A cancer-associated K-RAS mutant that preferentially binds GDP displays gain-of-function and atypical signaling

Xiaolan Qian
National Institutes of Health
Tatu Pantsar
University of Eastern Finland
Alex G. Papageorge1
National Institutes of Health
Brajendra K. Brajendra K. Tripathi
National Institutes of Health
Beatriz Sanchez-Solana
National Institutes of Health
Marian E. Durkin
National Institutes of Health
Yucheng Gao
National Institutes of Health
Dominic Esposito
National Cancer Institute RAS Initiative, Leidos Biomedical Research, Inc
Timothy Waybright
National Cancer Institute RAS Initiative, Leidos Biomedical Research, Inc
Andrew G. Stephen Stephen
National Institutes of Health
Dhirendra Simanshu
National Cancer Institute RAS Initiative, Leidos Biomedical Research, Inc
Dunrui Wang
National Institutes of Health
Douglas R. Lowy (lowyd@mail.nih.gov)
National Institutes of Health

Research Article

Keywords:

Posted Date: April 6th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1521657/v1

A cancer-associated K-RAS mutant that preferentially binds GDP displays gain-of-function and atypical signaling

Xiaolan Qian[1], Tatu Pantsar[2,3], Alex G. Papageorge[1], Brajendra K. Tripathi[1], Beatriz Sanchez-Solana[1], Marian E. Durkin[1], Yucheng Gao[1^], Dominic Esposito[4], Timothy Waybright[4], Andrew G. Stephen[4], Dhirendra Simanshu[4], Dunrui Wang[1], and Douglas R. Lowy[1]

[1]Laboratory of Cellular Oncology, Center for Cancer Research, National Cancer Institute, Bethesda, MD USA; [2]School of Pharmacy, Faculty of Health Sciences, University of Eastern Finland, Kuopio, Finland; [3]Department of Pharmaceutical and Medicinal Chemistry, Institute of Pharmaceutical Sciences, Eberhard Karls University Tübingen, Tübingen, Germany; [4]National Cancer Institute RAS Initiative, Cancer Research Technology Program, Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Inc., Frederick, MD USA

[^]Current address: Pinetree Theraputics; e-mail: Sophy Yucheng Gao <sophygao@pinetreetheraputics.com>

The authors declare no potential conlicts of interest

Running title: Gain-of-function K-RAS mutant protein that preferentially binds GDP

Abstract

Background: The *RAS* genes, which are the prototype for the RAS superfamily of small GTPases, encode guanine nucleotide binding proteins that are frequently mutated in cancer. The superfamily has five major branches: RAS, RHO, RAB, RAN, and ARF. Its wild type members function as molecular switches that are active when GTP-bound and inactive when GDP-bound. Previously analyzed cancer-associated point mutants of RAS, such as those affecting codons 12, 13, and 61, and mutants of other superfamily members have followed this paradigm, as their steady-state level of bound GTP is higher than that of the wild type protein. However, it is possible that a less common cancer-associated mutant might be active when bound to GDP.

Methods: We searched the Catalog of Somatic Mutations in Cancer (COSMIC) database and the TCGA database in the NCI Genomic Data Commons cancer-associated mutations in the RAS superfamily beyond the commonly mutated RAS codons, conducted biological and biochemical tissue culture studies and in vitro biochemical guanine nucleotide binding studies of a *K-RAS* mutant identified by this search, and complemented these findings with all-atom molecular dynamics simulations.

Results: The database search identified a recurrent cancer-associated *K-RAS* mutant, M67L, as well as another cancer-associated mutant that affect the same codon in *K-RAS* and *N-RAS* and additional point mutants in the analogous codon in 10 other members of the RAS superfamily of GTPases. Detailed analysis of the K-RAS M67L mutant indicated that it has a gain-of-function biologically, but its low GTP level is similar to that of wild type K-RAS. Unlike the GDP-bound wild type, the GDP-bound M67L mutant interacts with and activates at least three RAS effectors: RALGDS, PI3K-p110α, and RAPGEF6. However, it interacts weakly with the prototypic RAS

effector, RAF. All-atom molecular dynamics simulations of the mutant, conducted with and without RALGDS, provide structural insight into these characteristics of the GDP-bound mutant.

Conclusion: Cancer-associated mutations that are active when GDP-bound occur in K-RAS and probably in other members of the RAS superfamily.

Introduction

The wild type proteins encoded by the *RAS* superfamily of small GTPases, a group of more than 150 phylogenetically related genes, are highly regulated molecular switches that cycle between an active, GTP-bound form, and an inactive, GDP-bound form [1]. Most of the superfamily is divided into five major families: RAS, RHO, RAB, RAN, and ARF. These GTPases are positively regulated and activated by guanine nucleotide exchange factors (GEFs), which convert them from being GDP-bound and inactive to GTP-bound and active, and negatively regulated by GTPase-activating proteins (GAPs), which inactivate the small GTPases by converting them from being GTP-bound to GDP-bound [2]. The GEFs and GAPs are specific to each family, and some regulate only a subset of the GTPases in a given family. The guanine nucleotide binding domain of the superfamily have two regions, switch I and switch II, whose conformation changes according to whether the protein is bound to GDP or GTP [3]. Amino acids in switch I account for most of the GTP-dependent interactions with downstream effector proteins, while switch I and switch II, together with amino acids outside these regions, make critical contributions to the sensitivity of family members to their respective GEFs and GAPs [4, 5].

The most extensively studied *RAS* superfamily members have been the *RAS* GTPases, which are activated by somatic point mutation in more than one-quarter of human cancer [6]. Of the three genes that form the *RAS* family – *H-RAS*, *K-RAS*, and *N-RAS* – *K-RAS* has been studied in the more detail because it is frequently mutated in common cancers that account for a high proportion of cancer mortality. This includes ~85% of pancreatic adenocarcinomas, ~40% of colorectal cancers, and ~30% of lung adenocarcinomas [7]. The majority of cancer-associated mutations of *K-RAS* and the other *RAS* genes affect codons 12, 13, and 61, but there are also less frequent recurrent mutations of other codons [8]. The proteins encoded by the cancer-associated

K-RAS mutants that affect the commonly mutated codons have steady-state levels of bound GTP that are higher than that of wild type K-RAS, mainly because they have lower intrinsic GTP hydrolysis and lower sensitivity to GAPs than wild type K-RAS [9]. The mutant and wild type K-RAS proteins bind and activate, in a GTP-dependent manner, several direct downstream targets; three of the best studied are RAF, PI3K, and RALGDS [10, 11].

The *RHO* GTPases have also been widely implicated in cancer, although they are mutated less frequently than *RAS*. In the *RHO* family, *RHOA* is mutated most frequently, and when it is, the mutant alleles usually involve codons other than those that correspond to 12, 13, and 61 in *RAS* [12]. Therefore, the mechanism of action in at least some cancer-associated *RHOA* mutants might differ from those that underlie the commonly affected codons in *RAS*. In unpublished observations, we have identified recurrent cancer-associated RHOA mutants that affect a specific amino acid (codon-69) in the switch II region that have gain-of-function when bound to GDP. Codon-69 in *RHOA* corresponds to codon-67 in the *RAS* genes, as the N-terminus encoded by RHOA has two amino acids more than RAS.

In the current study, we have therefore screened a large cancer database for mutation of codon-67 in *K-RAS* and the corresponding codon in other members of the superfamily. The screen identified a cancer-associated mutation of the equivalent codon in more than 10 members of the superfamily, including a recurrent mutation of this codon in *K-RAS*. In contrast to previously described K-RAS mutant proteins, the in vivo level of GTP in the newly identified K-RAS mutant is similar to that of the wild type protein, but the mutant appears to be active when bound to GDP. The GDP-bound mutant interacts with and activates at least three K-RAS effectors: PI3K, RALGDS, and RAPGEF6. It also activates the RAF-MEK-MAPK pathway, but accomplishes it via alternate signaling mechanisms, rather than mainly by the canonical

mechanism of binding directly to a RAF protein. In addition, our structural investigations by all-atom molecular dynamics (MD) simulations can account for the ability of the GDP-bound mutant to interact with effectors such as RALGDS and PI3K-p110α as well as its reduced ability to interact with RAF.

Methods

Generation of cancer-associated missense mutations in K-RAS constructs by mutagenesis, plasmids and constructs

Myc-DDK tagged human K-RAS in pCMV6 vector (Origene) expressing wild-type K-RAS was used as template. The selected missense mutations based on bioinformatic analysis were mediated by the full list of designed primers as shown in Supplementary Table 1. Eighteen cycles of PCR were performed using a site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA) according to manufacturer's instructions. All the resultant mutations in the plasmids were confirmed by DNA sequencing. The PEBG mammalian expression vector [13] was used for GST-fusion proteins after subcloning the PCR products from full-length human *K-RAS*, *RAP1a*, *B-RAF-RBD* (151-233), *PI3K-p110α-RBD* (176-300) and *RAPGEF-RA* (737-845) using the primers listed in Supplementary Table 1. The primers included 5' KpnI and 3' NotI restriction sites. The pcDNA3-MyrSos1ΔDbl and p120RasGAP expression vectors were described previously [14]. All these constructs were verified by sequencing analysis.

RNA extraction and qPCR validation of microarray data

RNA was isolated from K-RAS stably expressing H358 clones at equivalent levels using TRIzol reagent (Invitrogen), according to the manufacturer's protocol. RNA quality was assessed using

an Agilent Tapestation, and high quality RNAs were sent for Microarray analysis at Laboratory of Molecular Technology, Advanced Technology Program, Leidos Biomedical Research, Inc. Frederick National Laboratory for Cancer Research. To validate the expression of genes of intereste from the microarray data, 1µg of original total RNA was used to synthesize cDNA using QuantiTech Reverse transcription kit, and PCR was run using FastStart Universal SYBR Green Master Mix (Roche) in a QuantStudio 5 Real-Time PCR system (Applied Biosystems). Actin expression was used to normalize the data. The human Actin, RB, c-MYC and EPCAM primers are listed in Supplementary Table 1.

Cell culture, Transfection, transwell migration and anchorage-independent growth assays

HEK 293TT [15] and colon cancer line SW620 cells (ATCC) were cultured in DMEM supplemented with 10% FBS. NSCLC lines (H1299, H358 and H1703 from ATCC) were cultured in RPMI-1640 supplemented with 10% FBS. Transient transfections were performed with Lipofectamine 3000 (Invitrogen), and cells were assayed after 48 hours. Stable clones expressing K-RAS mutants were made by transfecting SW620 or H358 cells with Lipofectamine 3000, followed by G418 selection (0.9 mg/ml) for three weeks.

The growth conditions for the Human NSCLC lines H358, H1299, and H1703 have been described previously [13], as have those for the SW620 colon cancer line [16]. H358 and SW620 cells were stably transfected with vector, wild-type, or mutant *K-RAS-DDK* constructs using Lipofectamine 3000 (Invitrogen), and colonies were generated after G418 (0.9 mg/ml) selection. H1299 cells were transiently transfected with *K-RAS* siRNA to knock down endogenous *K-RAS*, or transfected with *K-RAS-DDK* constructs and cultured for 48 h. 293TT cells[15] were transiently transfected with *K-RAS* siRNA to knock-down endogenous *K-RAS*,

then transfected with *PEBG-B-RAF-RBD, PEBG-p110α–RBD* or *PEBG-RAPGF6-RA* constructs and cultured for 48 h, followed by purification of GST-fusion proteins using glutathione beads as described[23]. The Transwell cell migration assay was performed using 6.5 mm diameter Falcon cell culture inserts (8 μM pore size, ThermoFisher) precoated with 0.01% gelatin, in 24 well cell culture plates as described [17]. The collected lysates from the stained migrated cells were quantified in a spectrophotometer using OD_{590nm}. For soft agar assay, $1X10^5$ of G418-resistant stable cell clones were mixed with complete media containing 0.4% of ultrapure agar (Invitrogen) and placed over 0.6% basal agar in 6 cm dishes in triplicate. Cells were fed weekly with fresh media and grown for 3 weeks, and colonies were photographed microscopically, quantified by a colony counter after staining with with Nitrotetrazolium Blue Chloride (1 mg/ml), and plotted with Prizm.

In vivo analysis of K-RAS GTP

In vivo analysis of GTP-Ras was performed as described previously, with minor modifications [18]. H1299 cells were transiently transfected for 36 hours and metabolically labeled with [^{32}P] orthophosphate for 10 hr. Cells were rinsed with PBS and lysed in 20 mM Tris-HCl (pH 7.4), 100 mM NaCl, 20 mM $MgCl_2$, 1%NP-40, 0.5% sodium deoxycholate. Immunoprecipitation of DDK tagged K-RAS proteins was performed with anti FLAG antibody and γBind Sepharose. Following chromatography on PEI cellulose plates, the chromatogram was visualized on Molecular Dynamics STORM 860 and analyzed in ImageQuant.

K-RAS and other small GTPase downstream effector RBD pull-down assay, co-immunoprecipitation (co-IP), immunoblotting (IB), and ERK in vitro kinase assays

Cells were lysed with Mg^{2+} lysis buffer (EMD Millipore) and the total proteins were estimated by BCA assay (Thermo Scientific) [13]. Equal amounts of protein from cell extracts were used for RAF1-RBD, PAK-1-RBD, and RALGDS-RBD pull-down assays according to manufacturer instructions, followed by immunoblotting with antibodies indicated in figures. RALBP1-RBD pull-down assay was performed using RAL activation kit (EMD Millipore). For co-immunoprecipitation (co-IP) experiments, equal amounts of protein from each cell extracts were precleared with protein A/G slurry (Thermo Fisher Scientific). After incubation with the indicated antibodies or control IgG for 1 h, 25 µl of protein A/G slurry was added to each immune reaction and rotated at 4°C overnight. The immunopellets were washed three times as described [17]. Co-IP proteins were eluted by boiling for 5 min in 30 µl Laemmli sample buffer and separated on a reducing SDS-PAGE. All endogenous proteins and those expressed from transfected plasmids were analyzed by immunoblotting with the following antibodies: anti-mouse and anti-rabbit DDK (Origene,cat # TA50011 and TA100023), RHOA, RAL, RAP1a (EMD Millipore, cat # 05-778, 05-586, 07-916, respectively), RAC (BD biosciences, cat# R56220), the antibodies from Cell Signaling: pERK (cat# 9101), ERK (cat# 9102), pAKT(S473, cat # 4668), AKT (cat# 2938), pmTOR (S2448), mTOR (cat# 2971), pJNK (T183/Y185, cat# 9252), JNK (cat# 9252), c-MYC (cat# 13987), EpCAM (cat# 3599), K-RAS (Protientech, cat#12063-1-AP), and the antibodies from Santa Cruz: GST (cat# sc-138), p110a (cat# sc1332), Sos1(cat# sc-256), p120RASGAP (cat# sc63-136481), CDC42 (cat# sc390210), RB (cat# sc-102), RAF-B (cat# sc-166), RAF-1 (cat# sc-227), RALGDS (cat# sc-393809), RAPGEF6 (cat# sc-398642), ACTIN (Sigma, cat # A4700) and HA antibody (BioLegend, cat# 902303) followed by secondary anti-IgG conjugated with HRP (1:10,000, GSS). The signals bound to the membranes were detected by ECL or ECL plus kit (Thermo Scientific). For the ERK in vitro

kinase assay, extracts from cells co-transfected with plasmids encoding HA-ERK and the K-RAS mutants or wild type were immunoprecipitated with HA antibody (BioLegend) followed by washing with three different buffers [18]. The immunopellets were subject to an ERK in vitro kinase assay, adding 2 μg of myelin basic protein (MBP, EMD Millipore), 20 μM ATP, 1 μCi of [γ-^{32}P]ATP (NEN-Dupont) in 25 μl of reaction buffer (12.5 mM MOPS [morpholinepropanesulfonic acid]; pH 7.4, 12.5 mM β-glycerophosphate, 7.5 mM $MgCl_2$, 0.5 mM EGTA, 0.5 mM sodium orthovanadate, 0.5 mM NaF). After incubation for 20 min at 30°C, kinase reactions were terminated by the addition of 2X Laemmli sample buffer. The samples were then resolved by SDS-PAGE, and the phosphorylated MBP was visualized by autoradiography.

Purification of recombinant proteins and preparation of exclusively GDP-bound and GTP-bound K-RAS proteins

PEBG-RBD derived from p110α or B-RAF and PEBG-RA derived from RAPGEF6 were expressed in 293TT cells and purified by glutathione beads pull-down followed by six times wash as described [17]. The quantity of purified GST-fusion proteins was estimated by Coomassie blue stained gel compared with the known concentrations of the albumin standard. Purified human Recombinant, 6xHis-Tag K-RAS G12D (Cytoskeleton), 6xHis-Tag K-RAS wild-type and M67L (Sino Biological) were used for in vitro analysis. K-RAS4b (1-169) expression clones for *E. coli* production were generated by fusing to His6-tagged maltose binding protein (MBP). K-RAS4b point mutants were generated by DNA synthesis and are based on the sequence of the wild-type construct (Addgene #159539). The His6-MBP-KRAS4b fusion proteins were expressed in *E. coli* as described [19]. All proteins were purified as outlined for G-

Hs.KRAS4b (1-169) as described [20]. Briefly, the His6-MBP-KRAS4b fusion proteins were purified from clarified lysates by immobilized metal-affinity chromatography (IMAC), treated with His6-TEV protease to release the target protein, the target protein was separated from other components of the TEV protease reaction by a second round of IMAC, and the final protein subjected to SEC (size exchange chromatography). GDP-bound avi-KRAS proteins were exchanged into an ammonium phosphatase buffer containing 40 mM Tris-HCl (pH 8.0), 200 mM $(NH_4)_2SO_4$, and 0.1 mM $ZnCl_2$. Alkaline phosphatase beads and a 10:1 molar ratio of a non-hydrolysable analog of GTP, GMPPNP, were then added and incubated at 25ºC for 3 h. A 0.22-micron centrifugal filter tube was used to remove the alkaline phosphatase beads by centrifugation at 14,000 rpm for 5 min. 30 mM $MgCl_2$ and a 5:1 molar ratio of GMPPNP (GTP) were then added and incubated at 25ºC for an additional 2 h. Excess nucleotide was removed by passing the reaction through a HiPrep 26/10 desalting column (GE Healthcare). The efficiency of nucleotide exchange of K-RAS was determined using high-performance liquid chromatography (HPLC). Exchanged proteins were diluted into 0.1 M K_2HPO_4 and 1 mM tetrabutyl ammonium hydrogen sulfate (buffer A), and injected onto an Ultrasphere 5 ODS, 250 x 4.6 mm column (Hichrom, Berkshire, UK). Bound nucleotides were eluted using a linear gradient of buffer A containing 30% acetonitrile at a flow rate of 0.6 mL/min. Standards of GDP and GMPPNP (GTP analog) were included to validate the identity of nucleotide bound. GMPPNP (GTP) exchange efficiency was routinely >95% pure as measured by this assay.

Immunofluorescent staining and Fluorescent confocal microscopy

H1703 cells were transfected with the pEGFP-C1 vector (Takara Bio Inc.) or the DDK-tagged indicated wild-type or mutant K-RAS constructs and seeded onto glass chamber slides with

75,000 cells per chamber. After 48 h of transfection, cells were fixed with 4% paraformaldehyde for 20 minutes, permeabilized with 0.25% Triton X-100 in PBS for 5 minutes and blocked with 3% BSA for 2 h. The cells were incubated with a 1:500 dilution (in 3% BSA) of the GFP mouse (Abcam, Inc.) or DDK (Origene) antibodies at 4°C overnight. After thoroughly washing with PBS, cells were incubated with 1:100 Alexa Fluor 568® phalloidin and 1:250 Alexa Flour 488 anti-mouse-IgG (A11001) secondary antibodies and DAPI (Thermo Fisher Scientific) for 1 hour. To visualize nuclei, cells were incubated with DAPI (1:2,500) for 1 h. After staining, cells were thoroughly washed with PBS and mounted with gel mounting solution (Biomeda Corporation). Confocal microscopy of fluorescent-labeled cells was performed using a microscope (LSM 780; Carl Zeiss) with an excitation wavelength of 488 nm and 568 nm to detect transfected GFP or DDK fusion proteins. Stress fibers (phalloidin) were viewed with an excitation wavelength of 568 nm. Images were made at room temperature using photomultiplier tubes with a Plan-Apochromat 63X/1.4 NA oil differential interference contrast objective lens. For quantification of representative morphology in each group, ~20 cells per condition randomly selected from several fields were analyzed.

siRNA transfection and inhibitor treatment

To knock-down endogenous K-RAS or RAPGEF6 expression, H1299 cells were transfected with 180 nM of siRNAs for human K-RAS, RAPGEF6, or with scrambled control siRNAs (Qiagen) for 24 hours followed by DNA construct transfection and incubation for anther 48 hours. Cell extracts were collected 48 h later and added to purified K-RAS loaded with GTP or GDP and pulled-down by p110α-RBD. Suppression of K-RAS protein expression, with two different siRNAs, was confirmed by immunoblotting. Validated siRNAs for human K-RAS (Hs_K-RAS

siRNA_8, SI02662051, target sequence 5′-AAGGAGAATTTAATAAAGATA-3′, Hs_K-RAS siRNA_2, SI00071015, target sequence 5′-CAGACGTATTGTATCATTT -3′) and siRNA for human RAPGEF6 (Hs_RAPGEF6_5, SI03019170) target sequence 5'-CAAGAGCATATCATTATAGAA-3 were from QIAGEN, as were negative control siRNA (control siRNA 1027280. RAF inhibitor Dabrafenib (20 mM), MEK inhibitor Trametinib (1mM), and PAK inhibitor PF-3758309 (10 mM) were all from Selleck Chemicals. The wild-type or mutant K-RAS transfected H1299 cells were incubated with single inhibitors for 24 h, followed by protein estimation and immunoblotting for ERK activity.

Bioinformatics, microarray data analysis, and Gene Set Enrichment Analysis (GSEA)

The data for the cancer-associated RAS superfamily mutants are from two sources: TCGA harmonized mutation data from NCI Genomic Data Commons Data Portal (GDC TCGA) and the Catalog of Somatic Mutations in Cancer (COSMIC version 92) database [21]. The cases with KRAS G60 or M67 mutations were identified with publicly annotated information. The proteins from the RAS superfamily with mutation of codons equivalent to K-RAS G60 and M67 were derived from NCBI's protein blast search results using short sequences with K-RAS G60 (LDTAGQEEYS) or M67 (EYSAMRDQYM) (https://blast.ncbi.nlm.nih.gov). Microarray global gene expression data from 4 K-RAS M67L stable clones and 3 wild-type K-RAS stable clones were loaded onto GSEA website (https://www.gsea-msigdb.org/gsea/index.jsp) for analysis [22]. The pathways with upregulated genes enriched from the M67L mutant clones relative to the wild types were identified by searching the oncological signature dataset in the system. Genes enriched with normal P value < 0.05 and false discovery rate (FDR) < 0.25 were considered as statistically significant.

In vitro GTP hydrolysis activity of purified K-RAS proteins

K-RAS constructs, full length His/Kras WT (wild type), M67L, and Q61H (Sino Biological) were bound to radioactive nucleotide in 50µl of 0.1M sodium phosphate (pH 6.8), 0.5mM EDTA, sodium cholate (0.0025%), 5mM $MgCl_2$, BSA 0.5 mg/ml. Then 1µg protein was added, and incubated at 37^0C for 10 min. Labeled nucleotide (^{32}P γGTP from Perkin Elmer, ^{3}H GDP and ^{33}P γGTP from ARC) was added and incubated at 30˚C for 15 min and then at 4˚C for 15 min. The intrinsic GTPase activity was performed with 5µl protein bound to ^{32}P γGTP in 50 µl TNMN buffer [23] at 30˚C for the indicated times. Timed aliquots were subjected to Cobalt agarose immobilization at 4˚C for one hour, TNMN buffer washed, and scintillation counted. GAP and GRF stimulated activities were initiated as above, but with the addition of 1µg of RAS-GRF GEF catalytic domain (Cytoskeleton) or 1µg p120 RasGAP (LS Bio). Incubations were at room temperature for the indicated times, the aliquots were immobilized on Cobalt agarose, and treated as indicated.

All-atom molecular dynamics simulations

Molecular modeling was conducted with Maestro (Schrödinger Release 2020-2, Maestro, Schrödinger, LLC, New York, NY, 2020) using OPLS3e force field [24, 25], and all MD simulations were run using Desmond [26]. As a starting structure for both monomeric simulations, we used the N-terminal processed K-RAS structure (PDB ID: 6p0z [27], which was prepared using Protein Preparation Wizard [28]. The C-terminas was capped (adding an N-methyl amide to the terminal residue), missing side chains were added with Prime, H-bonds were optimized, and finally after removing all water molecules no H-bonds to non-waters, the system

was energy minimized (restrained with heavy atom RMSD convergence of 0.3Å). To obtain the mutant system, the methionine in position 67 was mutated to leucine in Maestro, and a similar protein preparation procedure was conducted. Both energy minimized protein-GDP complexes were solvated to a cubic box with edge distances minimun of 15Å from the protein. The TIP3P [29] model was used for water and K^+ and Cl^--ions were added to obtain a final salt concentration of 0.15M with an overall neutral net charge to the system. Simulations were run using NpT ensemble in 310 K, with 1.01325 bar, with identical settings as in [30]. For M67L mutant, 20 replicates with different seed numbers were run (each for 10 μs, resulting in total of 200 μs) and for wild type 10 replicates were run (in total of 100 μs).

For the GDP-bound K-RAS M67L:RALGDS-RBD model generation, we utilized the NMR derived RALGDS-RBD structure (PDB ID: 2rgf [31] 1st conformation) and a MSM derived structure of K-RAS M67L, for which switch I conformation was modified so that T35 side chain was repositioned to be associated with the Mg^{2+}-ion. The conformational change of switch I was conducted manually using Maestro tools and energy minimized via Protein Preparation Wizard. The RALGDS-RBD was placed on top of the obtained KRAS M67L model following its orientation observed in the H-RAS:RALGDS-RBD crystal structure (PDB ID: 1lfd [32]). For the obtained K-RAS M67L:RALGDS-RBD complex, a similar simulation setup was applied as described above, except TIP4P [29] water model was used. Initially, 2 μs relaxation simulations were run for the built complex with five replicas, from which we selected a stable output configuration and ran additional 3 μs simulations with 10 replicas (data not shown for the model generation simulations). One of the stable output configurations was selected, and a final 1 μs simulation was run for this complex to represent the final putative model of the GDP-bound K-RAS M67L:RALGDS-RBD interaction. No constraints were used in the simulations.

Trajectories were analyzed by Maestro tools and visualization of the structures was conducted with PyMOL (The PyMOL Molecular Graphics System, Version 2.5 Schrödinger, LLC.). Crystal structure visualization with their electron density was conducted with Mol* Viewer [33].

Markov state model (MSM) generation

Bayesian MSM generation was conducted with PyEMMA2 [34] For featurization, we used the backbone torsions of residues 12–13, 30–40 and 58–71, resulting in a VAMP-2 score of 2.42 [35]. Time-lagged independent component analysis (TICA) [36] was used for dimension reduction. Lag time (τ) was set as 40 ns, and two dimensions were used, k-means clustering was applied for discretization of the data to microstates (with $\sqrt{N}$ used for the no. of clusters: 447), and the microstates were assigned to three macrostates, based on timescale separation (Supplementary Figure 7), using Perron-cluster cluster analysis (PCCA++) [37]. The implied timescales were converged at the selected lag time and Chapman-Kolmogorov test suggested an acceptable model (Supplementary Figure 7).

Results

Recurrent cancer-associated mutation of *K-RAS* codons 60 and 67 and the equivalent codons in other *RAS* superfamily members

We began this analysis by studying cancer-associated mutants of *RHOA*, whose encoded N-terminus is two amino acids longer than RAS. In unpublished observations, we have examined three cancer-associated gain-of-function mutants in the switch II region of *RHOA*: *G62E*, *L69P*, and *L69R*. While the G62E protein has high levels of bound GTP in vivo, L69P and L69R have

low levels of bound GTP, similar to that of wild type RHOA protein (Qian et al, manuscript in preparation).

Given the homology between *RHOA* and other members of the superfamily, we asked whether similar mutations might be present in *RAS* and other superfamily members. Because codons G62 and L69 in *RHOA* are analogous to codons G60 and M67 in *K-RAS,* respectively, we searched the Catalog of Somatic Mutations in Cancer (COSMIC) database and the NCI Genomic Data Commons Data Portal (GDC TCGA) for cancer-associated *K-RAS* mutations that affect codons 60 and 67 as well as for analogous mutations of the equivalent codons in other *RAS* superfamily members. (Table 1 and Supplementary Table 2). For *K-RAS*, we identified four tumors with point mutations in M67: three with an *M67L* mutation (two in hematopoietic tumors, and one in a pancreatic cancer) and one with an M67I mutation (in an hematopoietic tumor). *M67I* is also present in *N-RAS*, in an endometrial cancer. In addition, mutations that involve the equivalent codon occurred two *RAP* genes (in the RAS gene family), in RAC1 (in the RHO gene family), and seven different *RAB* genes. Thus, in the two databases, the codon equivalent to M67 *K-RAS* has been mutated in more than 10 members of the *RAS* superfamily. Codon-60 mutations occurred more commonly and affected all three *RAS* genes and other GTPase families (Table 1 and Supplementary Table 2). The vast majority of cancer-associated mutations in codon-60 arose following a single point mutation. For the three *RAS* genes, the codon-60 mutations encoded a negatively charged amino acid, either aspartic acid or glutamic acid. Whether the encoded mutant amino acid was D or E depended on the specific codon that encoded the wild type G60, as single mutation of some wild type triplets encode D while others encode E.

Table 1. KRAS M67 or G60 equivalent mutations of RAS superfamily in various tumor types

Mutation equivalent	Superfamily Member	Mutation Status	Number of Cases	Tumor Types
M67	RAS	KRAS (p.M67L)	3	he, pa
		KRAS (p.M67I)	1	he
		NRAS (p.M67I)	1	en
		RAP1B (p.M67I), RAP2B (p.M67T)	2	co, he
	RAB	RAB6B (p.L78M), RAB19 (p.I82Sfs*34), RAB23 (p.I74M), RAB34 (p.I117T),RAB38 (p.M75T), RAB40C (p.I79M), RASEF/RAB45 (p.I606S)	7	bo, en, es, lu, me, re, st
	RHO	RAC1 (p.L67I)	1	co
G60	RAS	KRAS (p.G60D)	11	bl, ce, co, hn, he, st, pa
		KRAS (p.G60A/R/V and others)	11	br, co, ge, he, lu, pa
		NRAS (p.G60E)	14	he, lu, sk
		NRAS (p.G60D/R/V and others)	9	co, he, me, re
		HRAS (p.G60D/S)	4	co, st,
		RALA (p.G71E), RALB (p.G71R)	2	li, hn
		RRAS (p.G86S)	1	bc
		RHEB (p.G63E/W)	2	me, lu
		RIT2 (p.G77V/S)	2	lu
	RAB	RAB11B (p.G69D), RAB14 (p.G69A)	17	bc, br, ce, co, en, he, li, lu, me, sk
		RAB17 (p.G76D), RAB18 (p.G66R)		
		RAB1A (p.G69D), RAB21 (p.G77S)		
		RAB33B (p.G91R), RAB39A (p.G71R)		
		RAB43 (p.G76Rfs*40), RAB5A (p.G78D)		
		RAB5A (p.G78R), RAB7A (p.G66*/E)		
		RAB8B (p.G66S)		
	RAN	RAN (p.G68D)	1	bc
	RHO	RAC1 (p.G60R), RAC3 (p.G60S)	4	br, co

Abbreviations: bc = breast; bl = bladder; ce = cervix; bo = bone; co = colon; en = endometrium; he = haematopoietic and lymphoid; hn = head and neck; li = liver; lu = lung; me = melanoma; pa = pancreas; re = rectum; sk = skin; st = stomach.

K-RAS M67L has a gain-of-function phenotype but low GTP-binding

For further study, we decided to focus on the G60E *K-RAS* mutant, which was equivalent to the G62E mutant in *RHOA*, and the recurrent M67L *K-RAS* mutant. To assess biological and biochemical properties of the *K-RAS* mutants, we constructed plasmids encoding epitope-tagged (DDK) versions of *K-RAS* G60E and M67L, as well as G12D and the wild type, which served, respectively, as the positive and negative controls. The plasmids were stably expressed in two human tumor cell lines – H358, which is from non-small cell lung cancer (NSCLC), and SW620, from colon cancer – and transiently expressed in the NSCLC H1299 line. Three bioassays were performed: cell migration in H1299 (Fig. 1A), anchorage-independent growth in soft agar in H358 (Fig. 1B) and SW620 (Supplementary Figure 1B), and G418-resistant colony growth in H358 (Supplementary Figure 1C). In each assay, the G60E, M67L, and G12D mutants displayed a gain-of-function phenotype compared to the wild type.

Using extracts from H1299 cells expressing the mutants, their RAS-GTP levels were determined by an in vitro RAF-RBD binding assay, the most commonly used approach to measure this parameter (Fig. 1A). The G60E and G12D mutants displayed high GTP-binding, consistent with their increased biological activity. However, RAF1-RBD binding for M67L appeared to be even lower than for the wild type, although this mutant had gain-of-function biologically. Qualitatively similar results were obtained when the analogous RBD assay was performed with the B-RAF-RBD in extracts from the transfected H358 cells (Fig. 1B). Similarly, GST-M67L binding to endogenous RAF1 and B-RAF was weaker than that of GST-WT (Supplementary Figure 1A).

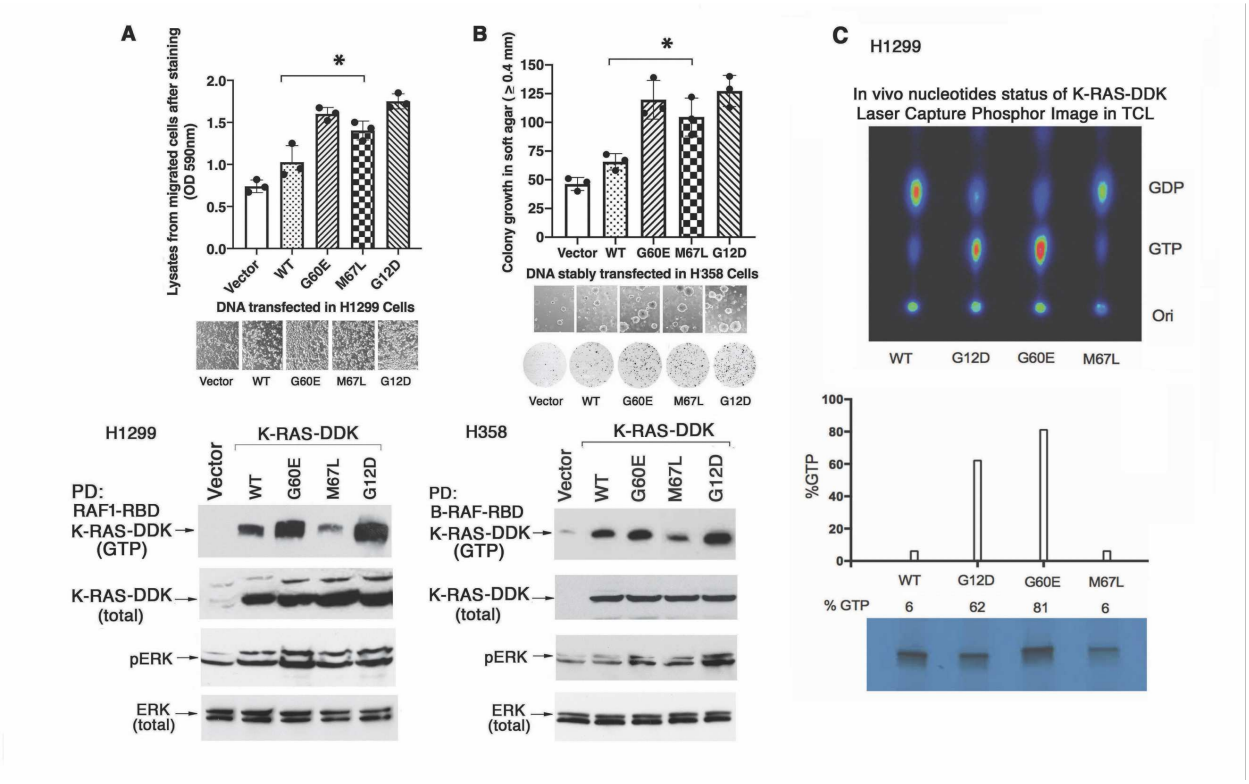

Fig. 1. *K-RAS* M67L mutant has gain-of-function biologically and its protein has low GTP-binding. **A** K-RAS M67L promotes cell migration but has reduced binding to Raf1-RBD. *K-RAS* constructs were transiently expressed in H1299 cells and subjected to transwell migration assay (top). The cell extracts of migrated cells were quantified along with representative images of migrated cells. Representative images of migrated cells are shown. RAF1-RBD pull-down and ERK activation assays were analyzed (bottom). **B** K-RAS M67L promoted anchorage-independent growth of H358 cells but reduced binding to B-RAF-RBD. Top: Stable clones of H358 cells expressing K-RAS wild type and mutants were subjected to anchorage-independent growth and the colonies were quantified and plotted as mean + SD (* p <0.05) after staining. Representative colony images under microscope and the stained whole dishes are shown (top). Cell extracts from these cells were analyzed similarly to **A**, except that binding to B-RAF-RBD was determined (bottom). **C** In vivo K-RAS GTP and GDP status in transfected H1299 cells. Top: Laser Capture Phospho images of TLC are shown (top). Bottom: The percent of K-RAS-GTP was calculated and plotted. K-RAS protein expression was also assayed.

To definitively determine the proportion of K-RAS GTP/GDP in vivo, extracts from the H1299 line expressing the mutants were metabolically labeled with ^{32}P-orthophosphate, and the nucleotide-bound K-RAS in cell extracts was subjected to thin layer chromatography to quantify the bound GTP and GDP (Fig. 1C). Similar to the RAF-RBD and B-RAF-RBD pull-down results, the G60E and G12D mutants had high GTP-binding in this assay, while the GTP-binding of M67L was as low as that of the wild type.

Given that that the M67L mutant had low GTP binding, we sought in vivo biochemical evidence to complement its observed gain-of-function biological phenotype. Consistent with this finding, M67L had increased activity for several downstream targets of K-RAS, similar to G60E and G12D (Fig. 1A and B and Supplementary Figure 1A and B). In the H1299, H358, and SW620 transfectants, this included increased pERK (Fig. 1A, 1B, Supplementary Figure 1A), despite the low binding of the M67L protein to the RAF1-RBD or the B-RAF-RBD, as well as increased CDC42 and JNK activation in SW620 (Supplementary Figure 1B) and increased CDC42, RAC1, and RHOA activation in H358 (Supplementary Figure 1C).

As an additional approach, we performed gene expression profiling and gene set enrichment analysis (GSEA) of several independently derived clones of cells stably transfected by wild type *K-RAS* and the *M67L* mutant. GSEA confirmed that expression of many growth regulatory genes differed between the two sets of transfectants (Supplementary Figure 2). Some of the regulated genes detected in the cells from the *M67L* mutant were enriched in particular pathways, including three oncologic signatures: K-RAS dependency, cell cycle, and RB1 knockout cells. Three of the genes were evaluated futher: RB and MYC, whose expression is frequently modified in RAS-transformed cells, and EPCAM, whose level of expression was several fold higher in the M67L transfectants compared with the wild type transfectants (Fig. 2).

Quantitative RT-PCR verified that in the M67L transfectants, RB mRNA expression was decreased, while MYC and EPCAM mRNA expression was increased. Immunoblotting confirmed the levels of the proteins encoded by the genes were comparable to their respective mRNA levels, and the K-RAS protein levels were similar in both original sets of transfectants (Fig. 2).

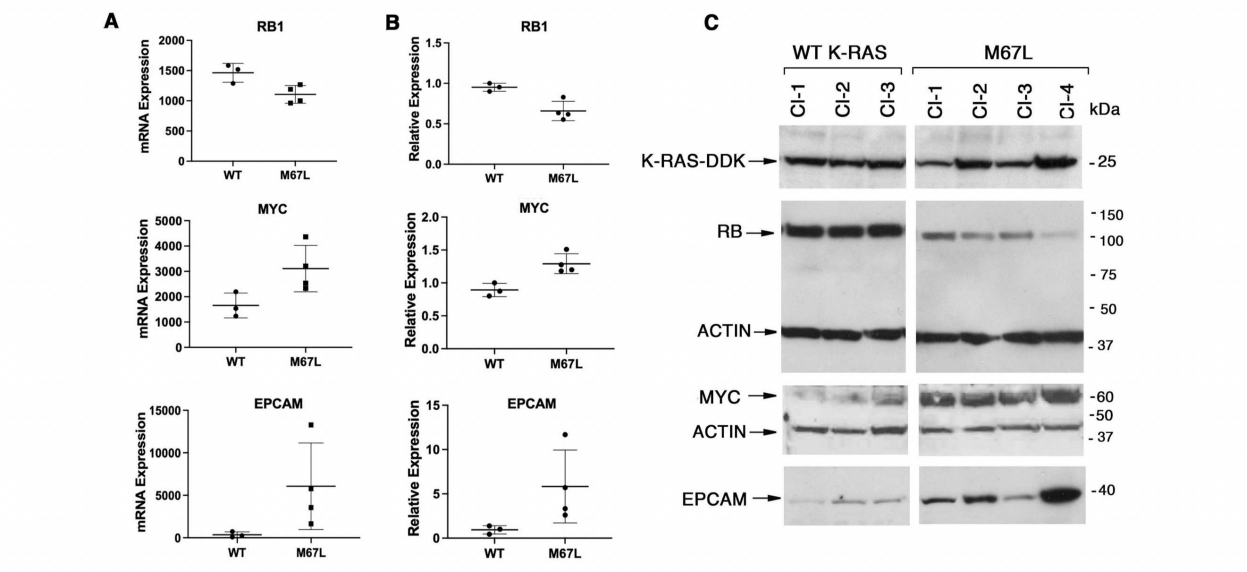

Fig. 2 Confirmation by qPCR and immunoblotting of three growth-regulatory gene expression altered by microarray analysis. **A** Microarray mRNA of RB, MYC, and EPCAM expression of four stable M67L clones vs. three stable wild type clones from transfected H358 cells was plotted as mean +/- SD (p <0.05, one tail test). **B** qPCR analysis of the three genes in stable clones as in **A**. (p <0.05, one tail test). **C** Protein levels for the three genes in stable clones were in the same direction as their respective mRNAs. Top: K-RAS protein levels for the M67L and wild type clones were similar to their respective mRNA levels.

M67L binds and activates RALGDS and PI3K

Although the M67L mutant interacted less efficiently with RAF1 and B-RAF compared with wild type, it seemed likely that the gain-of-function phenotype displayed by the M67L mutant might be mediated by its GDP-bound form being able to interact with other K-RAS effectors,

leading to downstream signaling. We therefore examined two other well-known K-RAS targets, RALGDS and PI3K. To avoid the potential in vivo ambiguity of whether binding was attributable to K-RAS-GTP- or K-RAS-GDP, wild type and mutant K-RAS proteins were expressed in bacteria, purified, and loaded in vitro exclusively with either GDP or GMPPNP (a non-hydrolyzable analog of GTP), and then reacted with purified GST-RBD fusion proteins from RAF1, RALGDS, and PI3K-p110α (Fig. 3A). In contrast to the weak binding between the RAF-RBD and M67L-GDP, the RALGDS-RBD and PI3K-p110α-RBD had readily detectable binding with M67L-GDP, whose interaction was higher than for the GDP-bound versions of the other K-RAS mutants or the wild type (Fig. 3A).

Given these in vitro binding results with the exclusively GDP- and GTP-bound proteins, we examined the in vivo status of a range of signaling molecules that are downstream from K-RAS effectors. In extracts from H1299 cells transiently transfected with the mutants, the M67L protein bound the RALGDS-RBD almost as well as the G12D mutant, and more strongly than the wild type or G60E (Fig. 3B). The RALGDS (a GEF for RAL) binding was correlated with stronger activation of RAL-GTP by M67L and G12D than G60E or wild type. For CDC42-GTP, which is downstream from RAL, the strongest activation was observed for G12D, followed by M67L and G60E, and then wild type. In addition, the three mutants bound endogenous RALGDS more efficiently than did the wild type (Fig. 3C). As noted earlier, RALGDS-associated signaling of CDC42 activation was also evaluated in extracts from SW620 (Supplementary Figure 1B) and H358 (Supplementary Figure 1C) cells stably transfected with the mutants. Compared with the wild type, the mutants in both lines were associated with higher CDC42-GTP and displayed higher levels of RAC-GTP, which lies downstream from TIAM1 and VAV and is increased by lipid messengers produced by the PI3K pathway or by activated RAP1

[38], than the wild type, while RHO-GTP was slightly higher in mutant expressing cells than in the wild type cells. For the p110α-RBD or endogenous p110α, the binding paralleled that of the RALGDS-RBD, with the M67L and G12D mutants binding more strongly than the wild type or G60E (Fig. 3D-F, top panels). To determine if this binding was associated with activation of a downstream target, we used phosphorylation of AKT (pS473) and phospho-mTOR (pmTOR S2248) as markers of PI3K activation. By these criteria, M67L activated PI3K with higher efficiency than WT K-RAS (Fig. 3F bottom panel and Supplementary Figure 3A).

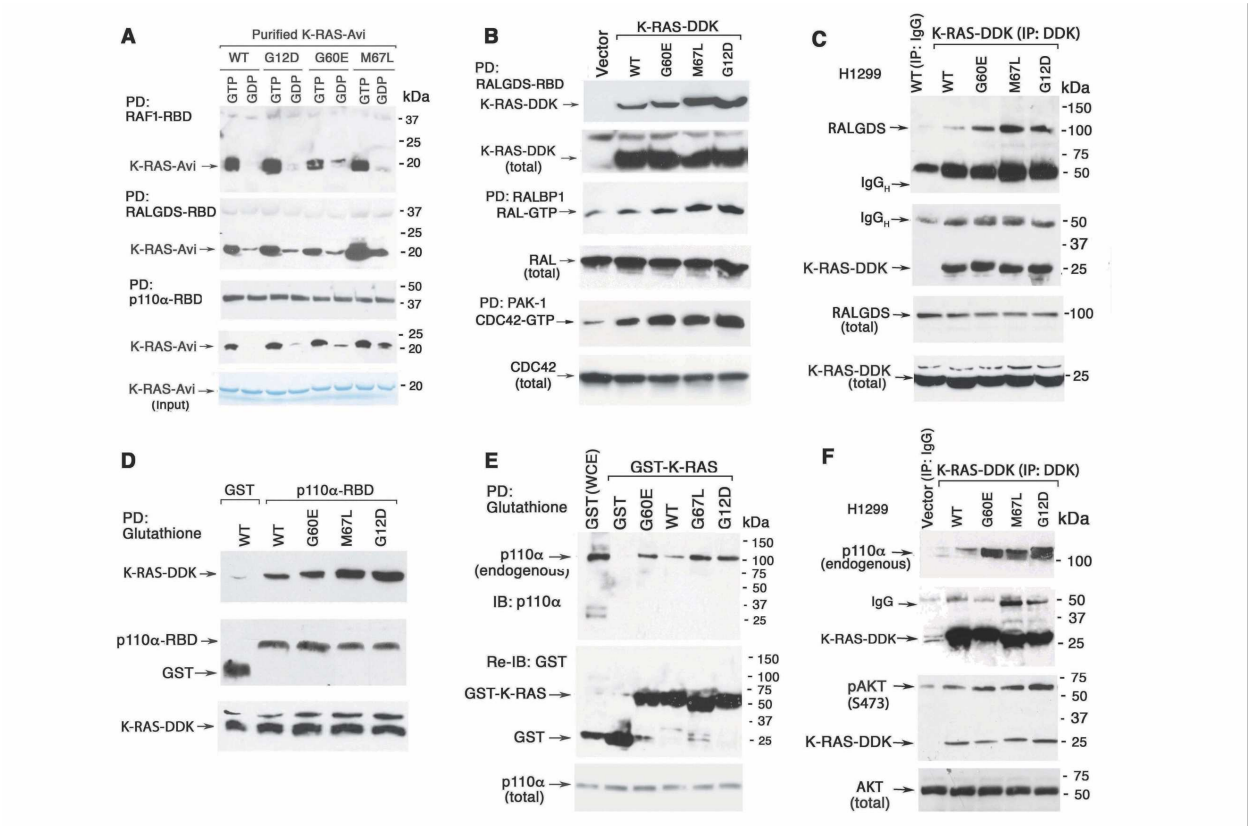

Fig. 3. M67L-GDP binds strongly to RalGDS and p110a and activates their signaling**. A** The GDP form of Avi-M67L can bind to the RBD domain of RalGDS and p110a. Purified recombinant wild type (WT) and mutant K-RAS-Avi proteins loaded exclusively with GTP analog or GDP were subjected to pull-down (PD) assay with glutathione beads bound to RAF1-RBD, RALGDS-RBD, and p110a-RBD (top rows). Equivalent amounts of K-RAS protein input are shown in the stained gel (bottom row, input). **B** M67L binds to RALGDS in vivo, which activates RAL-GTP and CDC42-GTP more efficiently than WT

K-RAS. Transfected H1299 cells were analyzed by RALGDS-RBD pull-down (PD) assay, and RAL and CDC42 activation assays as described. **C** M67L binds to endogenous RALGDS in vivo. Transfected H1299 cell extracts were immunoprecipitated (IP) with anti-RALGDS followed by immunoblotting with anti-DDK. The loading controls of expressed K-RAS and endogenous RALGDS are shown. **D** K-RAS binding to p110a- RBD in pull-down (PD) assay. The results of the pull-down by glutathioine beads are shown for the co-transfected GST-p110a-RBD and the K-RAS set in H1299 cells. **E** M67L binds endogenous p110a in vivo. The endogenous p110a was pulled-down by transfected GST-K-Ras set in H1299 cells. IB: immunoblot. **F** Complex formation of K-RAS and p110a results in activation of AKT kinase. Endogenous p110a was co-immunoprecipitated with the proteins from the transfected K-RAS set (top). The activated AKT and total AKT blots are shown (bottom). IP: Immunoprecipitation.

M67L binds and activates RAPGEF6

Given the positive results with RALGDS and PI3K, we tested a less extensively studied RAS effector, RAPGEF6 (also known as PDZGEF2 and RAGEF2), which is a RAP-specific GEF [39]. Similar to G12D, M67L bound strongly to endogenous RAPGEF6 (Fig 4A, top panels), and induced more RAP1 activation than WT K-RAS and G60E, as determined in a RALGDS pull-down assay as a measurement of RAP-GTP in H1299 cells (Fig. 4A, bottom panels). The binding mapped to the transfected RA domain of RAPGEF6 (Fig. 4B; see Supplemenary Figure 3B for domain structure of RAPGAP6). As had been true for RALGDS and PI3K, the GST-RA domain of RAPGEF6 was able to bind to purified M67L-GDP more readily detectable than to the wild type or other K-RAS mutants, although K-RAS wild type and mutants bound with the GTP analogue all bound well to the purified RA domain (Fig. 4C and Supplementary Figure 3C). RAPGEF6 made an important contribution to the activation of RAP1 by K-RAS, as siRNA knock-down of RAPGEF6 greatly reduced the level of RAP1-GTP (Supplementary Figure 3D).

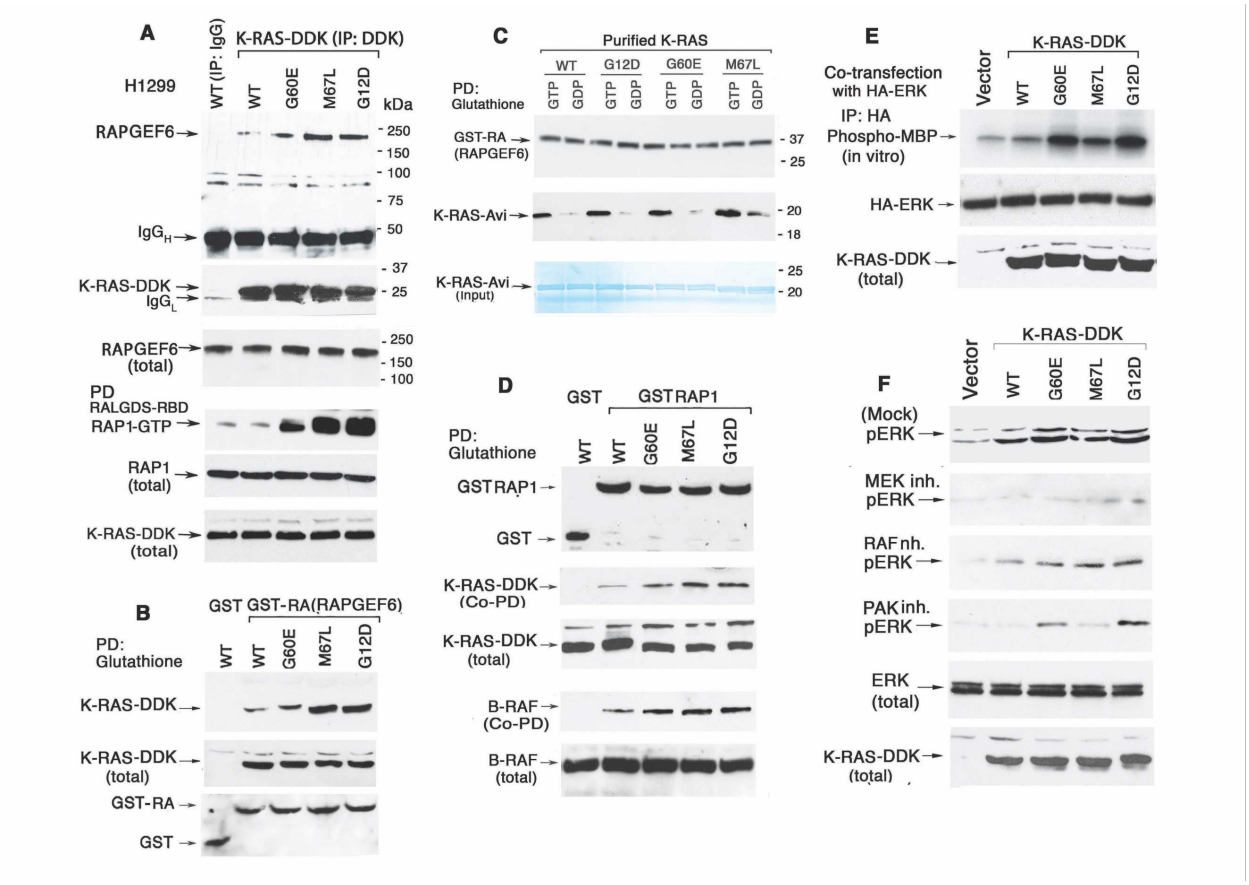

Fig. 4. M67L Induces RAP1-GTP through activation of RAPGEF6 and B-RAF and via the PAK to ERK pathway. A Association of K-RAS and RAPGEF6 results in RAP1 activation. The endogenous RAPGEF6 was co-immunoprecipitated with transiently transfected K-Ras set in H1299 cells (top panels). Cell extracts were also subjected to RALGDS pull-down assay followed by anti-RAP1 immunoblotting. The RAP1 and K-RAS loading controls are shown (bottom panels). **B** K-RAS binding to RAPGEF6-RA domain in pull-down assay. The results of the pull-down by glutathione beads are shown for the co-transfected GST- RA (RAPGEF6) and the K-RAS set in H1299 cells. **C** M67L-GDP binds to purified GST-RA. Purified recombinant wild type (WT) and mutant K-RAS-Avi proteins loaded exclusively with GTP analogue, or with GDP, were subjected pull-down assay with glutathione beads bound to RAPGEF6-RA (top rows). Equivalent amounts of K-RAS protein input are shown in the stained gel (bottom row). **D** M67L can mediate more complex formation between RAP1 and endogenous B-RAF than wild type K-RAS in K-RAS transfected H1299 cells. GST-RAP1 was pulled-down by glutathione beads and followed by immunoblotting to detect co-PD of transfected K-RAS with endogenous B-RAF. The expressed K-RAS and endogenous B-RAF protein loading controls are shown. **E** K-RAS mediated ERK activity in vitro. Extracts from H1299 cells co-transfected with HA-ERK and the K-RAS set were immunoprecipitated with HA followed by an ERK in vitro kinase assay using MBP as the substrate. The

transfected HA-ERK and K-RAS are shown. **F** M67L-induced ERK activity is sensitive to inhibitors of RAF, MEK, and PAK. The K-RAS transfected H1299 cells were treated with or without inhibitors followed by immunoblots with the indicated antibodies.

Two signaling pathways contribute to ERK activation by the M67L mutant

We wished to provide some insight into the unexpected finding that M67L activated ERK, although its binding to the RAF1-RBD or the B-RAF-RBD was weaker than wild type K-RAS. To complement the in vivo findings of ERK activation, this parameter was evaluated with an ERK in vitro kinase assay, using MBP as a substrate, for H1299 cells that had been co-transfected with HA-ERK and the *K-RAS* mutants or wild type. In this assay, the ERK activity in extracts from M67L expressing cells was higher than wild type K-RAS, but not as high as the G12D or G60E mutants (Fig. 4E).

K-RAS and RAP1 share many amino acids in switch I and have been reported to have some overlap in the downstream targets they bind and activate, such as B-RAF [40, 41]. As noted above, the M67L mutant bound the RA domain of RAPGEF6 to a similar degree as the G12D mutant, which was more than the G60E mutant, which in turn bound more than the wild type (Fig. 4B). We also evaluated whether K-RAS expression can increase the ability of GST-RAP1 in the extracts to form a complex with endogenous B-RAF, which indicated that the three K-RAS mutants, through their binding to GST-RAP1, induced more B-RAF binding than wild type K-RAS, with the potential to activate the ERK pathway [38] (Fig. 4D). Overexpression of wild type or constitutive RAP1(Q63L) can indeed activate ERK pathway as detected by ERK phosphorylation as well as its kinase activity to MBP as an in vitro substrate (Supplementary

Figure 3E), supporting the conclusion that activation of RAP1 by M67L can result in ERK activation.

RAL-GTP, which we found earlier to be activated by the M67L mutant through its activation of RALGDS (a RAL GEF), can stimulate CDC42 and its downstream target PAK1 (Fig. 3B), which is an upstream activator of ERK [42]. We therefore assessed the roles of B-RAF, MEK, and PAK in ERK activation by the M67L mutant by treating cells expressing the mutants with single inhibitors of each kinase. Each inhibitor was able to reduce ERK phosphorylation, suggesting that both B-RAF and PAK contribute to ERK activation by the M67L mutant (Fig. 4F).

The hierarchic response of the M67L mutant to GEF and GAP

We conducted in vitro studies to evaluate the intrinsic rate of GTP hydrolysis, GEF-dependent activation, and GAP-dependent inactivation of the M67L mutant, using His-tagged K-RAS proteins purified from *E. coli*. The intrinsic hydrolysis rate for M67L was similar to that of wild type K-RAS, which was much faster than that of a codon-61 activating mutant, Q61H (Fig. 5A). In vitro stimulation of GEF activity was evaluated by adding RAS-GRF. The results indicated M67L was comparable to wild type, while Q61H was not detectably stimulated (Fig. 5C). GAP sensitivity was assessed by incubating the K-RAS proteins with full-length p120 RAS-GAP. The data indicated M67L was stimulated by GAP, although the kinetics may be somewhat slower than wild type (Fig. 5B).

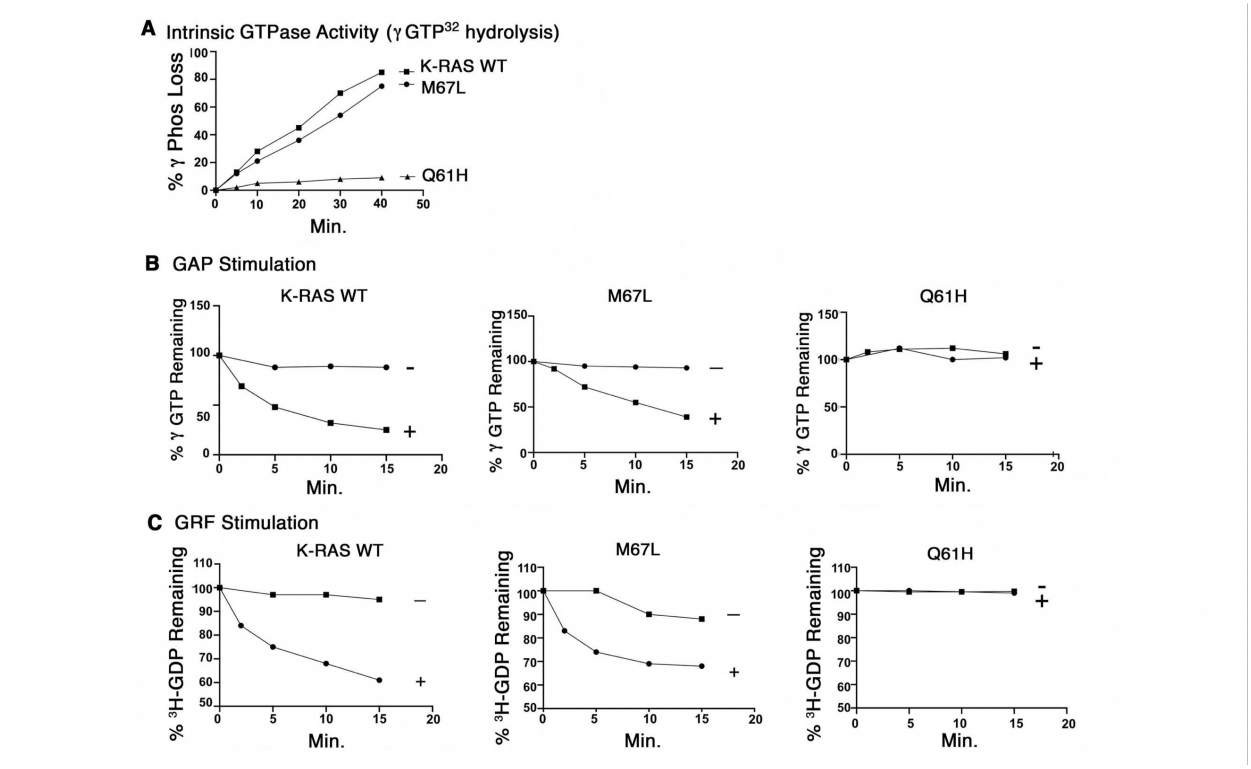

Fig. 5. In vitro intrinsic GTPase activity, GAP and GRF stimulation of wild type, M67L and Q61H. Full length N terminal His tagged and Baculovirus vector expressed K-RAS proteins were bound to ^{32}P γGTP (**A**), ^{33}P γGTP (**B**) and H-GDP (**C**). GAP hydrolysis reactions were performed with RASA1 GAP protein in **B**. GRF exchange reactions were performed with the GEF domain of RAS-GRF in **C**.

We also assessed the in vivo response of the mutants to co-transfection with a GAP, p120RASGAP, and a GEF, MyrSOS1ΔDbl, which is a constitutively activated version of SOS1 [14]. When bound GTP levels were measured by RAF1-RBD binding, the wild type, G60E, and M67L were sensitive to regulation by GEF and GAP, while G12D was resistant to them (Fig. 6). When binding to RALGDS-RBD, PI3K-p110α-RBD, or RAPGEF6-RA was used to assess these parameters, M67L was sensitive to the GEF, but appeared to be less sensitive to the GAP (Fig. 6). However, this apparent lack of GAP sensitivity is probably attributable to the relatively efficient binding of the GDP-bound M67L mutant to RALGDS-RBD and PI3K-p110α-RBD.

We also assessed the ability of a frequently used RAS monoclonal antibody, Y13-259 [43], to bind the mutants, as its binding has been mapped to specific amino acids in the switch II region [44]. The results indicated the M67L mutant was deficient for binding Y13-259, in contrast to G60E, G12D, and the wild type (Supplementary Figure 3F). The M67L result is consistent with the prior finding that an engineered M67I HRAS mutant does not bind Y13-259 and is biologically active [43].

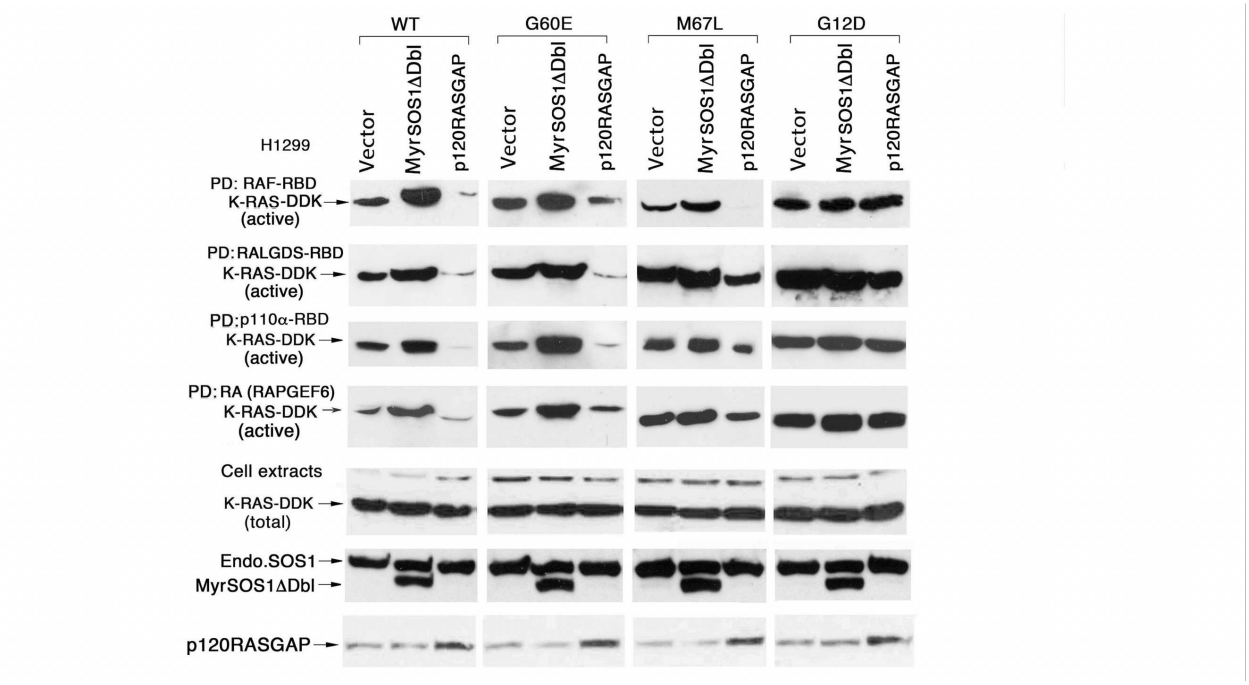

Fig. 6. The M67L mutant response to in vivo regulation by a GEF and a GAP. H1299 cells were co-transfected with wild type or mutant DDK-taggeted K-RAS together with either a constitutively active GEF (MyrSos1DDbl) or a GAP (p120RASGAP) for 48 hours, and the binding efficiency of the K-RAS proteins to the RAF1-RBD, RALGDS-RBD, p110a-RBD, and RAPGEF6-RA was determined with cell extracts in pull-down assays and anti-DDK immunoblots. Expression of K-RAS, MyrSos1DDbl and p120RAS GAP along with ACTIN is shown as controls.

The three well studied RHO GTPases – RHOA, RAC1 and CDC42 – are known to be central regulators of actin dynamics [45]. Among these, at least RhoA directly promotes stress

fiber assembly through its effectors, Rho-associated protein kinase (ROCK) and mDia [46, 47]. As the K-RAS mutants activate RHO, RAC, and CDC42, we used confocal microscopy to analyze the effect of the mutants on stress fiber architecture (Supplementary Figure 4A). By immunofluorescence, cells transfected with wild type KRAS had well-formed stress fibers (as measured by red phalloidin), unlike cells transfected with GFP vector, which had fewer stress fibers. Cells transfected with the mutants K-RAS had more stress fibers compared to wild type KRAS, especially in the center region of the cells, which is consistent with increased Rho activity (Supplementary Figure 4B).

The GDP-bound M67L mutant displays a semi-closed switch I configuration that enables RALGDS and PI3K interaction but not RAF binding

It has been well-established, until now, that only GTP-bound K-RAS, in its closed state 2 conformation [48], can efficiently bind and activate its downstream effectors. In publicly available crystal structures of RAS in complex with RBDs of RAF, RALGDS, or PI3K, RAS displays a nearly identical state 2 conformation regardless of the bound effector protein (Fig. 7A). Residue 67 of K-RAS, located in α2-helix of the switch II region, is quite distant from the main effector protein binding interface (Fig. 7A). The M67L mutation does not induce a dramatic change in the residue, as M and L are relatively similar amino acids, both nonpolar, but L is a bit shorter and branched, and is lacking a sulphur-atom (Fig. 7B). These facts make it somewhat perplexing how M67L may facilitate binding of specific effector proteins in its GDP-bound form. However, even P-loop mutations, which are distant from the effector binding interface, may affect RAS dynamics on this interface [49, 50]. Therefore, to see if the M67L would affect the conformation dynamics of GDP-bound K-RAS in a manner that would allow

specific effector protein binding, we first conducted simulations of monomeric K-RAS M67L bound to GDP (aggregate of 200 μs) and compared it to GDP-bound wild-type (aggregate of 100 μs).

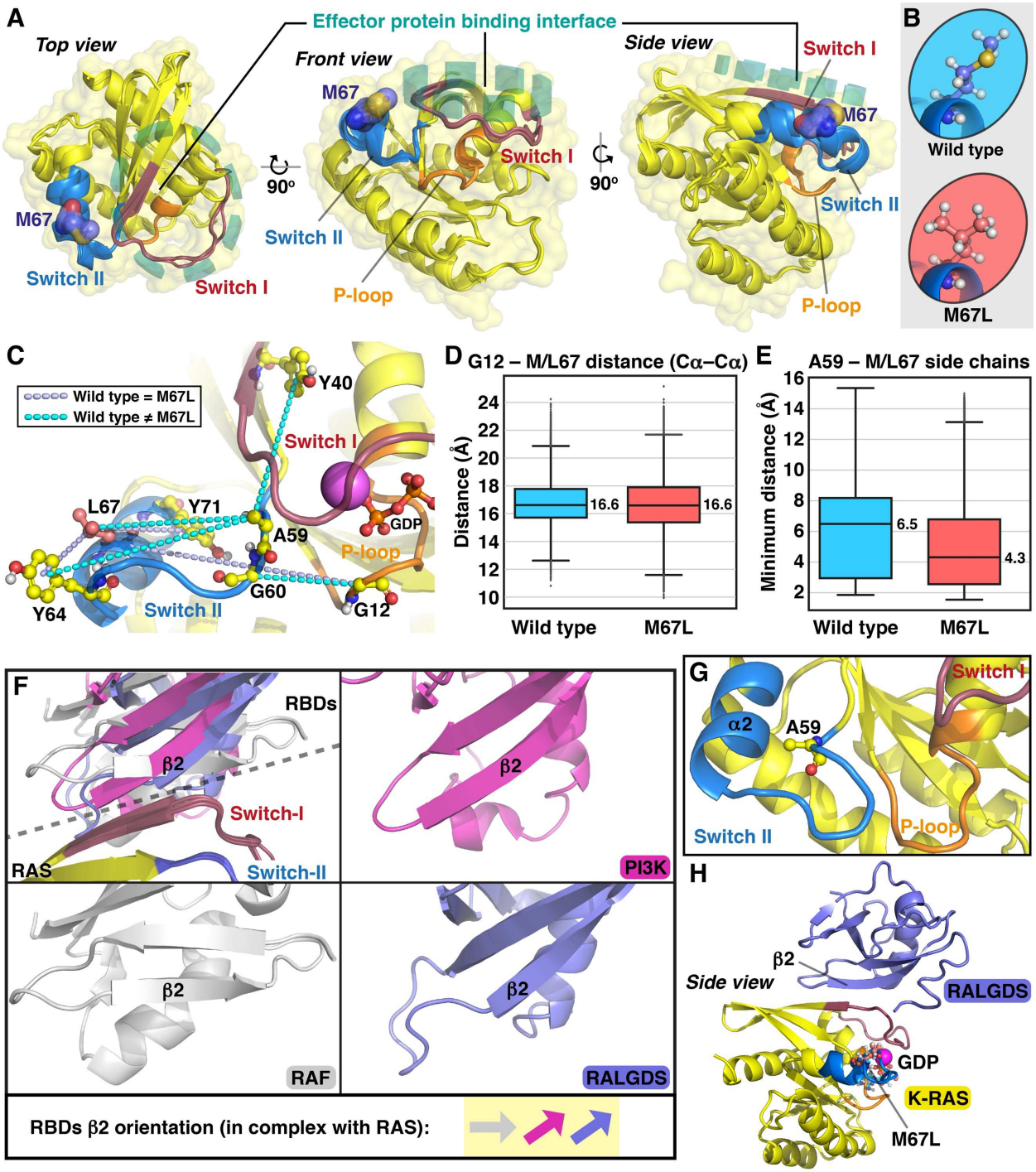

Fig. 7 A putative mechanistic explanation why GDP-bound M67L is capable of binding specific effectors. **A** Superimposed RAS conformations from the crystal structures where RAS is in complex with RALGDS, c-RAF, and PI3Kg (PDB IDs: 1he8, 1lfd, 6vjj). RAS conformation is highly similar regardless of the bound effector. M67 is located far from the main effector protein binding interface (green dashed circle). RAS is shown in yellow cartoon with transparent molecular surface, with specific regions highlighted: P-loop, orange; switch I, red; switch II, blue; M67 molecular surface. **B** Comparison of the wild type methionine (blue) and mutant leucine (red). **C** The analyzed distances of the MD simulations, plotted in **D**, **E,** and Supplementary Figure 5. Distances with light blue dashed lines appear similar between wild type and M67L, while distances with differences are shown in cyan. **D** Distance between Ca atoms of G12 and M/L67. Distance was measured every 1 ns; i.e. the data consist of 100,100 and 200,200 data points for wild type and M67L, respectively. Box displays the quartiles of the dataset (25%–75%), horizontal line in the box represents the median, and whiskers the rest of the data with maximum 1.5 IQR, outliers indicated with black diamonds. **E** Minimum distance between sidechain atoms of A59 and M/L67. **F** The effector protein differences in their RBDs b2-sheet orientation. RAF appears in more tightly bound orientation compared to the more loose RALGDS and PI3K (superimposed structures as in **A**). **G** MSM-derived conformation of M67L, where sidechain of A59 is pointing towards a2-helix. This A59 orientation is normally seen with state 2 GTP-bound K-RAS and not with GDP-bound (see Supplementrary Figure 6). **H** The binding orientation of RALGDS is compatible with GDP-bound M67L. A snapshot of the MD simulation (at 1 μs) of GDP-bound K-RAS(M67L) in complex with RALGDS is shown (see Supplementary Movies 1 and 2 for the full simulation). Stable binding of RALGDS is observed throughout the simulation.

We first focused on the putative local changes in close proximity to residue 67. Overall, the simulations suggested that M67L does not induce any major conformational rearrangement in the absolute position of residue 67, as exemplified by the identical distance between the Cα-atoms of M/L67 and G12 of the P-loop (Fig. 7C, D). The mutation also did not influence the location of two hydrophobic tyrosine residues, Y64 and Y71, that flank M/L67 (Figs. 7C; Supplementary Figure 5). Therefore, M67L does not seem to have any dramatic local effect on K-RAS conformation. We noticed, however, that M67L appears to have an effect on the

conformational behavior of the N-terminus of switch II, located on its opposite side, including residues A59 and G60 (Fig. 7C). These residues play an important role in RAS activation [51], and can also contribute to switch I dynamics [27, 52]. First, the mutant promoted configurations where the A59 side chain is pointing away from switch I and towards L67 and Y64 (Fig. 7D; Supplementary Figure 5C,E,F). Furthermore, configurations where G60 is located in close proximity to G12 were observed with increased frequency compared to the wild type (Supplementary Figure 5B). Therefore, these data suggest that M67L induces a conformational shift in A59 orientation. Interestingly, a similar A59 configuration is normally observed with GTP-bound state 2 RAS conformations, where the backbone NH group of G60 is oriented towards the γ-phosphate of the GTP, forcing A59 sidechain to orient towards α2-helix of K-RAS, away from switch I (Fig. S6A,B). Of note, in GDP-bound K-RAS crystal structures, A59 is mainly pointing towards switch I (Supplementary Figure 6C,D; Supplementary Table 3).

To confirm the relevance of this A59 configuration, we used a Markov state modeling (MSM) approach [52, 53], which aims to identify the most relevant conformations of a system, designated metastable states. M67L appears quite dynamic in all of MSM derived kinetic states, and these states displayed unclear conformational consistency, especially on A59, highlighting the highly-dynamic nature of GDP bound K-RAS (Supplementary Figure 7). However, the configuration where A59 is pointing away from switch I towards the α2-helix was found among all metastable state derived structures of M67L (Fig. 7G; Supplementary Figure 7). Although the switch II appears to mimic an active K-RAS configuration in this context, M67L conformations are obviously totally distinct from the GTP-bound state 2 RAS conformations found in the effector protein binding complexes. Importantly, GDP-bound K-RAS is never able to achieve

this fully closed switch I state 2 conformation, as it is lacking the option to form the essential H-bond(s) from the backbone of T35 and/or G60 to (the absent) γ-phosphate.

Puzzled by these findings and especially how this GDP-bound M67L conformation could offer a platform for specific effector binding, we next looked at the GTP (GNP) bound RAS–effector protein complexes and their putative differences. Surprisingly, we identified a notable difference in effector protein β-sheet interface binding orientations (Fig. 7F). While the RBD of RAF displays a more tightly bound β-sheet interface on top of RAS, RALGDS and PI3K appear in a totally different orientation, which is looser, especially above the bound GTP. Therefore, we next assessed how well-suited GDP-bound M67L could be with the different types of β-sheet interface orientations. To evaluate if this putative semi-closed complex-orientation is really compatible with these "more loose" binding effectors, we constructed a model of GDP-bound K-RAS M67L in complex with the RALGDS-RBD that was based on the MSM-derived M67L structure in its specific A59 conformation (see methods). The complex is stable in microsecond timescale MD simulation (Fig. 7E, Supplementary Figure 8; Supplementary Movie 1), thus providing a putative mechanistic explanation for the observed binding of RALGDS in biological experiments. Conversely, a clear clash with RAF, related to its more tightly bound orientation, is evident with this GDP-bound M67L K-RAS semi-closed configuration that is compatible with RALGDS (Supplementary Figure 8A). Moreover, we noted that RALGDS and PI3K have a polar side chain in the position at the end of their β2 interface close to the nucleotide that is exposed to the solvent in our model (RALGDS: S816; PI3Kα: T208; PI3Kγ: T232), while RAFs have a hydrophobic valine in this location (RAF1: V69; B-RAF: V169; A-RAF: V32) (Supplementary Figure 8B). This difference may also facilitate the tolerability of the semi-closed configuration of GDP bound KRAS M67L for specific effectors. As a concequence of

their looser binding orientation, RALGDS and PI3K may facilitate a better possibility for binding to the GDP bound mutant. Finally, as structural data are unavailable for the RA domain of RAPGEF6 (residues 749–835), we utilized AlphaFold [54, 55] to evaluate its putative configuration (reasonable confidence for RA domain prediction). The configuration of this domain follows closely to that observed for RALGDS and PI3K, and it also presents an amino acid capable of hydrogen bonding, Y764, at the end of its putative RAS-binding β-sheet interface (Supplementary Figure 9). Overall, these results provide a putative mechanistic explanation why GDP-bound M67L is capable of binding RALGDS and not RAF, where the classical tightly closed state 2 GTP-bound K-RAS conformation is a necessity (Fig. 7F, Supplementary Figure 8).

Discussion

This study has identified a cancer-associated recurrent K-RAS mutant, M67L, that has a gain-of-function phenotype although its proportion of GTP bound to the protein is similar to that of the wild type protein. Our analysis indicates that, in contrast to the wild type protein and previously described gain-of-function mutants, M67L, when bound to GDP, can interact with and activate two of the three downstream effectors that were evaluated, namely RALGDS and PI3K, but is deficient for RAF interaction. All-atom molecular dynamics simulations of M67L can account for these findings, as the simulations indicate that the mutation, induces conformational changes in the switch regions that enable the GDP-bound mutant protein to interact with RALGDS and PI3K, but not RAF. Additionally, analogous cancer-associated mutations of the equivalent codon are present in several other members of the RAS superfamily, suggesting this mechanism is likely relevant to them as well.

The M67L *K-RAS* mutation is recurrent in cancer, but occurs less commonly than many of the allele-specific mutations that affect codons 12, 13, and 61. In the COSMIC and TCGA databases, M67L (ATG>CTG) has been found in two different classes of tumors, hematopoietic and pancreatic. A similar mutation, M67I (ATG>ATA), has also occurred in K-*RAS* and *N-RAS*, in a hematopoietic tumor and endometrial cancer, respectively.

The relatively low frequency of the *M67L* mutation suggests it may be less oncogenic than more commonly mutated in *K-RAS* codons, but that it contributes to cancer in the tissue settings where it has been identified. There are ample precedents for allele-specific *K-RAS* mutants to be favored in certain tissues. For example, the frequency of specific mutant alleles of codons 12, 13, and 61 differs between pancreatic, colorectal, and lung cancer [6], and experimental studies have documented distinct tissue-specific signaling for the same mutant allele can account for the tissue-associated spectrum of specific mutant alleles [8]. In addition, some K-RAS mutants have been found to preferentially activate RAF. For example, K-RAS Q61H, which is a relatively uncommon cancer-associated mutation, activates RAF more strongly than it activates PI3K, whereas the K-RAS G12D mutant, which occurs much more commonly in cancer, activates strongly both downstream pathways [56]. Similarly, analysis of K-RAS G12R, which is relatively frequent in pancreatic cancer but less so in other cancer types commonly associated with mutant *K-RAS*, has determined that the mutant strongly activates RAF but is deficient for activating PI3K [57]. However, in pancreatic cancer, this mutation frequently occurs in association with a cooperating *PI3K* mutation, which presumably compensates for the specific “defect” in the G12R mutant [57]. Consequently, it is possible that analogous complementation may coexist with the M67L mutation.

Beyond its low percentage of steady-state GTP-binding, the M67L mutant protein displays other unusual biochemical phenotypes for a cancer-associated mutant. One is that the mutant is sensitive to the GAP-dependent hydrolysis in vitro, which contributes to its steady-state level of GTP-bound protein being similar to that of the wild type. This result contrasts with that of the vast majority of cancer-associated RAS mutants, which have high GTP-binding mainly because they are resistant to GAP-dependent negative regulation, although there are some exceptions [58]. Alternatively, the principal alteration with other mutants appears to be an increased guanine nucleotide dissociation rate, which leads to high steady-state GTP-binding because the intracellular concentration of unbound GTP is much higher than GDP [51]. However, the intrinsic dissociation rate of M67L is similar to that of wild type K-RAS, as is its sensitivity to RAS-specific GEF. Another unusual aspect of the M67L mutant is that it appears to bind weakly to RAF1 or B-RAF, yet it activates ERK. Our data indicate the M67L mutant activates B-RAF through its stronger activation of RAP1 compared with the wild type, which leads to the binding of RAP1-GTP to B-RAF [38, 40, 41].

The most unusual biochemical feature of M67L is that M67L-GDP interacts in vitro with RALGDS and with PI3K with an efficiency similar to G12D-GTP, while M67L-GDP does not interact strongly with the RAF1-RBD or B-RAF-RBD. The in vitro RALGDS results are correlated with in vivo binding of M67L to RAPGEF6 and RALGDS, and the activation of these exchange factors, as measured by an increase in their cognate downstream effectors RAP1-GTP and CDC42, respectively, to a degree similar to that seen with G12D.

The in vivo situation with RAF1 and B-RAF is more complicated for M67L, as ERK is activated in cells expressing the M67L mutant, despite the poor binding of M67L-GDP to RAF. Our analysis indicated ERK activation probably resulted from the combined effects of the ability

of GDP-bound M67L to activate RAP-dependent and PAK-dependent pathways (Fig. 8). In one pathway, RAPGEF6-mediated RAP1-GTP can bind and activate B-RAF, leading to MEK-ERK activation. In the other, RALGDS, RAL and CDC42-mediated PAK can increase ERK phosphorylation. The cooperative roles of these two pathways for ERK activation were confirmed by pharmacologic inhibition of RAF, MEK, and PAK1.

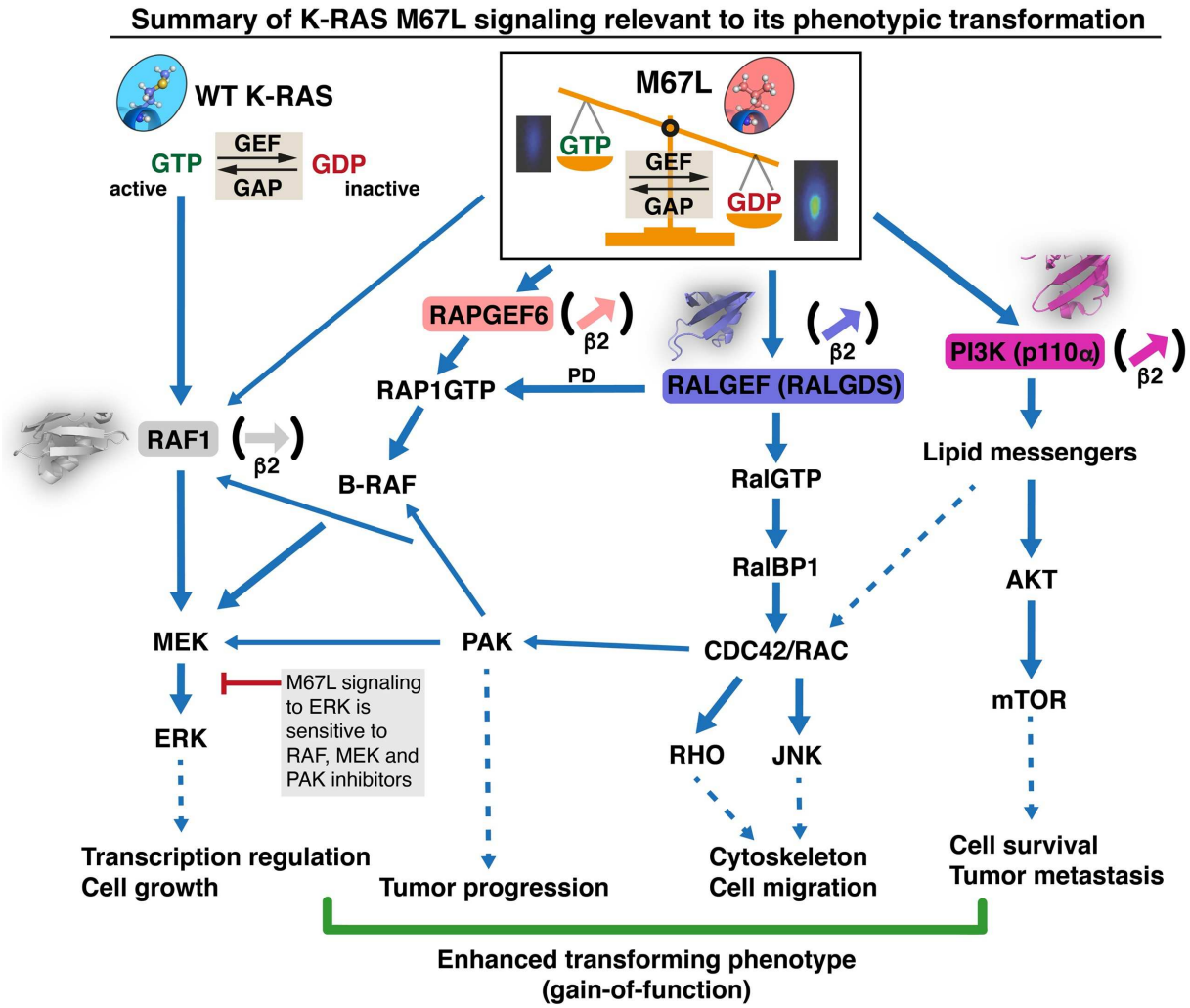

Fig. 8. Summary of K-RAS M67L signaling relevant to its phenotypic transformation. M67L, even in its GDP-bound form, binds and activates RALGDS, p110a and RAPGEF6 more efficiently than WT K-RAS, resulting in CDC42/RAC/RHO and AKT/mTOR activation, respectively. M67L also activates

more RAP1-GTP based, which promotes the activation of B-RAF, MEK, and ERK. The enhanced activation of CDC42/RAC can also activate MEK/ERK via PAK activation. Thus, the gain-of function of mutant K-RAS M67L mediates cell migration and transformation mainly by these three effectors (RALGDS, p110a and RAPGEF6 with representative b2 structures).

We used all-atom molecular dynamic (MD) simulations to understand the mechanism underlying the conformational change in switch I that results in the GDP-bound M67L protein interacting strongly with RALGDS and PI3K, and poorly with RAF. MD analysis of the M67L mutant protein by itself provided a helpful but incomplete answer. As previously described computationally and experimentally, the simulations of wild type and M67L mutant K-RAS demonstrated the highly-dynamic nature of the protein, especially in its switch regions. Specifically, the simulations suggested that the dynamics of switch II, especially related to the A59 region, is altered by the M67L mutation. The main clear difference of the mutant was the orientation of residue 59, at the N-terminus of switch II, in a manner similar to its orientation when K-RAS is GTP-bound. Thus, when the mutant is GDP-bound, residue 59 resembles a state 2 GTP-like configuration [48] more than a GDP-like one, but without a clear impact on switch I. The more intrinsically dynamic properties of switch II in GDP-bound wild type K-RAS have been experimentally assigned to residues G60–S65 [59]. As A59 is just N-terminal to G60, perhaps it is the reason why the M67L mutation, which is two amino acids C-terminal to S65, can induce changes in A59 dynamics. The side chain of A59 has an important role in regulating K-RAS dynamics, especially as related to switch I [52], with A59T mutations being present in the Harvey (H-RAS) and Kirsten (K-RAS) sarcoma viruses and possessing a gain-of-function phenotype [51, 60], and K-RAS A59T being cancer-associated [61]. These features suggest that almost any disturbance on A59 may have an effect on K-RAS behavior.

We found that the various effectors engage RAS differently, as can be seen from their distinct RBD orientations on top of RAS. Consistent with biochemical studies that have indicated wild type K-RAS binds RAF with a higher affinity than with RALGDS and PI3K [62], we observed more aligned orientation in the interaction of RAF-RBD's and RAS β-sheet interfaces than was true of this interaction interface between RAS and RALGDS or PI3K (and putatively for RASGEF6), which are associated with lower affinity. For the M67L mutant, our simulations with the model based on the specific A59 conformation found in the mutant suggested a compatible and stable binding of the effectors with the distinct interface orientation from RAF with M67L-GDP. Importantly, the tight binding RAF is incompatible with this semi-closed switch I GDP-bound M67L. However, we cannot ignore the possibility that M67L would not promote any other beneficial direct or sovent-mediated interactions to the specific effectors on top of switch II. To determine if such an additional feature exists, a more complete picture of structural information related to K-RAS–effector interactions is required than is currently available. In addition, it is possible that the lack of proper K-RAS dimerization and oligomerization might contribute to the poor RAF binding and signaling with the M67L mutant [63-66].

A mutant with some similarities to the M67L mutant has been described previously for a heterotrimeric G protein, R201G Gαs [67]. Wild type G proteins share structural similarities with members of the RAS superfamily, although the G proteins are substantially larger. While GAP activity is crucial to the regulation of wild type RAS, the wild type R201 residue in Gαs catalyzes the rapid hydrolysis of bound GTP to GDP. The R201C Gαs mutant is cancer-associated and appears to be active when GDP-bound. However, structural analyses showed that the conformation of the switch regions and nucleotide-binding pocket observed in the GDP-

bound structure of R201C Gαs is very similar to that in the structure of GTPγS-bound wild-type Gαs. By contrast, as noted above, no structural similarity in the switch I region was seen between GDP-bound K-RAS-M67L and GMPPNP-bound wild type K-RAS. This structural divergence suggests there may be substantial differences between the conformation inherently assumed by the GDP-bound R201C mutant Gαs compared with that of the GDP-bound K-RAS M67L mutant, whose switch I change was uncovered only when studied together with an effector that is permissive for stable interation with the GDP-bound mutant.

Although the frequency of M67L in cancer is low, it is noteworthy that a scan of the COSMIC database enabled identification of analogous cancer-associated mutations of the equivalent codon in 10 other members of the RAS superfamily, including N-RAS. These results strongly suggest that cancer has recurrently selected for mutation of RAS-related genes that lead to proteins with gain-of-function when GDP-bound. In addition, the possibility should now be considered that some mutations of other codons might result in a similar phenotype.

Availability of data and materials

Raw trajectories of the MD simulations are freely available at

https://doi.org/10.5281/zenodo.5506845

https://doi.org/10.5281/zenodo.5507212

https://doi.org/10.5281/zenodo.5507131

Supplementary Movie S1. MD simulation of the GDP-bound K-RAS(M67L)–RALGDS complex. The movie displays a single 1 μs simulation from two orientations (side and front orientations, see Fig. 7A for orientation definitions). Stable binding of the effector protein is observed throughout the simulation. RALGDS is shown in blue cartoon; K-RAS GDP and L67 shown in ball and stick model (uploaded separately).

References

1. Wennerberg K, Rossman KL, Der CJ. The Ras superfamily at a glance. J Cell Sci. 2005;118(Pt 5):843-6.

2. Vigil D, Cherfils J, Rossman KL, Der CJ. Ras superfamily GEFs and GAPs: validated and tractable targets for cancer therapy? Nat Rev Cancer. 2010;10(12):842-57.

3. Wittinghofer A, Vetter IR. Structure-function relationships of the G domain, a canonical switch motif. Annu Rev Biochem. 2011;80:943-71.

4. Lowy DR, Willumsen BM. Function and regulation of ras. Annu Rev Biochem. 1993;62:851-91.

5. Simanshu DK, Nissley DV, McCormick F. RAS Proteins and Their Regulators in Human Disease. Cell. 2017;170(1):17-33.

6. Stephen AG, Esposito D, Bagni RK, McCormick F. Dragging ras back in the ring. Cancer Cell. 2014;25(3):272-81.

7. Moore AR, Rosenberg SC, McCormick F, Malek S. RAS-targeted therapies: is the undruggable drugged? Nat Rev Drug Discov. 2020;19(8):533-52.

8. Poulin EJ, Bera AK, Lu J, Lin YJ, Strasser SD, Paulo JA, et al. Tissue-Specific Oncogenic Activity of KRAS(A146T). Cancer Discov. 2019;9(6):738-55.

9. Munoz-Maldonado C, Zimmer Y, Medova M. A Comparative Analysis of Individual RAS Mutations in Cancer Biology. Front Oncol. 2019;9:1088.

10. Buday L, Downward J. Many faces of Ras activation. Biochim Biophys Acta. 2008;1786(2):178-87.

11. Nussinov R, Tsai CJ, Muratcioglu S, Jang H, Gursoy A, Keskin O. Principles of K-Ras effector organization and the role of oncogenic K-Ras in cancer initiation through G1 cell cycle deregulation. Expert Rev Proteomics. 2015;12(6):669-82.

12. Hodge RG, Schaefer A, Howard SV, Der CJ. RAS and RHO family GTPase mutations in cancer: twin sons of different mothers? Crit Rev Biochem Mol Biol. 2020;55(4):386-407.

13. Qian X, Li G, Asmussen HK, Asnaghi L, Vass WC, Braverman R, et al. Oncogenic inhibition by a deleted in liver cancer gene requires cooperation between tensin binding and Rho-specific GTPase-activating protein activities. Proc Natl Acad Sci U S A. 2007;104(21):9012-7.

14. Qian X, Vass WC, Papageorge AG, Anborgh PH, Lowy DR. N terminus of Sos1 Ras exchange factor: critical roles for the Dbl and pleckstrin homology domains. Mol Cell Biol. 1998;18(2):771-8.

15. Buck CB, Pastrana DV, Lowy DR, Schiller JT. Efficient intracellular assembly of papillomaviral vectors. J Virol. 2004;78(2):751-7.

16. Wang D, Qian X, Sanchez-Solana B, Tripathi BK, Durkin ME, Lowy DR. Cancer-Associated Point Mutations in the DLC1 Tumor Suppressor and Other Rho-GAPs Occur Frequently and Are Associated with Decreased Function. Cancer Res. 2020;80(17):3568-79.

17. Qian X, Li G, Vass WC, Papageorge A, Walker RC, Asnaghi L, et al. The Tensin-3 protein, including its SH2 domain, is phosphorylated by Src and contributes to tumorigenesis and metastasis. Cancer Cell. 2009;16(3):246-58.

18. Anborgh PH, Qian X, Papageorge AG, Vass WC, DeClue JE, Lowy DR. Ras-specific exchange factor GRF: oligomerization through its Dbl homology domain and calcium-dependent activation of Raf. Mol Cell Biol. 1999;19(7):4611-22.

19. Taylor T, Denson JP, Esposito D. Optimizing Expression and Solubility of Proteins in E. coli Using Modified Media and Induction Parameters. Methods Mol Biol. 2017;1586:65-82.

20. Kopra K, Vuorinen E, Abreu-Blanco M, Wang Q, Eskonen V, Gillette W, et al. Homogeneous Dual-Parametric-Coupled Assay for Simultaneous Nucleotide Exchange and KRAS/RAF-RBD Interaction Monitoring. Anal Chem. 2020;92(7):4971-9.

21. Forbes S, Clements J, Dawson E, Bamford S, Webb T, Dogan A, et al. Cosmic 2005. British journal of cancer. 2006;94(2):318-22.

22. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. GSEA-P: a desktop application for Gene Set Enrichment Analysis. Bioinformatics. 2007;23(23):3251-3.

23. Zhang K, Papageorge AG, Lowy DR. Mechanistic aspects of signaling through Ras in NIH 3T3 cells. Science. 1992;257(5070):671-4.

24. Harder E, Damm W, Maple J, Wu C, Reboul M, Xiang JY, et al. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. J Chem Theory Comput. 2016;12(1):281-96.

25. Roos K, Wu C, Damm W, Reboul M, Stevenson JM, Lu C, et al. OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. J Chem Theory Comput. 2019;15(3):1863-74.

26. Bowers K, Chow, E, Xu, H, Dror, RO, Eastwood, MP, Gregersen, BA, Klepeis, JL, Kolossvary, I, Moraes, MA. Scalable algorithms for molecular dynamics simulations on commodity clusters. Proceedings of the 2006 ACM/IEEE Conference on Supercomputing. 2006.

27. Dharmaiah S, Tran TH, Messing S, Agamasu C, Gillette WK, Yan W, et al. Structures of N-terminally processed KRAS provide insight into the role of N-acetylation. Sci Rep. 2019;9(1):10512.

28. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des. 2013;27(3):221-34.

29. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. The Journal of Chemical Physics. 1983;79(2):926-35.

30. Pantsar T. KRAS(G12C)-AMG 510 interaction dynamics revealed by all-atom molecular dynamics simulations. Sci Rep. 2020;10(1):11992.

31. Geyer M, Herrmann C, Wohlgemuth S, Wittinghofer A, Kalbitzer HR. Structure of the Ras-binding domain of RalGEF and implications for Ras binding and signalling. Nat Struct Biol. 1997;4(9):694-9.

32. Huang L, Hofer F, Martin GS, Kim SH. Structural basis for the interaction of Ras with RalGDS. Nat Struct Biol. 1998;5(6):422-6.

33. Sehnal D, Bittrich S, Deshpande M, Svobodova R, Berka K, Bazgier V, et al. Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. Nucleic Acids Res. 2021;49(W1):W431-W7.

34. Scherer MK, Trendelkamp-Schroer B, Paul F, Perez-Hernandez 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;11(11):5525-42.

35. Wu H, Noe, F. Variational approach for learning Markov processes from time series data. arXiv:170704659v3. 2019.

36. Perez-Hernandez G, Paul F, Giorgino T, De Fabritiis G, Noe F. Identification of slow molecular order parameters for Markov model construction. J Chem Phys. 2013;139(1):015102.

37. Röblitz S, Weber M. Fuzzy spectral clustering by PCCA+: application to Markov state models and data classification. Advances in Data Analysis and Classification. 2013;7(2):147-79.

38. Zhang YL, Wang RC, Cheng K, Ring BZ, Su L. Roles of Rap1 signaling in tumor cell migration and invasion. Cancer Biol Med. 2017;14(1):90-9.

39. Ibanez Gaspar V, Catozzi S, Ternet C, Luthert PJ, Kiel C. Analysis of Ras-effector interaction competition in large intestine and colorectal cancer context. Small GTPases. 2021;12(3):209-25.

40. Hariharan IK. Ras and Rap: are former enemies now friends? Dev Cell. 2005;8(3):303-4.

41. Carey KD, Stork PJ. Nonisotopic methods for detecting activation of small G proteins. Methods Enzymol. 2002;345:383-97.

42. Bar-Sagi D, Hall A. Ras and Rho GTPases: a family reunion. Cell. 2000;103(2):227-38.

43. Furth ME, Davis LJ, Fleurdelys B, Scolnick EM. Monoclonal antibodies to the p21 products of the transforming gene of Harvey murine sarcoma virus and of the cellular ras gene family. J Virol. 1982;43(1):294-304.

44. Sigal IS, Gibbs JB, D'Alonzo JS, Scolnick EM. Identification of effector residues and a neutralizing epitope of Ha-ras-encoded p21. Proc Natl Acad Sci U S A. 1986;83(13):4725-9.

45. Clayton NS, Ridley AJ. Targeting Rho GTPase Signaling Networks in Cancer. Front Cell Dev Biol. 2020;8:222.

46. Leung T, Chen XQ, Manser E, Lim L. The p160 RhoA-binding kinase ROK alpha is a member of a kinase family and is involved in the reorganization of the cytoskeleton. Mol Cell Biol. 1996;16(10):5313-27.

47. Watanabe N, Madaule P, Reid T, Ishizaki T, Watanabe G, Kakizuka A, et al. p140mDia, a mammalian homolog of Drosophila diaphanous, is a target protein for Rho small GTPase and is a ligand for profilin. EMBO J. 1997;16(11):3044-56.

48. Milburn MV, Tong L, deVos AM, Brunger 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;247(4945):939-45.

49. Pantsar T, Rissanen S, Dauch D, Laitinen T, Vattulainen I, Poso A. Assessment of mutation probabilities of KRAS G12 missense mutants and their long-timescale dynamics by atomistic molecular simulations and Markov state modeling. PLoS Comput Biol. 2018;14(9):e1006458.

50. Smith MJ, Ikura M. Integrated RAS signaling defined by parallel NMR detection of effectors and regulators. Nat Chem Biol. 2014;10(3):223-30.

51. Feig LA, Cooper GM. Relationship among guanine nucleotide exchange, GTP hydrolysis, and transforming potential of mutated ras proteins. Mol Cell Biol. 1988;8(6):2472-8.

52. Lu M, Wang C, Chen W, Mao C, Wang J. miR-654-5p Targets GRAP to Promote Proliferation, Metastasis, and Chemoresistance of Oral Squamous Cell Carcinoma Through Ras/MAPK Signaling. DNA Cell Biol. 2018;37(4):381-8.

53. Chodera JD, Noe F. Markov state models of biomolecular conformational dynamics. Curr Opin Struct Biol. 2014;25:135-44.

54. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-9.

55. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 2022;50(D1):D439-D44.

56. Zhou ZW, Ambrogio C, Bera AK, Li Q, Li XX, Li L, et al. KRAS(Q61H) Preferentially Signals through MAPK in a RAF Dimer-Dependent Manner in Non-Small Cell Lung Cancer. Cancer Res. 2020;80(17):3719-31.

57. Hobbs GA, Baker NM, Miermont AM, Thurman RD, Pierobon M, Tran TH, et al. Atypical KRAS(G12R) Mutant Is Impaired in PI3K Signaling and Macropinocytosis in Pancreatic Cancer. Cancer Discov. 2020;10(1):104-23.

58. Rabara D, Tran TH, Dharmaiah S, Stephens RM, McCormick F, Simanshu DK, et al. KRAS G13D sensitivity to neurofibromin-mediated GTP hydrolysis. Proc Natl Acad Sci U S A. 2019;116(44):22122-31.

59. Menyhard DK, Palfy G, Orgovan Z, Vida I, Keseru GM, Perczel A. Structural impact of GTP binding on downstream KRAS signaling. Chem Sci. 2020;11(34):9272-89.

60. Velu TJ, Vass WC, Lowy DR, Tambourin PE. Harvey murine sarcoma virus: influences of coding and noncoding sequences on cell transformation in vitro and oncogenicity in vivo. J Virol. 1989;63(3):1384-92.

61. Lou E, Xiu J, Baca Y, Nelson AC, Weinberg BA, Beg MS, et al. Expression of Immuno-Oncologic Biomarkers Is Enriched in Colorectal Cancers and Other Solid Tumors Harboring the A59T Variant of KRAS. Cells. 2021;10(6).

62. Nakhaeizadeh H, Amin E, Nakhaei-Rad S, Dvorsky R, Ahmadian MR. The RAS-Effector Interface: Isoform-Specific Differences in the Effector Binding Regions. PloS one. 2016;11(12):e0167145.

63. Lee KY, Fang Z, Enomoto M, Gasmi-Seabrook G, Zheng L, Koide S, et al. Two Distinct Structures of Membrane-Associated Homodimers of GTP- and GDP-Bound KRAS4B Revealed by Paramagnetic Relaxation Enhancement. Angew Chem Int Ed Engl. 2020;59(27):11037-45.

64. Ambrogio C, Kohler J, Zhou ZW, Wang H, Paranal R, Li J, et al. KRAS Dimerization Impacts MEK Inhibitor Sensitivity and Oncogenic Activity of Mutant KRAS. Cell. 2018;172(4):857-68 e15.

65. Mysore VP, Zhou ZW, Ambrogio C, Li L, Kapp JN, Lu C, et al. A structural model of a Ras-Raf signalosome. Nat Struct Mol Biol. 2021;28(10):847-57.

66. Sarkar-Banerjee S, Sayyed-Ahmad A, Prakash P, Cho KJ, Waxham MN, Hancock JF, et al. Spatiotemporal Analysis of K-Ras Plasma Membrane Interactions Reveals Multiple High Order Homo-oligomeric Complexes. J Am Chem Soc. 2017;139(38):13466-75.

67. Hu Q, Shokat KM. Disease-Causing Mutations in the G Protein Galphas Subvert the Roles of GDP and GTP. Cell. 2018;173(5):1254-64 e11.

Key words: K-RAS mutations, Cancer-associated, gain-of-function, GDP-bound activity, RAS effector binding, RAS signaling, all-atom molecular dynamics simulations.

Abbreviations:

COSMIC: Catalog of Somatic Mutations in Cancer

GEFs: guanine nucleotide exchange factors

GAPs: GTPase-activating proteins

MD: molecular dynamics

GSEA: gene set enrichment analysis

FDR: false discovery rate

PD: pull-down

IP: Immunoprecipitation

IB: Immunoblotting

DECLARATIONS

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable

CONSENT FOR PUBLICATION

Not applicable

AVAILABILITY OF DATA AND MATERIALS

All data that support the findings of this study are available from the corresponding authors upon reasonable request.

COMPETING INTERESTS

The authors declare no competing interests**.**

FUNDING

This study was supported by the intramural research program of the Center for Cancer Research, National Cancer Institute.

AUTHORS' CONTRIBUTIONS

X.Q., T.P., and D.R.L conceived the project and designed the experiments. X.Q., T.P, A.P., B.T., and B.S-S. performed the experiments. D.E., T.W, and A.G.S. provided reagents. M.D., Y.G., D.W., and D.S. provided technical support, scientific insights, and suggestions. X.Q., T.P., D.E, A.G.S., D.W., and D.R.L wrote the manuscript. All the authors discussed and read the manuscript.

ACKNOWLEDGEMENTS

This study was supported by the intramural research program of the Center for Cancer Research, National Cancer Institute, and in part with Federal funds from the National Cancer Institute,

National Institutes of Health, under contract number HHSN261200800001E. We thank summer intern Shivani Nellore for technical assistance, the National Cancer Institute Center Genomic Core Facility for DNA sequencing, and Agilent TapeStation analysis, Fluorescent Confocal Microscopy Core Facility for image analysis, Emily Alberico, Matt Drew, Jen Mehalko, Zhaojing Meng, Simon Messing, Shelley Perkins, Lauren Procter, Nitya Ramakrishnan, Troy Taylor, and Vanessa Wall from the NCI RAS Initiative, Frederick National Laboratory for Cancer Research, for assistance in protein production and validation, and Laboratory of Molecular Technology, Advanced Technology Program, Leidos Biomedical Research, Inc. Frederick National Laboratory for Cancer Research for Microarray analysis, and Frank McComick for critical reading and helpful suggestions. The authors wish to acknowledge CSC – IT Center for Science, Finland, for computational resources. T.P. acknowledges financial support from the Orion Research Foundation sr, European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie (grant no 839230) and Academy of Finland GeneCellNano Flagship (337120).

AUTHORS' INFORMATION

The complete authors' information has been filled out in Authors Section

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFiguresS1S4.pdf
- SupplementaryFiguresS5S9.pdf
- SupplementaryMovieM1.m2v
- SupplementaryTablesS1S3.docx