PNAS
www.pnas.org

Supplementary Information for

KRAS G13D sensitivity to neurofibromin-mediated GTP hydrolysis

Dana Rabara[a,1], Timothy H. Tran[a,1], Srisathiyanarayanan Dharmaiah[a], Robert M. Stephens[a],

Frank McCormick[a,b,2], Dhirendra K. Simanshu[a,2], and Matthew Holderfield[a,c,2]

Matthew Holderfield
Email: matthew.holderfield@quantatx.com

Dhirendra K. Simanshu
Email: dhirendra.simanshu@nih.gov

Frank McCormick
Email: frank.mccormick@ucsf.edu

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.1908353116

Supplementary Information Text

Bioinformatics:

COAD TOP Mutations:
PIK3CA, SMAD4, PTCH1, CSMD3, RYR2, OBSCN, CARD11, EGFR, DNAH8, PCLO, NEB, FGFR2, LRP2, ROBO2, GRP98, KMT2D, APOB, AKAP9, PKHD1, KMT2C, CNTNAP5, PTPRK, SRCAP, BRCA2, FMN2, PREX2, MBD6, UNC13C, ZC3H13, CACNA1E, ACACB, LAMA1, KRTAP5-1, SYNE2, MDN1, VPS13B, RP1L1, SETD2, DNAH10, ANK3, ALK, TMEM132D, KCNB2, PIK3R1, UBR4, DDX60L, NYAP2, ARID1A, PIWIL1, EPHB1, NBEA, TENM3, SALL1, CDH18, ZDBF2, KRTAP10-7, EZH2, LMO7, INSR, GIMAP8, MAP3K19, ASH1L, ARHGAP32

READ TOP Mutations:
KIAA1109, SLC9C1, SMAD2, CCDC168, FGD4, CNKSR2, RPTOR, DKK2, SPATA31E1, LRGUK, DAPK1, PSD2, ITPR2, CCDC178, KCNB2, TMPRSS15, RGAG1, CDK12, F8, CMTR1, ZFYVE9, NCOR2, PAK7, RNF157, OSBPL6, CYLD, MYH1

STAD TOP Mutations:
TRIM51, FLNC, KMT2D, RIMS2, CSMD1, RPL22, PCLO, PGM5, ZBTB20, ATAD2, MDN1, LARP4B, MBD6, RIF1, RANBP2, ANK3, CACNA1E, TCERG1, ADD3, B4GALNT4, CUL9, SLC4A3, ZDBF2, ELAVL3, DOCK1, CIC, PHF2, PRPF4B, ZNF106, HDLBP, PTPN23, ARID1B, PXDNL, FCGBP, INVS, TPTE, BAI1, SPG20, LRRIQ3, ELMSAN1, FAT1, KCNA4

Supplementary Materials and Methods:

cDNA sequences and construction:

HA-NF1-Caax STD (R711-M06-370)

Protein encoded:

MGYPYDVPDYAGTTLYKKVGENLYFQGETVLADRFERLVELVTMMGDQGELPIAMALANVVPC
SQWDELARVLVTLFDSRHLLYQLLWNMFSKEVELADSMQTLFRGNSLASKIMTFCFKVYGATYL
QKLLDPLLRIVITSSDWQHVSFEVDPTRLEPSESLEENQRNLLQMTEKFFHAIISSSSEFPPQLRS
VCHCLYQVVSQRFPQNSIGAVGSAMFLRFINPAIVSPYEAGILDKKPPPRIERGLKLMSKILQSIAN
HVLFTKEEHMRPFNDFVKSNFDAARRFFLDIASDCPTSDAVNHSLSFISDGNVLALHRLLWNNQE
KIGQYLSSNRDHKAVGRRPFDKMATLLAYLGPPEHGCVIM

Construction: Human NF1 GAP domain (residues 1198-1509) was cloned downstream of a tobacco etch virus protease site (ENLYFQG) and upstream of a farnesylation signal sequence (CVIM) in a Gateway Entry clone using Gateway BP recombination (ThermoFisher) from a PCR template. The sequence validated Entry clone was recombined into pDest-370, a mammalian expression vector based on pcDNA3.1 and utilizing a strong CMV promoter (CMV51).

GFP (R931-M01-366)

Protein encoded:

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTL
TYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGI
DFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPV
LLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK

Construction: An enhanced monomeric GFP Gateway entry clone was recombined into pDest-366, a mammalian expression vector based on pcDNA3.1 and utilizing a strong CMV promoter (CMV51).

3x-FLAG-KRAS WT (R750-M67-304)

Protein encoded:

MGDYKDHDGDYKDHDIDYKDDDDKGSTTLYKKVGTMTEYKLVVVGAGGVGKSALTIQLIQNHFV
DEYDPTIEDSYRKQVVIDGETCLLDILDTAGQEEYSAMRDQYMRTGEGFLCVFAINNTKSFEDIH
HYREQIKRVKDSEDVPMVLVGNKCDLPSRTVDTKQAQDLARSYGIPFIETSAKTRQGVDDAFYTL
VREIRKHKEKMSKDGKKKKKKSKTKCVIM

Construction: Gateway multisite recombination was used to generate an expression clone containing a CMV51 promoter, 3xFLAG epitope tag, and the coding sequence for human KRAS4b (Addgene: 83129) in the backbone of pDest-304, a puromycin resistant mammalian expression vector.

3x-FLAG-KRAS G12D (R750-M03-304)

Protein encoded:

MGDYKDHDGDYKDHDIDYKDDDDKGSTTLYKKVGTMTEYKLVVVGADGVGKSALTIQLIQNHFV
DEYDPTIEDSYRKQVVIDGETCLLDILDTAGQEEYSAMRDQYMRTGEGFLCVFAINNTKSFEDIH
HYREQIKRVKDSEDVPMVLVGNKCDLPSRTVDTKQAQDLARSYGIPFIETSAKTRQGVDDAFYTL
VREIRKHKEKMSKDGKKKKKKSKTKCVIM

Construction: Gateway multisite recombination was used to generate an expression clone containing a CMV51 promoter, 3xFLAG epitope tag, and the coding sequence for human KRAS4b G12D (Addgene: 83131) in the backbone of pDest-304, a puromycin resistant mammalian expression vector.

3x-FLAG-KRAS G13D (R750-M05-304)

Protein encoded:

MGDYKDHDGDYKDHDIDYKDDDDKGSTTLYKKVGTMTEYKLVVVGAGDVGKSALTIQLIQNHFV
DEYDPTIEDSYRKQVVIDGETCLLDILDTAGQEEYSAMRDQYMRTGEGFLCVFAINNTKSFEDIH
HYREQIKRVKDSEDVPMVLVGNKCDLPSRTVDTKQAQDLARSYGIPFIETSAKTRQGVDDAFYTL
VREIRKHKEKMSKDGKKKKKKSKTKCVIM

Construction: Gateway multisite recombination was used to generate an expression clone containing a CMV51 promoter, 3xFLAG epitope tag, and the coding sequence for human KRAS4b G13D (Addgene: 83133) in the backbone of pDest-304, a puromycin resistant mammalian expression vector.

Final protein: NF1(1209-1463)
Clone: R711-X07-636 (His6-MBP-tev-NF1(1209-1463)
Protein encoded:
GVELVTMMGDQGELPIAMALANVVPCSQWDELARVLVTLFDSRHLLYQLLWNMFSKEVELADS
MQTLFRGNSLASKIMTFCFKVYGATYLQKLLDPLLRIVITSSDWQHVSFEVDPTRLEPSESLEENQ
RNLLQMTEKFFHAIISSSSEFPPQLRSVCHCLYQVVSQRFPQNSIGAVGSAMFLRFINPAIVSPYE
AGILDKKPPPRIERGLKLMSKILQSIANHVLFTKEEHMRPFNDFVKSNFDAARRFFLDIAS
Construction: Human NF1 domain (residues 1209-1463) was synthesized by ATUM (Newark, CA) as an insect optimized Gateway Entry clone with an upstream tobacco etch virus protease site (ENLYFQG). The sequence validated Entry clone was subcloned into pDest-636, a baculovirus expression vector based on pFastBac1 (Thermo Fisher) with an amino-terminal His6-MBP (maltose binding protein) fusion.

Final protein: NF1(1198-1530)
Clone: R711-X04-566 (His6-MBP-tev-NF1 (1198-1530))
Protein encoded:
GETVLADRFERLVELVTMMGDQGELPIAMALANVVPCSQWDELARVLVTLFDSRHLLYQLLWN
MFSKEVELADSMQTLFRGNSLASKIMTFCFKVYGATYLQKLLDPLLRIVITSSDWQHVSFEVDPT
RLEPSESLEENQRNLLQMTEKFFHAIISSSSEFPPQLRSVCHCLYQVVSQRFPQNSIGAVGSAMF
LRFINPAIVSPYEAGILDKKPPPRIERGLKLMSKILQSIANHVLFTKEEHMRPFNDFVKSNFDAARR
FFLDIASDCPTSDAVNHSLSFISDGNVLALHRLLWNNQEKIGQYLSSNRDHKAVGRRPFDKMATL
LAYLGPPEH
Construction: Human NF1 domain (residues 1198-1530) was synthesized by ATUM (Newark, CA) as an *E. coli* optimized Gateway Entry clone with an upstream tobacco etch virus protease site (ENLYFQG). The sequence validated Entry clone was subcloned into pDest-566 (Addgene #11517), an *E. coli* T7 expression vector with an amino-terminal His6-MBP (maltose binding protein) fusion.

Final protein: KRAS (1-169)
Clone: R989-X02-566 (His6-MBP-tev-KRAS (1-169))
Protein encoded:
GMTEYKLVVVGAGGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLLDILDTAGQEEYS
AMRDQYMRTGEGFLCVFAINNTKSFEDIHHYREQIKRVKDSEDVPMVLVGNKCDLPSRTVDTKQ
AQDLARSYGIPFIETSAKTRQGVDDAFYTLVREIRKHKEK
Construction: Human KRAS4b (residues 1-169) was synthesized by ATUM (Newark, CA) as an *E. coli* optimized Gateway Entry clone with an upstream tobacco etch virus protease site (ENLYFQG). The sequence validated Entry clone was subcloned into pDest-566 (Addgene #11517), an *E. coli* T7 expression vector with an amino-terminal His6-MBP (maltose binding protein) fusion.

Cloning, expression and purification of recombinant proteins

Gateway Entry clones for all proteins were generated by standard cloning methods and incorporate an upstream tobacco etch virus (TEV) protease cleavage site (ENLYFQG) followed by the appropriate protein coding sequences. Mutant forms of KRAS were generated using clones from the KRAS Mutant Entry clone library (Addgene) as templates. NF1 GAP contained residues (1198-1530 or 1209-1463) of isoform 2 of human NF1 (cloned from Addgene #70423), while RASA1-GAP contained residues 714-1046 of isoform 1 of human RASA1 (cloned from Addgene #70511). The catalytic domain of SOS (SOScat) contained residues 564-1048 of isoform 1 of human SOS1 (cloned from Addgene #70601). Sequence validated Entry clones were sub-cloned into pDest-566, a Gateway Destination vector containing a His6 and maltose-binding protein tag or pDest-527, a Gateway Destination vector containing a His6 protein tag to produce the final *E. coli* expression clones (1). The BL21 STAR (rne131) *E. coli* strain containing the DE3 lysogen and rare tRNAs (pRare plasmid CmR) was transformed with the expression plasmid (His6-MBP-TEV-POI, AmpR). Proteins were expressed and purified using the procedure described previously (2). Briefly, the expressed proteins of the form His6-MBP-tev-target, were purified from clarified lysates by IMAC, treated with His6-TEV protease to release the target protein, and the target protein separated from other components of the TEV protease reaction by a second round of IMAC. Proteins were further purified by gel-filtration chromatography in buffer containing 20 mM HEPES pH 7.3, 150 mM NaCl, 5 mM MgCl2 and 1 mM TCEP. The peak fractions containing pure protein were pooled, flash-frozen in liquid nitrogen and stored at -80ºC. The KRAS4b (1-169) G13D was purified in the presence of 250 μM GDP, including the final buffer, to enhance protein stability. His6-tev-SOScat was purified in a similar manner but was not digested with TEV and thus the pool from the initial IMAC was taken to the gel-filtration chromatography step (final buffer - 20 mM HEPES, pH 7.3, 150 mM NaCl, 1 mM TCEP).

Nucleotide exchange, complex formation, crystallization, and data collection

To crystallize $NF1^{GRD}$ bound to KRAS bound to GTP analog, GMPPNP, we carried out nucleotide exchange to replace GDP with GMPPNP. $NF1^{GRD}$ and KRAS-GMPPNP complexes were prepared by mixing these two proteins in a ratio of 1:1.2 and then passing the mixture through the gel-filtration column to remove unbound proteins. The elution profile from the gel-filtration column followed by SDS-PAGE analysis for both the wild-type and the G13D mutant of KRAS showed the formation of a stable protein-protein complex. Crystallization screenings were carried out using the sitting-drop vapor diffusion method using sparse matrix screens. The wild-type KRAS-$NF1^{GRD}$ complex was crystallized in 0.1 M Tris pH 8.0 and 19% PAA-co-maleic acid (Midas screen, Molecular Dimensions), and for the $KRAS^{G13D}$-$NF1^{GRD}$ complex, the crystallization condition contains 0.1 M MES pH 6.0 and 25% Pentaerythritol propoxylate (5/4 PO/OH). Crystals were harvested for data collection and cryoprotected with crystallization condition containing an additional 25% (v/v)

solution of glycerol for the wild-type complex and 25% pentaerythritol propoxylate (5/4 PO/OH) for the $KRAS^{G13D}$-$NF1^{GRD}$ complex, before being flash-cooled in liquid nitrogen. Diffraction data sets were collected on 24-ID-C/E beamlines at the Advanced Photon Source (APS), Argonne National Laboratory. Crystallographic datasets were integrated and scaled using XDS (3). The crystal parameters and the data collection statistics are summarized in Table S1.

RNA extraction, RT-PCR and qPCR

RNA extraction: Total RNA was extracted 24-hours post transfection with siRNA using the RNeasy Mini Kit (Qiagen, 74106) and the RNase Free DNase Set (Qiagen, 79254) per manufacturer's protocol and total RNA samples were frozen at -80C. Concentrations of total RNA were determined by nanodrop.

Reverse Transcription: 500 ng of total RNA was reversed transcribed in a 20 µl reaction using the High-Capacity cDNA Reverse Transcription Kit (ThermoFisher, 4368814) and the reaction was run on a Veriti 96 Well Thermal Cycler by Thermo Fisher Scientific. The cDNA was stored at -20°C for future use.

qPCR: PCR reactions were run on Quant Studio 3. Reactions consisted of 10 ng of cDNA with either SYBR Green and forward and reverse primers for human NF1 (Sigma, Primer set H_NF1_3) or TaqMan with 18S in 20 µl reactions. Samples were run in triplicate and analyzed using the comparative C_t method. Measurements were expressed as a percentage of mRNA expressed compared to scramble siRNA at 100%. Reagents were purchased from ThermoFisher.
Forward Primer Sequence: 5' - AGATCCCACAGACTGATATG - 3'
Reverse Primer Sequence: 5' - GGACACAAGATAAGGAGAATG - 3'

ACL4 Media

ATCC DMEM: F12 Medium (Cat# 30-2006) is used as the base of the medium.
The following are added to the medium at the listed final concentration:
0.02 mg/ml insulin (Sigma, I9278-5mL)
0.01 mg/ml transferrin (Sigma, T8158)
25 nM sodium selenite (Sigma, S5261-10G)
50 nM hydrocortisone (Sigma, H6909-10mL)
0.01 mM ethanolamine (2-aminoethanol) (Sigma, E0135-100mL)
0.01 mM O-phosphorylethanolamine (Sigma, P0503-5G)
100 pM 3,3',5-triiodo-L-thyronine [T3] (Sigma, T6397)
0.5% FBS (GE Life Sciences, SH30070.03)
0.5 mM sodium pyruvate (ThermoFisher, 11360070)

10 mM HEPES (in addition to HEPES already present in the basal medium) (ThermoFisher, 15630080)
extra 2mM L-glutamine (for final conc. of 4.5mM) (ThermoFisher, 25030081)
1 ng/ml epidermal growth factor (EGF) (do not filter) (Sigma E9644)
Filter media through a 0.2-micron filter and then add EGF.
Media expires in 4 weeks.

Antibodies

Vinculin (Cell Signaling 19301)
HA-tag (Cell Signaling 2367)
KRAS (Sigma/Novus Biologicals WH0003845M1/H0003845M1)
pMEK (Cell Signaling 9121)
pEGFR (Cell Signaling 3777)
NF1 (Bethyl A300-140A)
Li-Cor Secondary Antibodies (IRDye® Secondary Antibodies 800CW and 680RD)

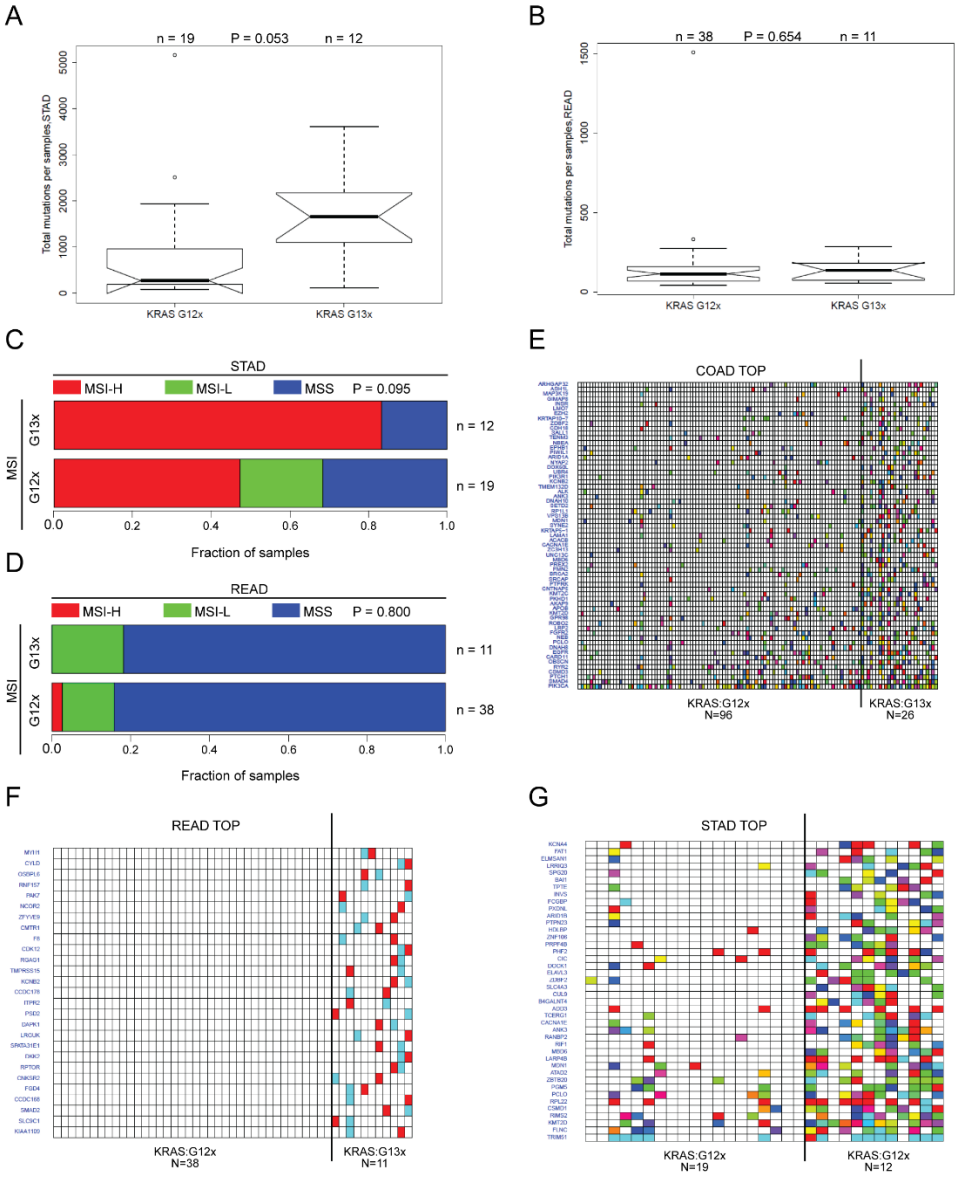


Figure S1. Co-mutations in KRAS mutated STAD and READ samples

(A) Number of Raw Mutations per sample in STAD (P = 0.053)

(B) Number of Raw Mutations per sample in READ (P = 0.654)

(C) Mutational burden comparing KRAS codon 12 and codon 13 mutations with microsatellite instability (MSI High: MSI-H; MSI Low: MSI-L; MS Stable: MSS) in TCGA STAD samples.

(D) Mutational burden comparing KRAS codon 12 and codon 13 mutations with microsatellite instability (MSI High: MSI-H; MSI Low: MSI-L; MS Stable: MSS) in TCGA READ samples.

(E) COAD TOP 100 mutations in G12x and G13x

(F) READ TOP 100 mutations in G12x and G13x

(G) STAD TOP 100 mutations in G12x and G13x

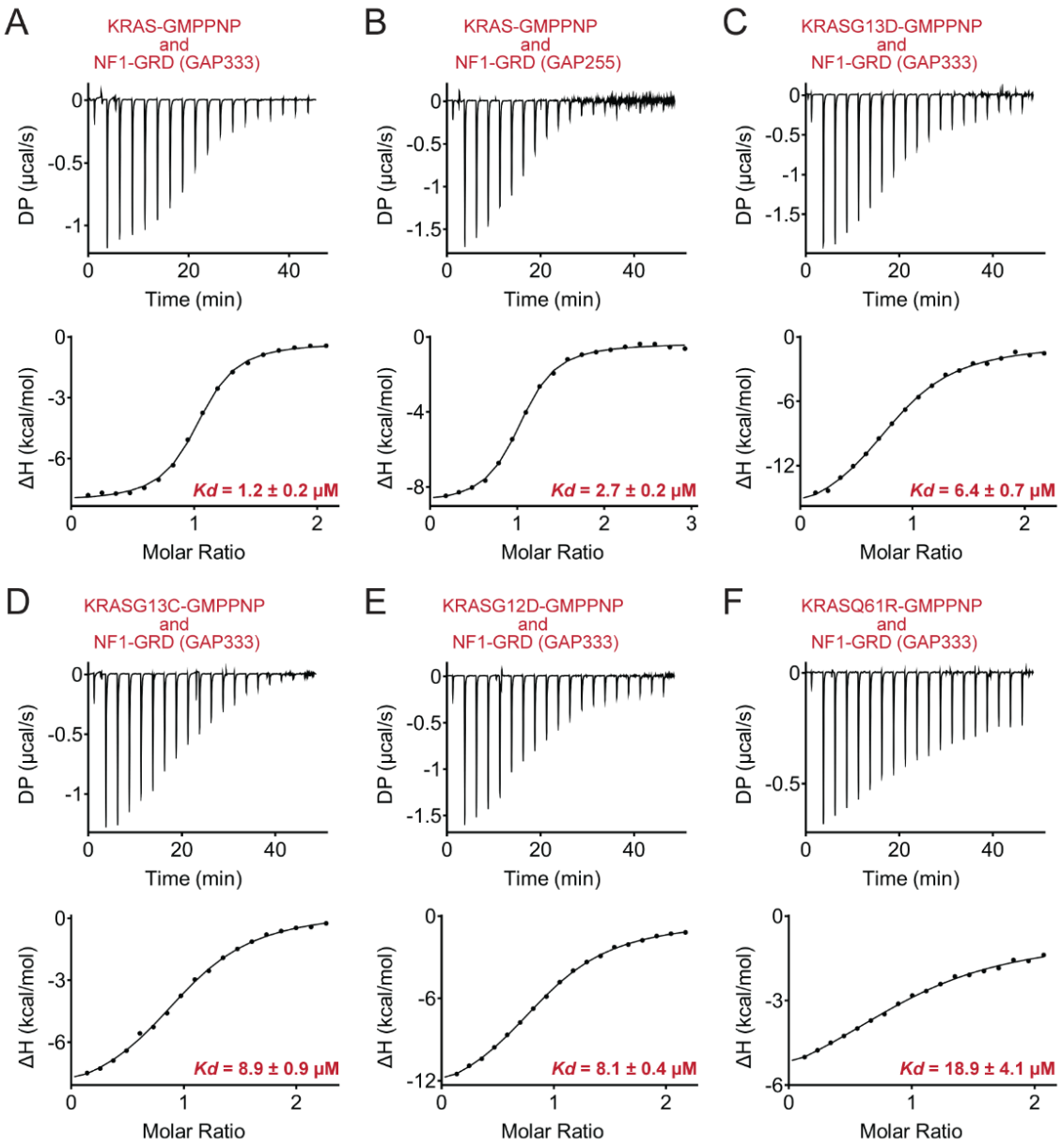


Figure S2. ITC experiments showing binding affinity of KRAS mutants for NF1^{GRD} (GAP333)

(A) GMPPNP-bound wild-type KRAS with NF1^{GRD} (GAP333)

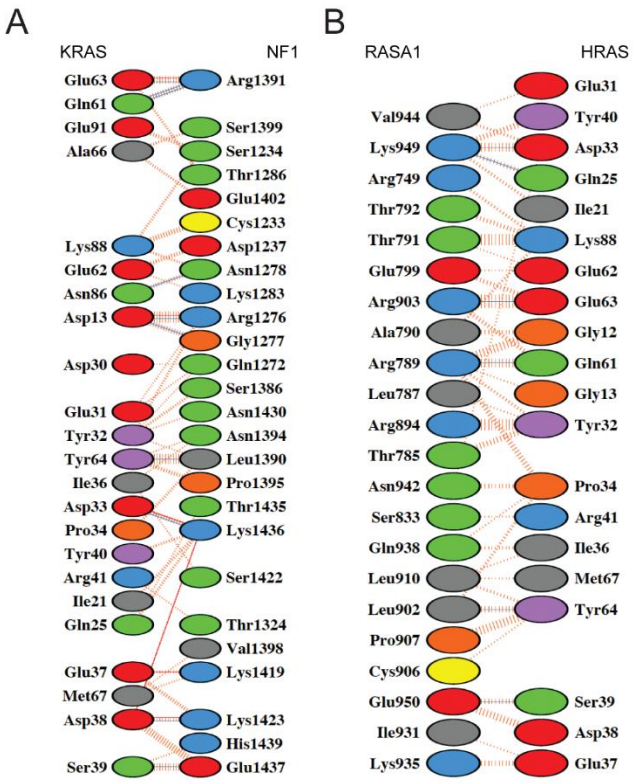
(B) GMPPNP-bound wild-type KRAS with NF1^{GRD} (GAP255)

(C) GMPPNP-bound KRAS G13D with NF1^{GRD} (GAP333)

(D) GMPPNP-bound KRAS G13C with NF1^{GRD} (GAP333)

(E) GMPPNP-bound KRAS G12D with NF1^{GRD} (GAP333)

(F) GMPPNP-bound KRAS Q61R with NF1^{GRD} (GAP333)



C Interface statistics for KRAS-NF1 complex

	No. of interface residues	Interface area (A^2)	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts
KRAS	21	1233	4	-	6	121
NF1	21	1182				

D Interface statistics for HRAS-RASA1 complex

	No. of interface residues	Interface area (A^2)	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts
RASA1	20	1522	2	-	6	123
HRAS	22	1392				

Figure S3. Protein-protein interaction at the interface of KRAS-NF1^{GRD} and HRAS-RASA1^{GRD} complexes. The figure was prepared using PDBsum server(http://www.ebi.ac.uk/pdbsum).
(A) Schematic representation of the protein-protein interaction interface of KRAS-NF1^{GRD} complex.
(B) Schematic representation of the protein-protein interaction interface of HRAS-RASA1^{GRD} complex (PDB ID: 1WQ1).
(C) Interface statistics for KRAS-NF1^{GRD} complex.
(D) Interface statistics for HRAS-RASA1^{GRD} complex.

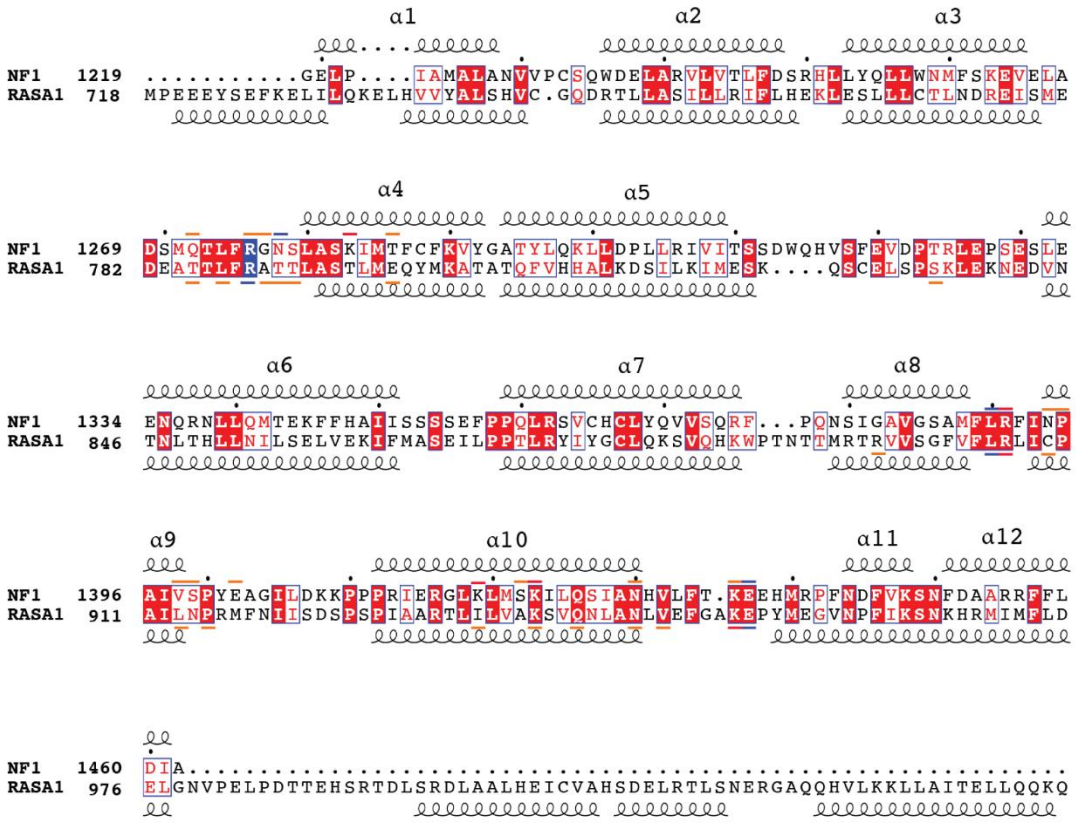


Figure S4. Structure-based sequence alignment of NF1^{GRD}(GAP255) and RASA1^{GRD}(GAP334). The orange, blue, and red color bars indicate residues forming Van der Waals/hydrophobic, hydrogen bond, and salt-bridge interactions, respectively, at the RAS-GAP interface.

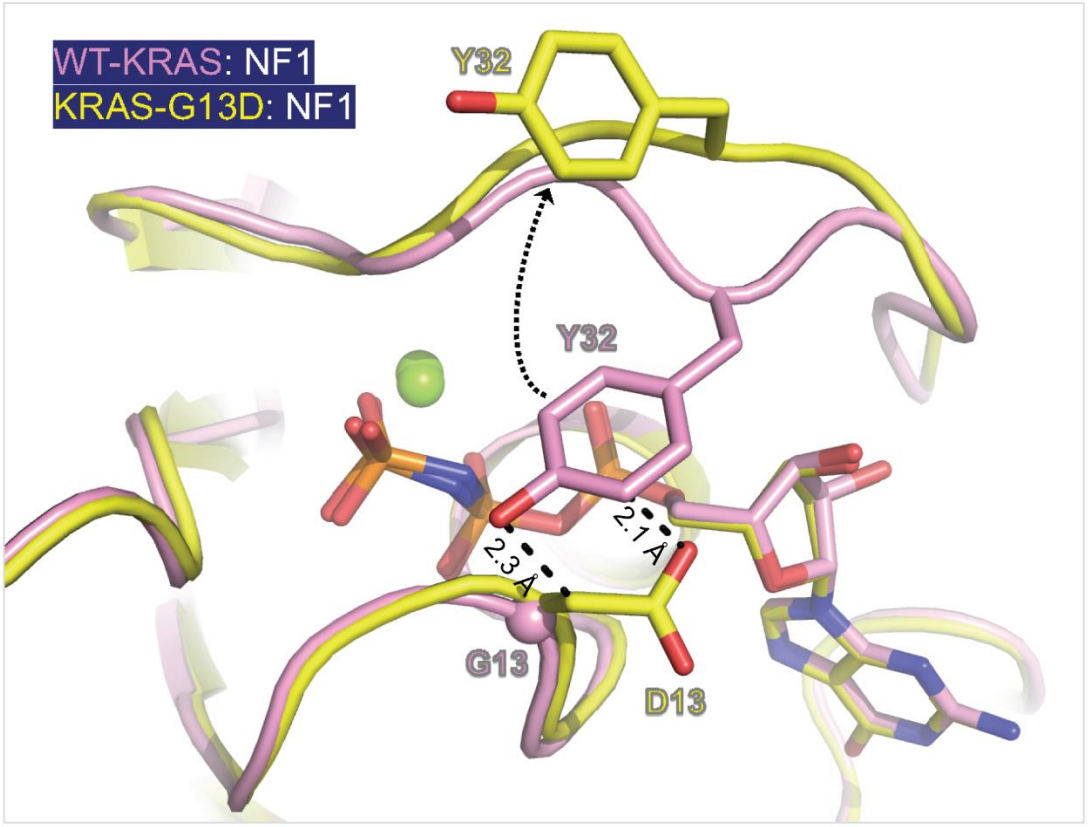


Figure S5. Structural superposition of WT KRAS and KRAS^{G13D} complexed with NF1 (GAP255) showing displacement of Y32. The protein atoms of WT KRAS and KRAS^{G13D} are colored pink and yellow, respectively. Dashed lines indicate the distance (stearic clash) between D13 and Y32 atoms and the dotted arrow shows displacement of Y32.

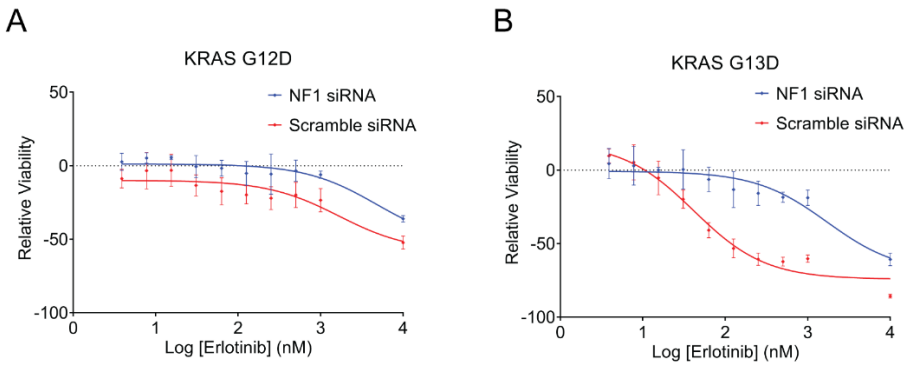


Figure S6. Erlotinib data from SW48 isogenic cell lines

(A) G12D siRNA treated shows no change in response to Erlotinib

(B) G13D shows resistance to drug when NF1 is knocked down.

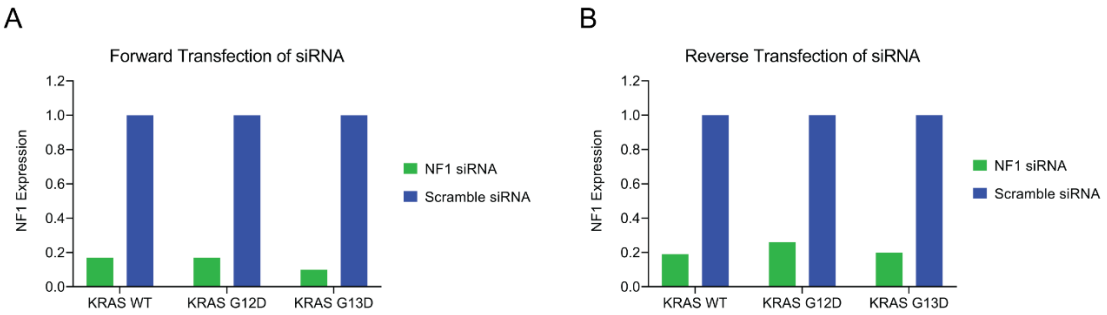


Figure S7. Confirmation of NF1 knockdown by qPCR

(A) NF1 expression levels 24 hours post siRNA forward transfection in SW48 isogenic cell lines (KRAS WT/G12D/G13D) - A control for western blot and active RAS pulldown experiments.

(B) NF1 expression levels 24 hours post siRNA reverse transfection in SW48 isogenic cell lines (KRAS WT/G12D/G13D) - A control for cell viability experiments.

Table S1. Crystallographic data collection and refinement statistics.

	Wild-type KRAS-NF1(GAP255)	$KRAS^{G13D}$-NF1(GAP255)
Data collection		
Resolution range (Å)	46.00 - 2.85 (3.02 - 2.85)	77.76 - 2.1 (2.23 - 2.10)
Space group	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Unit cell a, b, c (Å) α, β, γ (°)	 83.73, 99.74, 155.51 90, 90, 90	 81.68, 97.17, 155.53 90, 90, 90
Total reflections	197084 (29260)	350416 (56430)
Unique reflections	30699 (4720)	72838 (11552)
Multiplicity	6.4 (6.2)	4.8 (4.9)
Completeness (%)	98.10 (95.2)	99.6 (98.9)
Mean I/sigma(I)	13.88 (2.0)	17.26 (2.68)
Wilson B-factor ($Å^2$)	68.21	38.27
R-merge	0.090 (0.824)	0.052 (0.679)
R-meas	0.098 (0.899)	0.059 (0.761)
CC1/2	0.999 (0.664)	0.999 (0.891)
Refinement		
Reflections used in refinement	30688 (2848)	72751 (7161)
Reflections used for R-free	2000 (186)	1779 (174)
R-work	0.1993 (0.3193)	0.1860 (0.3225)
R-free	0.2471 (0.3522)	0.2259 (0.3926)
Number of non-H atoms	6416	7037
macromolecules	6290	6593
ligands	84	129
solvent	42	315
Protein residues	823	828
RMS - bond length (Å)	0.005	0.007
RMS - bond angle (°)	0.75	0.96
Ramachandran favored (%)	98.02	96.81
Ramachandran allowed (%)	1.98	3.19
Ramachandran outliers (%)	0.00	0.00
Average B-factor ($Å^2$)	81.56	57.75
macromolecules	81.68	57.35
ligands	78.20	71.99
solvent	65.28	60.3

Statistics for the highest-resolution shell are shown in parentheses.

Supplementary References

1. Taylor T, Denson JP, & Esposito D (2017) Optimizing Expression and Solubility of Proteins in E. coli Using Modified Media and Induction Parameters. *Methods Mol Biol* 1586:65-82.
2. Dharmaiah S*, et al.* (2016) Structural basis of recognition of farnesylated and methylated KRAS4b by PDEdelta. *Proc Natl Acad Sci U S A* 113(44):E6766-E6775.
3. Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125-132.