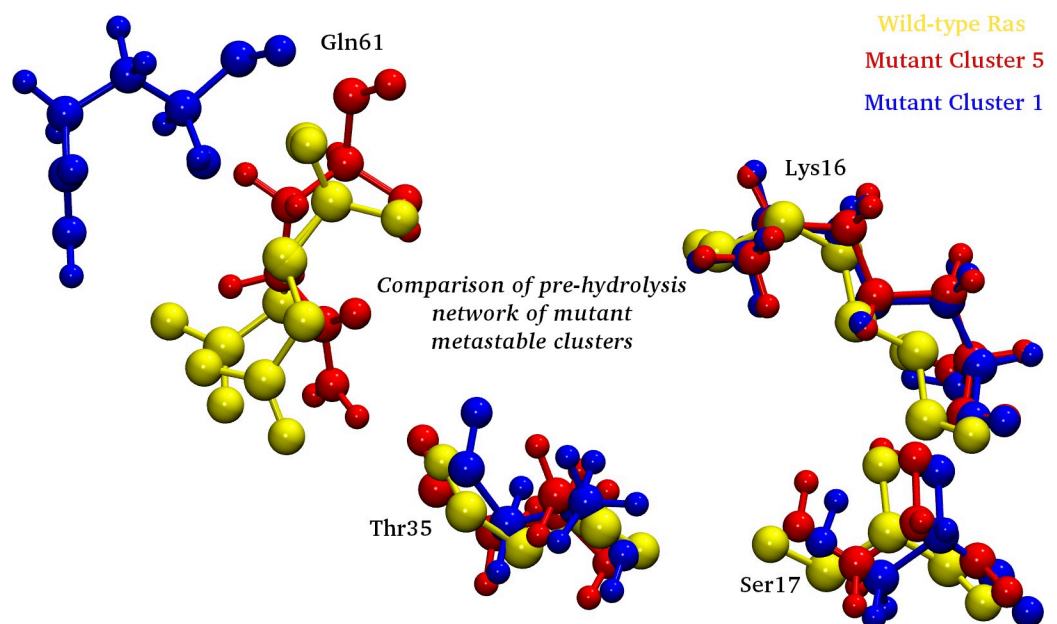


Fig 12. Fig showing the overlap of representative structures of the cluster 5 (mutant), cluster 1 (wild-type and mutant), from the respective MSM based analysis. Yellow CPK representation shows the mutant conformation from cluster 5 (sparsely populated) and the most probable conformations of wild-type and mutant trajectories, from the respective cluster 1.

<https://doi.org/10.1371/journal.pone.0234836.g012>

Other additional relevant plots for the MSM analysis like VAMP score, lag-time calculation, implied time-scales and cktest are provided in the [S1 File](#).

Discussion

In the present study, the wild-type and mutant A59G-HRas were simulated using classical MD simulations for microseconds level time-scale, for each of the two systems. Through this study, an attempt was made to assess and characterize the distinct features of the wild-type and mutant A59G ensembles along with the comparative analyses of the pre-hydrolysis state. The absence of pre-hydrolysis state in the A59G-HRas mutant plays a pivotal role in the aberrant hydrolysis activity [48]. In this backdrop, 5 μ s of classical MD simulations were carried out and in-depth analyses using advanced analytics techniques like feature-based PCA and MSM were carried out. The preliminary RMSD and RMSF based analyses showed distinct trends of the conserved regions, especially SwII for the wild-type and mutant trajectories. The structural analyses showed remarkably different orientation of the SwII region, where the mutation (A59G) lies, which also correlates with the RMSF results. The impact of the hot-spot mutation is clearly evident on otherwise conserved regions as well (especially SwII) [63]. The comparison also demonstrated that the regions encompassing residues 45–50 and 104–108 had considerable variance for the wild-type system, which also correlates with the simulation study on the solution structure of the H-ras p21-GTP complex [33]. The study also suggested that the loss of important intermolecular contacts upon solvation at or near the residues, might be responsible for this enhanced variation. Following this, an extended analysis was also carried out on the residues of the SwI, SwII and GBR regions. The analyses showed important residues like Tyr32, Thr35 (SwI), Gly60, Gln61 and Tyr64 (SwII) assist in maintaining their respective interactions and hence are capable while selecting features for biased simulations like umbrella sampling, steered MD, metadynamics etc. Moreover, for advanced analytics approaches like feature-based PCA (where multiple variates are required) and for MSM based techniques where CVs are required, features based on these important residues can be defined. A very recent study has also reported the importance of reaction coordinates specifically for differentiating the wild-type from G12V, T35S and Q61K mutant HRas ensembles using long 1.4 μ s accelerated MD simulations [64]. The study also reveals the effect of underlying mutations on the dynamic conformational changes of the conserved regions (SwI and SwII) in the mutant HRas systems. Additionally upon visual analysis, the residue Tyr64 showed quite interesting behavior in terms of side chain fluctuation and to quantify this behavior dihedral angle analysis was carried out for Tyr64 residue. A flip along the ψ dihedral of the Tyr64 residue was observed, where a complete flip from the negative (-60° to 60°) to positive quadrant (60° to 120°) was observed for wild-type and mutant ensembles, respectively. The above values for the simulation trajectories correlated well with the experimental crystal structures available at RCSB, where ψ dihedral was -24° (for wild-type pdb id: 1QRA [21]) and 143° (for A59G mutant pdb id: 1LF0 [47]). This also strengthens the hypothesis about the strong impact of A59G mutation on the SwII region. Another recent study also talks about the individual yet crucial impact of specific residues on the structural arrangement for the KRas isoform [65]. The study talks about wild-type, G12C and G12D mutants of KRas and roles of specific residues like Tyr32, Gln61 and Tyr64 in destabilization of water molecules in mutant systems. Of these, Tyr32 (SwI) came out to be a central contributor in the dynamics. The quantifiable flip of the Tyr64 residue's side-chain is a crucial characteristic in A59G-HRas mutants, which also participates in making direct contacts with other important residues like Tyr32, Tyr64 and Gln61. A similar trend of SwII change and flip of Gln61 side-chain is also reported previously [25, 47, 48]. Hence, a highly pronounced effect of the A59G mutation (belonging to SwII) on the crucial

residues like Gly60, Gln61, Tyr64 is observed. The role of Gln61 residue and other residues of SwII region (Tyr64) and SwI region helps in maintaining the interaction network for GTP hydrolysis [29]. The residue Gly60's (SwII) role in GTP hydrolysis is well established [39, 41, 46]. In agreement to this, the impact of A59G mutation on neighbouring residue Gly60 is highly pronounced in the current simulation as well where 2 distinct maxima were observed for Gly60's interaction with GBR's residues Lys16 and Ser17, in mutant *simulationM*. Additionally, residues namely Tyr32, Thr35 also showed distinct wild-type and mutant behavior. In addition, the majority of the structural differences for the available Ras isoforms are localized in the C-terminal stretch and for the simulation studies in the present work we have excluded this hyper-variable C-terminal region. Furthermore, BLASTp between HRas and KRas proteins sequences shows that the residues ranging between 1–166 (i.e. excluding the hyper-variable C-terminal region) shares 94% identities and 98% similarities amongst the two isoforms (data not shown) [66]. Though the HRas isoform was used for the simulations and analyses for the present study, similar trends and behavior could also be speculated for the KRas isoform as well. A recent review on the available conformational information for the KRas isoform pointed out the role of G12D mutation in populating the metastable states for the wild-type and mutant isoforms [67]. The review outlines the experimental, structural and simulation studies on KRas in a great detail. Main emphasis on the conserved switch regions is also shed, which signifies the switch regions to be highly dynamic in KRas conformers. Further, it has also been postulated that KRas mutations affect the structure in an allosteric manner leading to changes visible at a distant site. Another study which mainly focussed on SwII region's mutants: D33E, A59G demonstrated that these conformers adopt nearly identical conformational as that of HRas [68]. These mutations induced KRas to get crystallized in open state 1 conformation. The DXXGQ motif also showed an increased flexibility across the SwII rearrangement for the A59G KRas mutants. After understanding and enlisting the crucial distinguishing features of the two systems, a multivariate PCA analysis was performed using the R statistical package. The main aim of performing this feature-based PCA is dimensionality reduction, which could give more refined features for the biased simulations like steered MD, metadynamics as well for feature-based analytics package like MSM based analysis tool for example PyEMMA. In the present study, a total of eight features were chosen for this namely: Gln61 and Tyr64 Φ and Ψ dihedrals, distances between Tyr32-Tyr64, Tyr64-Thr35, Tyr64-Gln61 and total energy calculated using MMGBSA. The selected features could segregate the wild-type and mutant trajectories and hence can be used as reaction coordinate/CV for featured based simulations or analyses. The study also observed that the dihedrals of the residue Gln61 are negatively correlated with the rest of the features. To further understand the origin of these distinct behaviors and to comprehensively analyze the two trajectories, Markov State Model (MSM) based analysis of the simulation trajectories was performed using PyEMMA. MSM based analyses can help in systematically re-define the simulation data from long MD simulations in few important metastable structures [56]. Four metastable were observed for the wild-type simulation and five for A59G mutant simulation. The MSM based analysis that resulted in four broad metastable clusters for wild-type, showed little differences in the structures of these states. For the mutant simulation, cluster 5 was an interesting finding, where few structures with the interaction network characteristic of pre-hydrolysis state were also observed, as depicted in the overlap of structures from this cluster with respect to one of the wild-type states.

The present study reconciles that the SwII region is highly influenced by the neighboring A59G mutation and a great degree of restructuring occurs due to this mutation. The absence of pre-hydrolysis state in the mutant conformation is the major reason for this, where the crucial bonding network is not present. The subsequent MSM analyses resulted in few

energetically expensive pre-hydrolysis states even in the mutant simulations. This suggests that the absence of pre-hydrolysis state in the mutant system, required for regular cycling of active-inactive state transition in Ras isoforms, reasons for its permanently active oncogenic state.

Conclusion

The present study focussed on exploring the wild-type and mutant A59G-HRas systems with a perspective of enlisting crucial distinguishing features that could classify and segregate the two systems, in classical or biased MD simulations. The overlap of the available crystal structures of the wild-type and mutant Ras showed visible difference in SwII orientation, especially in side-chain orientation of Gln61 and Tyr64 residues. In the wild-type crystal structure (1QRA), the side chains of these residues point inwards, towards GTP and Mg (i.e. towards the pre-hydrolysis network region). On the other hand, mutant conformation had exactly the opposite orientation of these residues pointing towards the solvent. The MD simulations of these conformations also showed the same trend, which was maintained for the 5 μ s of the respective simulation. The preliminary analyses on these trajectories demonstrated clear discrimination between the SwII region, along with few residue stretches between 25–29, 45–50 and 104–108. Structural analyses of the two trajectories were also carried out to quantify the changes in these residue stretches, as depicted by the relative thickness in the three-dimensional structural overlaps of the two systems. The orientation of SwII and thickness of SwII along with residues 25–29 and 45–50 was clearly observed in this overlap. The findings of basic RMSD, RMSF analyses complement with the crucial interactions of the important residues from conserved SwI, SwII and GBR regions. Residues Gly60, Gln61, Tyr64 (all belonging to SwII region) showed distinct interaction trends with important residues of SwI and GBR. The dihedral angle analysis of Tyr64 was also calculated as this residue showed major differences in its respective orientations in the crystal structure overlaps. MMGBSA calculation was also done for the two trajectories and the coulombic and polar solvation energy components showed a difference of \sim 1000 KJ/mol and \sim 1400 KJ/mol between the wild-type and mutant systems. Combining the outcomes of the above discussed preliminary analyses, a multivariate PCA was carried out using a combination of eight features. The features like ψ dihedral of Gln61 and Tyr64 along with interactions of Tyr64 residue (with Tyr32, Thr35 and Gln61) and total energy component from MM-GBSA analysis were chosen for this PCA. Interestingly, Φ and ψ dihedrals of Gln61 negatively correlated with the other six variates. To summarize, on the basis of these eight features, the mutant and wild-type systems could be clearly differentiated. This suggests that an optimal combination of any of these eight features could be used as reaction coordinate/CV for metadynamics, steered MD, umbrella sampling like biased simulation. Moreover, advanced analyses methods like MSM based techniques of kinetic model generation and thermodynamics assessment could also be used taking into account the above listed features. As feature selection is a key factor in advanced MD simulations, these features might prove useful while selecting the CVs. Lastly, MSM based analysis of the individual trajectories was also carried out using Cadist as the selected feature. The main finding of this analysis was observed in the mutant simulation, where a sparsely populated cluster of energetically expensive metastable states was observed where pre-hydrolysis interacting network was observed. This pre-hydrolysis network is essentially a feature of wild-type ensemble, where the regular switching of active and inactive state conformation is intact and hence is not observed in mutant ensembles as they tend to be stuck in permanently active state. This wild-type-like conformation (having pre-hydrolysis conformation) was missed while performing classical analyses for the 5 μ s mutant simulation trajectory, being very few in number, but MSM analyses could capture this state as well. Upon reconciliation of the basic and advanced analyses and correlating the same with existing

experimental evidence available, the impact of A59G mutation of nearby residues like Gln61 and Tyr64 is established, which also plays a crucial role in GTP hydrolysis.

To conclude, the present study postulates the role of A59G mutation in destabilizing the pre-hydrolysis state normally observed in the wild-type RAS state. MSM analytics could capture the very short lived pre-hydrolysis state in case of A59G mutation. Hence, drug/inhibitor docking and re-purposing studies with this Ras mutant might help to bring back stability to the pre-hydrolysis state essential for the hydrolysis of Ras.

Supporting information

S1 File.

(DOCX)

Acknowledgments

This work was performed using the “Bioinformatics Resources and Applications Facility (BRAAF)” and “National PARAM Supercomputing Facility (NPSF)” at C-DAC, Pune.

Author Contributions

Conceptualization: Neeru Sharma, Uddhavesh Sonavane.

Data curation: Neeru Sharma.

Formal analysis: Neeru Sharma.

Investigation: Neeru Sharma, Uddhavesh Sonavane.

Methodology: Neeru Sharma, Uddhavesh Sonavane.

Supervision: Rajendra Joshi.

Validation: Uddhavesh Sonavane.

Visualization: Neeru Sharma.

Writing – review & editing: Neeru Sharma, Uddhavesh Sonavane, Rajendra Joshi.

References

1. Lu S, Jang H, Muratcioglu S, Gursoy A, Keskin O, Nussinov R, et al. Ras Conformational Ensembles, Allostery, and Signaling. *Chem Rev*. 2016 Jun 8; 116(11):6607–65. <https://doi.org/10.1021/acs.chemrev.5b00542> Epub 2016 1515Jan 27. PMID: 26815308.
2. Castellano E, Santos E. Functional specificity of ras isoforms: so similar but so different. *Genes Cancer*. 2011 Mar; 2(3):216–31. <https://doi.org/10.1177/1947601911408081> PMID: 21779495; PMCID: PMC3128637.
3. Hobbs GA, Der CJ, Rossman KL. RAS isoforms and mutations in cancer at a glance. *J Cell Sci*. 2016 Apr 1; 129(7):1287–92. <https://doi.org/10.1242/jcs.182873> Epub 2016 Mar 16. PMID: 26985062; PMCID: PMC4869631.
4. Olson MF, Marais R. Ras protein signalling. *Semin Immunol*. 2000 Feb; 12(1):63–73. <https://doi.org/10.1006/smim.2000.0208> PMID: 10723799.
5. Stephen AG, Esposito D, Bagni RK, McCormick F. Dragging ras back in the ring. *Cancer Cell*. 2014 Mar 17; 25(3):272–81. <https://doi.org/10.1016/j.ccr.2014.02.017> PMID: 24651010.
6. Takai Y, Sasaki T, Matozaki T. Small GTP-binding proteins. *Physiol Rev*. 2001 Jan; 81(1):153–208. <https://doi.org/10.1152/physrev.2001.81.1.153> PMID: 11152757.
7. Karnoub AE, Weinberg RA. Ras oncogenes: split personalities. *Nat Rev Mol Cell Biol*. 2008 Jul; 9(7):517–31. <https://doi.org/10.1038/nrm2438> PMID: 18568040; PMCID: PMC3915522.

8. Matallanas D, Birtwistle M, Romano D, Zebisch A, Rauch J, von Kriegsheim A, et al. Raf family kinases: old dogs have learned new tricks. *Genes Cancer*. 2011 Mar; 2(3):232–60. <https://doi.org/10.1177/1947601911407323> PMID: 21779496; PMCID: PMC3128629.
9. Kraulis PJ, Domaille PJ, Campbell-Burk SL, Van Aken T, Laue ED. Solution structure and dynamics of ras p21.GDP determined by heteronuclear three- and four-dimensional NMR spectroscopy. *Biochemistry*. 1994 Mar 29; 33(12):3515–31. <https://doi.org/10.1021/bi00178a008> PMID: 8142349.
10. Milburn MV, Tong L, deVos AM, Brünger A, Yamaizumi Z, Nishimura S, et al. Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. *Science*. 1990 Feb 23; 247(4945):939–45. <https://doi.org/10.1126/science.2406906> PMID: 2406906.
11. Barbacid M. ras genes. *Annu Rev Biochem*. 1987; 56:779–827. <https://doi.org/10.1146/annurev.bi.56.070187.004023> PMID: 3304147.
12. Raimondi F, Portella G, Orozco M, Fanelli F. Nucleotide binding switches the information flow in ras GTPases. *PLoS Comput Biol*. 2011 Mar; 7(3):e1001098. <https://doi.org/10.1371/journal.pcbi.1001098> Epub 2011 Mar 3. PMID: 21390270; PMCID: PMC3048383.
13. Simanshu DK, Nissley DV, McCormick F. RAS Proteins and Their Regulators in Human Disease. *Cell*. 2017 Jun 29; 170(1):17–33. <https://doi.org/10.1016/j.cell.2017.06.009> PMID: 28666118; PMCID: PMC5555610.
14. Gasper R, Sot B, Wittinghofer A. GTPase activity of Di-Ras proteins is stimulated by Rap1GAP proteins. *Small GTPases*. 2010 Nov; 1(3):133–141. <https://doi.org/10.4161/sgtp.1.3.14742> PMID: 21686267; PMCID: PMC3116609.
15. Goodsell DS. The molecular perspective: the ras oncogene. *Oncologist*. 1999; 4(3):263–4. PMID: 10394594.
16. Grant BJ, Gorfe AA, McCammon JA. Ras conformational switching: simulating nucleotide-dependent conformational transitions with accelerated molecular dynamics. *PLoS Comput Biol*. 2009 Mar; 5(3):e1000325. <https://doi.org/10.1371/journal.pcbi.1000325> Epub 2009 Mar 20. PMID: 19300489; PMCID: PMC2651530.
17. Scheffzek K, Ahmadian MR, Kabsch W, Wiesmüller L, Lautwein A, Schmitz F, et al. The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science*. 1997 Jul 18; 277(5324):333–8. <https://doi.org/10.1126/science.277.5324.333> PMID: 9219684.
18. Bos JL, Rehmann H, Wittinghofer A. GEFs and GAPs: critical elements in the control of small G proteins. *Cell*. 2007 Jun 1; 129(5):865–77. <https://doi.org/10.1016/j.cell.2007.05.018> Erratum in: *Cell*. 2007 Jul 27; 130(2):385. PMID: 17540168.
19. Glennon TM, Villà J, Warshel A. How does GAP catalyze the GTPase reaction of Ras? A computer simulation study. *Biochemistry*. 2000 Aug 15; 39(32):9641–51. <https://doi.org/10.1021/bi000640e> PMID: 10933780.
20. Cherfils J, Chardin P. GEFs: structural basis for their activation of small GTP-binding proteins. *Trends Biochem Sci*. 1999 Aug; 24(8):306–11. [https://doi.org/10.1016/s0968-0004\(99\)01429-2](https://doi.org/10.1016/s0968-0004(99)01429-2) PMID: 10431174.
21. Scheidig AJ, Burmester C, Goody RS. The pre-hydrolysis state of p21(ras) in complex with GTP: new insights into the role of water molecules in the GTP hydrolysis reaction of ras-like proteins. *Structure*. 1999 Nov 15; 7(11):1311–24. [https://doi.org/10.1016/s0969-2126\(00\)80021-0](https://doi.org/10.1016/s0969-2126(00)80021-0) PMID: 10574788.
22. Heesen H, Gerwert K, Schlitter J. Role of the arginine finger in Ras.RasGAP revealed by QM/MM calculations. *FEBS Lett*. 2007 Dec 11; 581(29):5677–84. <https://doi.org/10.1016/j.febslet.2007.11.026> Epub 2007 Nov 20. PMID: 18022389.
23. Khrenova MG, Grigorenko BL, Kolomeisky AB, Nemukhin AV. Hydrolysis of Guanosine Triphosphate (GTP) by the Ras-GAP Protein Complex: Reaction Mechanism and Kinetic Scheme. *J Phys Chem B*. 2015 Oct 8; 119(40):12838–45. <https://doi.org/10.1021/acs.jpcc.5b07238> Epub 2015 Sep 25. PMID: 26374425.
24. Schubert S, Shannon K, Bollag G. Hyperactive Ras in developmental disorders and cancer. *Nat Rev Cancer*. 2007 Apr; 7(4):295–308. <https://doi.org/10.1038/nrc2109> Erratum in: *Nat Rev Cancer*. 2007 Jul; 7(7):563. PMID: 17384584.
25. Lukman S, Grant BJ, Gorfe AA, Grant GH, McCammon JA. The distinct conformational dynamics of K-Ras and H-Ras A59G. *PLoS Comput Biol*. 2010 Sep 9; 6(9):e1000922. <https://doi.org/10.1371/journal.pcbi.1000922> PMID: 20838576; PMCID: PMC2936511.
26. Grant BJ, Lukman S, Hocker HJ, Sayyah J, Brown JH, McCammon JA, et al. Novel allosteric sites on Ras for lead generation. *PLoS One*. 2011; 6(10):e25711. <https://doi.org/10.1371/journal.pone.0025711> Epub 2011 Oct 25. PMID: 22046245; PMCID: PMC3201956.
27. Hocker HJ, Cho KJ, Chen CY, Rambahal N, Sagineedu SR, Shaari K, et al. Andrographolide derivatives inhibit guanine nucleotide exchange and abrogate oncogenic Ras function. *Proc Natl Acad Sci U S A*.

- 2013 Jun 18; 110(25):10201–6. <https://doi.org/10.1073/pnas.1300016110> Epub 2013 Jun 4. PMID: [23737504](https://pubmed.ncbi.nlm.nih.gov/23737504/); PMCID: PMC3690838.
28. Lito P, Solomon M, Li LS, Hansen R, Rosen N. Allele-specific inhibitors inactivate mutant KRAS G12C by a trapping mechanism. *Science*. 2016 Feb 5; 351(6273):604–8. <https://doi.org/10.1126/science.aad6204> Epub 2016 Jan 14. PMID: [26841430](https://pubmed.ncbi.nlm.nih.gov/26841430/); PMCID: PMC4955282.
 29. Shurki A, Warshel A. Why does the Ras switch "break" by oncogenic mutations? *Proteins*. 2004 Apr 1; 55(1):1–10. <https://doi.org/10.1002/prot.20004> PMID: [14997535](https://pubmed.ncbi.nlm.nih.gov/14997535/).
 30. Gao C, Eriksson L. Impact of Mutations on K-Ras-p120GAP Interaction. *Computational Molecular Bio-science*. 2013 May 30; 3(2), 9–17. <https://doi.org/10.4236/cmb.2013.32002>
 31. Adjei AA. Blocking oncogenic Ras signaling for cancer therapy. *J Natl Cancer Inst*. 2001 Jul 18; 93(14):1062–74. <https://doi.org/10.1093/jnci/93.14.1062> PMID: [11459867](https://pubmed.ncbi.nlm.nih.gov/11459867/).
 32. Downward J. Targeting RAS signalling pathways in cancer therapy. *Nat Rev Cancer*. 2003 Jan; 3(1):11–22. <https://doi.org/10.1038/nrc969> PMID: [12509763](https://pubmed.ncbi.nlm.nih.gov/12509763/).
 33. Foley CK, Pedersen LG, Charifson PS, Darden TA, Wittinghofer A, Pai EF, et al. Simulation of the solution structure of the H-ras p21-GTP complex. *Biochemistry*. 1992 Jun 2; 31(21):4951–9. <https://doi.org/10.1021/bi00136a005> PMID: [1599919](https://pubmed.ncbi.nlm.nih.gov/1599919/).
 34. Díaz JF, Wroblowski B, Engelborghs Y. 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*. 1995 Sep 19; 34(37):12038–47. <https://doi.org/10.1021/bi00037a047> PMID: [7547942](https://pubmed.ncbi.nlm.nih.gov/7547942/).
 35. Kuppens S, Hellings M, Jordens J, Verheyden S, Engelborghs Y. Conformational states of the switch I region of Ha-ras-p21 in hinge residue mutants studied by fluorescence lifetime and fluorescence anisotropy measurements. *Protein Sci*. 2003 May; 12(5):930–8. <https://doi.org/10.1110/ps.0236303> PMID: [12717016](https://pubmed.ncbi.nlm.nih.gov/12717016/); PMCID: PMC2323864.
 36. Grigorenko BL, Nemukhin AV, Shadrina MS, Topol IA, Burt SK. Mechanisms of guanosine triphosphate hydrolysis by Ras and Ras-GAP proteins as rationalized by ab initio QM/MM simulations. *Proteins*. 2007 Feb 1; 66(2):456–66. <https://doi.org/10.1002/prot.21228> PMID: [17094109](https://pubmed.ncbi.nlm.nih.gov/17094109/).
 37. Xia F, Rudack T, Kötting C, Schlitter J, Gerwert K. The specific vibrational modes of GTP in solution and bound to Ras: a detailed theoretical analysis by QM/MM simulations. *Phys Chem Chem Phys*. 2011 Dec 28; 13(48):21451–60. <https://doi.org/10.1039/c1cp22741f> Epub 2011 Nov 2. PMID: [22048726](https://pubmed.ncbi.nlm.nih.gov/22048726/).
 38. Prakash P, Gorfe AA. Overview of simulation studies on the enzymatic activity and conformational dynamics of the GTPase Ras. *Mol Simul*. 2014; 40(10–11):839–847. <https://doi.org/10.1080/08927022.2014.895000> Epub 2014 Mar 19. PMID: [26491216](https://pubmed.ncbi.nlm.nih.gov/26491216/); PMCID: PMC4610817.
 39. Sharma N, Sonavane U, Joshi R. Probing the wild-type HRas activation mechanism using steered molecular dynamics, understanding the energy barrier and role of water in the activation. *Eur Biophys J*. 2014 Mar; 43(2–3):81–95. <https://doi.org/10.1007/s00249-014-0942-4> Epub 2014 Jan 20. PMID: [24442446](https://pubmed.ncbi.nlm.nih.gov/24442446/).
 40. Prakash P, Sayyed-Ahmad A, Gorfe AA. The role of conserved waters in conformational transitions of Q61H K-ras. *PLoS Comput Biol*. 2012; 8(2):e1002394. <https://doi.org/10.1371/journal.pcbi.1002394> Epub 2012 Feb 16. PMID: [22359497](https://pubmed.ncbi.nlm.nih.gov/22359497/); PMCID: PMC3280954.
 41. Pai EF, Krengel U, Petsko GA, Goody RS, Kabsch W, Wittinghofer A. Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *EMBO J*. 1990 Aug; 9(8):2351–9. PMID: [2196171](https://pubmed.ncbi.nlm.nih.gov/2196171/); PMCID: PMC552258.
 42. Smith MJ, Neel BG, Ikura M. NMR-based functional profiling of RASopathies and oncogenic RAS mutations. *Proc Natl Acad Sci U S A*. 2013 Mar 19; 110(12):4574–9. <https://doi.org/10.1073/pnas.1218173110> Epub 2013 Mar 4. PMID: [23487764](https://pubmed.ncbi.nlm.nih.gov/23487764/); PMCID: PMC3607025.
 43. Khrenova MG, Mironov VA, Grigorenko BL, Nemukhin AV. Modeling the role of G12V and G13V Ras mutations in the Ras-GAP-catalyzed hydrolysis reaction of guanosine triphosphate. *Biochemistry*. 2014 Nov 18; 53(45):7093–9. <https://doi.org/10.1021/bi5011333> Epub 2014 Nov 4. PMID: [25339142](https://pubmed.ncbi.nlm.nih.gov/25339142/).
 44. Grigorenko BL, Nemukhin AV, Topol IA, Cachau RE, Burt SK. QM/MM modeling the Ras-GAP catalyzed hydrolysis of guanosine triphosphate. *Proteins*. 2005 Aug 15; 60(3):495–503. <https://doi.org/10.1002/prot.20472> PMID: [15906320](https://pubmed.ncbi.nlm.nih.gov/15906320/).
 45. Martín-García F, Mendieta-Moreno JI, López-Viñas E, Gómez-Puertas P, Mendieta J. The Role of Gln61 in HRas GTP hydrolysis: a quantum mechanics/molecular mechanics study. *Biophys J*. 2012 Jan 4; 102(1):152–7. <https://doi.org/10.1016/j.bpj.2011.11.4005> Epub 2012 Jan 3. Erratum in: *Biophys J*. 2012 Mar 7; 102(5):1234. PMID: [22225809](https://pubmed.ncbi.nlm.nih.gov/22225809/); PMCID: PMC3250681.
 46. Muraoka S, Shima F, Araki M, Inoue T, Yoshimoto A, Ijiri Y, 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

- Jun 12; 586(12):1715–8. <https://doi.org/10.1016/j.febslet.2012.04.058> Epub 2012 May 11. PMID: 22584058.
47. Hall BE, Bar-Sagi D, Nassar N. The structural basis for the transition from Ras-GTP to Ras-GDP. *Proc Natl Acad Sci U S A*. 2002 Sep 17; 99(19):12138–42. <https://doi.org/10.1073/pnas.192453199> Epub 2002 Sep 4. PMID: 12213964; PMCID: PMC129411.
 48. Sharma N, Sonavane U, Joshi R. Differentiating the pre-hydrolysis states of wild-type and A59G mutant HRas: An insight through MD simulations. *Comput Biol Chem*. 2017 Aug; 69:96–109. <https://doi.org/10.1016/j.compbiolchem.2017.05.008> Epub 2017 Jun 1. PMID: 28600956.
 49. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF Jr, Brice MD, Rodgers JR, et al. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J Mol Biol*. 1977 May 25; 112(3):535–42. [https://doi.org/10.1016/s0022-2836\(77\)80200-3](https://doi.org/10.1016/s0022-2836(77)80200-3) PMID: 875032.
 50. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput*. 2008 Mar; 4(3):435–47. <https://doi.org/10.1021/ct700301q> PMID: 26620784.
 51. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*. 2006 Nov 15; 65(3):712–25. <https://doi.org/10.1002/prot.21123> PMID: 16981200; PMCID: PMC4805110.
 52. Schüttelkopf AW, van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr*. 2004 Aug; 60(Pt 8):1355–63. <https://doi.org/10.1107/S0907444904011679> Epub 2004 Jul 21. PMID: 15272157.
 53. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004 Oct; 25(13):1605–12. <https://doi.org/10.1002/jcc.20084> PMID: 15264254.
 54. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996 Feb; 14(1):33–8, 27–8. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) PMID: 8744570.
 55. <http://plasma-gate.weizmann.ac.il/Grace/>
 56. Scherer MK, Trendelkamp-Schroer B, Paul F, Pérez-Hernández G, Hoffmann M, Plattner N, et al. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J Chem Theory Comput*. 2015 Nov 10; 11(11):5525–42. <https://doi.org/10.1021/acs.jctc.5b00743> Epub 2015 Oct 14. PMID: 26574340.
 57. Prinz JH, Wu H, Sarich M, Keller B, Senne M, Held M, et al. Markov models of molecular kinetics: generation and validation. *J Chem Phys*. 2011 May 7; 134(17):174105. <https://doi.org/10.1063/1.3565032> PMID: 21548671.
 58. Chodera JD, Noé F. Markov state models of biomolecular conformational dynamics. *Curr Opin Struct Biol*. 2014 Apr; 25:135–44. <https://doi.org/10.1016/j.sbi.2014.04.002> Epub 2014 May 16. PMID: 24836551; PMCID: PMC4124001.
 59. Chodera JD, Singhal N, Pande VS, Dill KA, Swope WC. Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics. *J Chem Phys*. 2007 Apr 21; 126(15):155101. <https://doi.org/10.1063/1.2714538> PMID: 17461665.
 60. Sirur A, De Sancho D, Best RB. Markov state models of protein misfolding. *J Chem Phys*. 2016 Feb 21; 144(7):075101. <https://doi.org/10.1063/1.4941579> PMID: 26897000; PMCID: PMC4760979.
 61. Wu H, Noé F. Variational Approach for Learning Markov Processes from Time Series Data. *J Nonlinear Sci*. 2020 Feb 01; 30, 23–66.
 62. Noé F, Schütte C, Vanden-Eijnden E, Reich L, Weikl TR. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. Version 2. *Proc Natl Acad Sci U S A*. 2009 Nov 10; 106(45):19011–6. <https://doi.org/10.1073/pnas.0905466106> Epub 2009 Nov 3. PMID: 19887634; PMCID: PMC2772816.
 63. Ford B, Hornak V, Kleinman H, Nassar N. Structure of a transient intermediate for GTP hydrolysis by ras. *Structure*. 2006 Mar; 14(3):427–36. <https://doi.org/10.1016/j.str.2005.12.010> PMID: 16531227.
 64. Chen J, Wang W, Pang L, Zhu W. Unveiling conformational dynamics changes of H-Ras induced by mutations based on accelerated molecular dynamics. *Phys Chem Chem Phys*. 2020 Sep 15. <https://doi.org/10.1039/d0cp03766d> Epub ahead of print. PMID: 32930679.
 65. Menyhárd DK, Pálffy G, Orgován Z, Vida I, Keserű GM, Perczel A. Structural impact of GTP binding on downstream KRAS signaling. *Chem Sci*. 2020 Aug; 11:9272–9289. <https://doi.org/10.1039/d0sc03441j>
 66. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990 Oct 5; 215(3):403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: 2231712.
 67. Pantsar T. The current understanding of KRAS protein structure and dynamics. *Comput Struct Biotechnol J*. 2019 Dec 26; 18:189–198. <https://doi.org/10.1016/j.csbj.2019.12.004> PMID: 31988705; PMCID: PMC6965201.

68. Lu J, Bera AK, Gondi S, Westover KD. KRAS Switch Mutants D33E and A59G Crystallize in the State 1 Conformation. *Biochemistry*. 2018 Jan 23; 57(3):324–333. <https://doi.org/10.1021/acs.biochem.7b00974> Epub 2017 Dec 28. PMID: 29235861.