Supporting Information

Hunter et al. 10.1073/pnas.1404639111

SI Methods

Protein Expression and Purification. A construct encoding codon-optimized N-terminal His-tobacco etch virus (TEV)-G12C V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (K-Ras) in the pJExpress vector (DNA2.0) was synthesized and used to transform BL21(DE3) cells. Cells were grown in Luria broth (LB) to OD 600 0.7 and induced with 250 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 16 h at 16 °C. Cells were pelleted and resuspended in lysis buffer [20 mM sodium phosphate (pH 8.0), 500 mM NaCl, 10 mM imidazole, 1 mM 2-mercaptoethanol (BME), 5% (vol/vol) glycerol] containing PMSF, benzamidine, and 1 mg/mL lysozyme. Lysates were flash-frozen and stored at −80 °C until use.

Protein was purified over an IMAC cartridge (BioRad) following standard Ni-affinity protocols and desalted into 1× crystallization buffer [20 mM Hepes (pH 8.0), 150 mM NaCl, 5 mM $MgCl_2$, 0.5 mM DTT]. The N-terminal His tag was cleaved by overnight digestion with a 1:5 ratio of TEV protease at 4 °C, and the TEV and Tag were removed by reverse purification over an IMAC cartridge. Protein was concentrated to 30–40 mg/mL in a 10-kDa cutoff Amicon filter (Millipore), aliquoted, and then flash-frozen and stored under liquid nitrogen. Yields were ∼15 mg of purified K-Ras per liter of culture.

A WT expression construct was obtained by site-directed mutagenesis of the G12C construct using GenArt site-directed mutagenesis (Life Technologies) following the manufacturer's protocol. Protein was expressed and purified in the same manner as outlined for G12C K-Ras protein.

Liquid Chromatography-Electrospray Ionization-MS Analysis of Intact K-Ras G12C. Labeling and liquid chromatography-electrospray ionization-MS analysis of G12C K-Ras by SML-8-73-1 (SML) for crystallography were performed as described previously (1). Briefly, a solution of G12C K-Ras (1 mg/mL) was mixed with a 10-fold molar excess of compound and incubated at 37 °C for 2 h. The percentage of labeling was determined to be greater than 95% by intact MS analysis. Mass spectra were deconvoluted using MagTran software (version 1.03b2) (2).

X-Ray Crystallography. GDP-bound G12C and WT K-Ras protein crystals were obtained from hanging drop vapor diffusion against a well solution containing 0.2 M sodium acetate, 0.1 M Tris (pH 8.5), and 26% (wt/vol) PEG 3,350 at 20 °C. Crystals appeared in about 1 wk, and were between 100 and 200 μm in size. Crystals of SML-labeled G12C K-Ras were obtained through sitting-drop vapor diffusion against a well solution containing 0.2 M MMT buffer (Malic acid, 2-ethanesulfonic acid, Tris buffer) (pH 4.0) and 28% (wt/vol) PEG 6,000 at 4 °C with microseeding, and grew to about 10 × 20 μm in size. Crystals were cryoprotected with 10% (vol/vol) glycerol.

X-ray diffraction data were collected at the Advanced Photon Source beamline 19-ID. Images were processed, integrated, and scaled with HKL-2000/3000 packages (HKL Research Inc.) (3). Molecular replacement and model refinement were performed using Phenix and CCP4 software with 4EPV as the initial search model (4, 5). Final structures were submitted to the Research Collaboratory for Structural Bioinformatics Protein Data Bank under ID codes 4OBE, 4LDJ, and 4NMM. Collection and refinement statistics of the final structures are listed in Table S1. Images were prepared using PyMOL, version 1.5.0.4 (Schrödinger, LLC).

Chemosensor [7-Diethylamino-3-(4-Maleimidophenyl)-4-Methylcoumarin] Assay. G12C or WT K-Ras protein was diluted to 0.25 mg/mL (15 μM) in buffer containing 20 mM Tris (pH 7.5), 50 mM NaCl, and no additional magnesium. SML was added to a concentration of 150 μM, and GTP and GDP were added to a concentration of 1.5 mM each in the respective samples. Samples were incubated at 37 °C for the indicated length of time. For each time point, 27 μL of sample was mixed with 3 μL of 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin dissolved in DMSO (100 μM final concentration) in a black 384-well plate and fluorescence was read at 384/470 nm. The assay was run in triplicate, and normalized data were fit to a one-phase decay nonlinear regression curve in GraphPad Prism.

Sequence Conservation Analysis. Amino acid sequences of Ras family proteins were obtained from the National Center for Biotechnology Information Protein Data Bank (PDB) and aligned using the multiple alignment server, Clustal-Omega (v.1.2.0). These files were submitted to the ConSurf server to generate relative conservation scores for each amino acid position using the Bayesian calculation method (6). Residues with the potential to interact with SML or switch II pocket (SIIP) compound 9 (7) (within 4 Å) and that were included in calculation of the conservation score were residues 11–18 [guanine nucleotide (GN)-binding pocket 1], 28–34 (GN-binding pocket 2), 58–59 (GN-binding pocket 3), 116–120 (GN-binding pocket 4), and 145–147 (GN-binding pocket 5), as well as residues 8–12 (switch II pocket 1S), 60–63 (switch II pocket 2S), 68–72 (switch II pocket 3S), and 96–100 (switch II pocket 4S), respectively. Residues included in the calculation of the kinase ATP binding site conservation score were residues 376–385, 400–402, 432, 447–451, 455, 501, and 511 (based on the structure sequence of PDB ID code 4FL1).

To color-code by conservation score, the ConSurf algorithm binned the continuous conservation scores into nine discrete groups ranging from highly conserved (9) to variable (1). Images of 3D structures color-coded by binned conservation score were generated using the PyMOL Molecular Graphics System. A Ras family dendrogram was generated using the Interactive Tree of Life (iTOL) online phylogenetic display tool (8, 9) with the Clustal-Omega–generated alignment and phylogenetic tree files as inputs.

SML MS-Based Chemical Profiling. MIA PaCa-2 cells were cultured in DMEM with 10% (vol/vol) FBS and 0.5% penicillin/streptomycin, harvested, and flash-frozen as a cell pellet. Cells were lysed, treated, and analyzed by MS according to previously reported methods at ActivX Biosciences, Inc., with minor modifications (10, 11, 12). Briefly, the cell pellets were lysed by sonication in lysis buffer [50 mM Hepes (pH 7.5), 150 mM NaCl, 0.1% Triton X-100, phosphatase inhibitors (mixture II AG Scientific no. P-1518)]. After lysis, the samples were cleared by centrifugation, the supernatant was gel-filtered using the BioRad 10DG columns, and the eluate was collected. The protein concentration was determined using a BCA Protein Assay Kit (Thermo Scientific), and adjusted to 5 mg/mL.

Lysates were then treated with compound (100 μM) or an equivalent volume of DMSO for 15 min. EDTA was added to a final concentration of 5 mM, followed by a desthiobiotin-GTP-acylphosphate probe (GTP probe) to a final concentration of 20 μM. One minute after addition of the probe, $MgCl_2$ was added to a final concentration of 20 mM, and the samples were incubated for an additional 10 min. Duplicate control and treated sample lysates were used for statistical analysis. The percentage of change in the MS signal was reported as statistically significant if the Student t test score was <0.04.

1. Lim SM, et al. (2014) Therapeutic targeting of oncogenic K-Ras by a covalent catalytic site inhibitor. *Angew Chem Int Ed Engl* 53(1):199–204.
2. Zhang Z, Marshall AG (1998) A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. *J Am Soc Mass Spectrom* 9(3):225–233.
3. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
4. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
5. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235–242.
6. Landau M, et al. (2005) ConSurf 2005: The projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* 33(Web Server issue): W299–W302.
7. Ostrem JM, Peters U, Sos ML, Wells JA, Shokat KM (2013) K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* 503(7477):548–551.
8. Letunic I, Bork P (2011) Interactive Tree Of Life v2: Online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* 39(Web Server issue):W475–W478.
9. Letunic I, Bork P (2007) Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics* 23(1):127–128.
10. Patricelli MP, et al. (2007) Functional interrogation of the kinome using nucleotide acyl phosphates. *Biochemistry* 46(2):350–358.
11. Patricelli MP, et al. (2011) In situ kinase profiling reveals functionally relevant properties of native kinases. *Chem Biol* 18(6):699–710.
12. Okerberg ES, et al. (2005) High-resolution functional proteomics by active-site peptide profiling. *Proc Natl Acad Sci USA* 102(14):4996–5001.

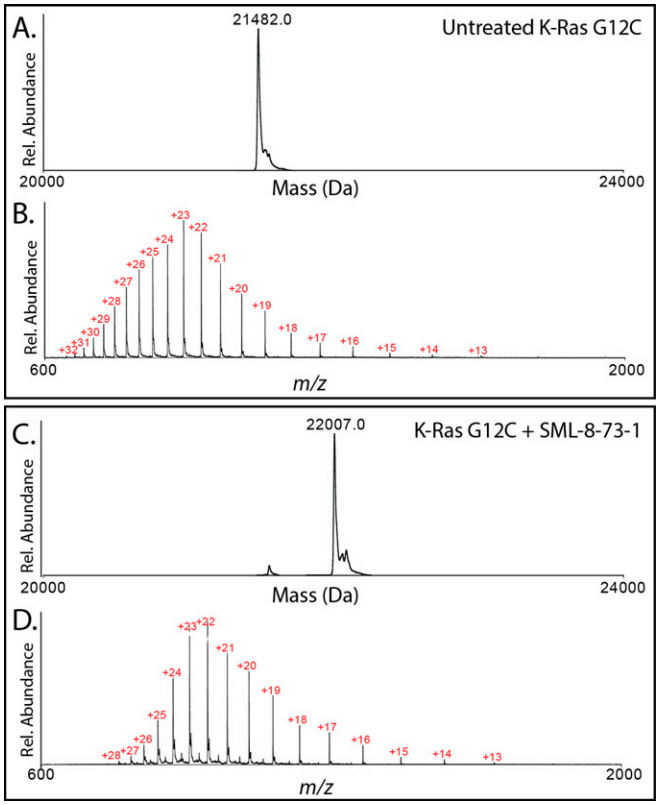

Fig. S1. Near-complete K-Ras G12C labeling by SML. Zero-charge mass spectra (*A* and *C*) and mass spectra (*B* and *D*) derived from analysis of untreated K-Ras G12C (*A* and *B*) or K-Ras G12C treated for 2 h with a 10-fold excess of SML (*C* and *D*). Treatment produces a shift in mass consistent with covalent labeling by SML (compound-HCl). Rel., relative.

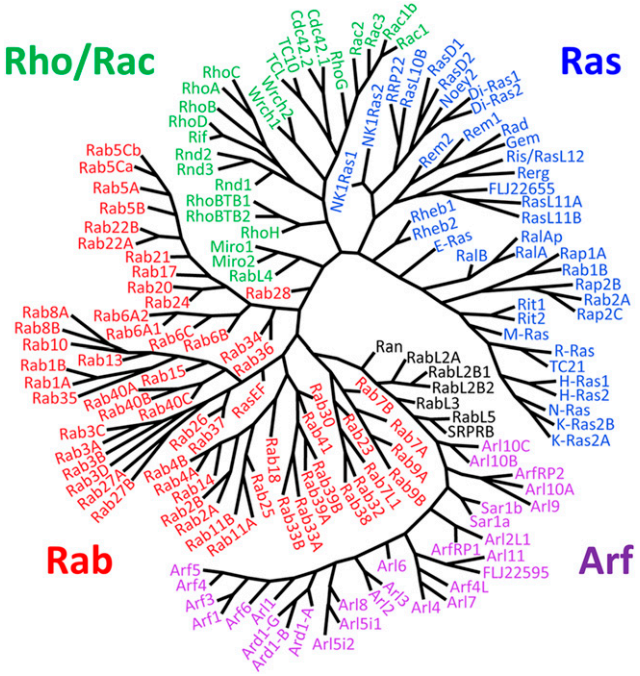

Fig. S2. Ras family dendrogram. The Ras superfamily consists of 160 members, which can be subdivided into the Ras, ADP ribosylation factor (Arf), Ras-related proteins in brain (Rab), Ras-homologous (Rho/Rac), and Ras-related nuclear protein (Ran) subfamilies. Clustering analysis was performed on amino acid sequences obtained from the National Center for Biotechnology Information PDB using the Clustal-Omega multiple sequence alignment tool.

Table S1. Data collection and refinement statistics

	WT (PDB ID code 4OBE)	G12C (PDB ID code 4LDJ)	SML (PDB ID code 4NMM)
Data collection			
Source	APS 19-1D	APS 19-1D	APS 19-1D
Wavelength, Å	0.97926	0.97918	0.97921
Space group	C2	$P2_12_12_1$	$P2_12_12_1$
Unit cell			
a, b, c; Å	66.2, 42.1, 114.4	39.1, 40.7, 91.4	39.1, 42.0, 91.2
α, β, γ; °	90, 105.3, 90	90, 90, 90	90, 90, 90
Resolution, Å	50–1.24	50–1.15	50–1.89
Unique reflections	81,203	52,661	12,363
Redundancy*	4.3 (3.2)	7.0 (3.0)	5.8 (5.1)
Completeness, %	100 (94.1)	99.2 (94.8)	99.8 (67.8)
R-merge	0.049 (0.395)	0.075 (0.317)	0.126 (0.883)
<I/σ>	33.2 (2.07)	9.6 (2.48)	14.47 (2.04)
Wilson B-factor	11.5	9.66	23.3
Refinement			
Resolution	22.1–1.24	20.7–1.15	45.6–1.89
Reflections used	77,093	50,614	11,695
Reflections for R-free	4074	1977	615
Nonhydrogen atoms	3,093	1,666	1,482
Protein	2,680	1,386	1,357
Water	355	280	125
R-work	0.157	0.136	0.184
R-free	0.169	0.158	0.244
rmsd			
Bond lengths, Å	0.011	0.015	0.016
Bond angles, °	1.3	1.7	1.4
Average B-factor	17.7	12.9	26.0
Ramachandran plot, %			
Favored/allowed/disallowed	98.8/1.2/0.0	94.2/5.8/0.0	99.4/0.6/0.0
MolProbity score	0.82 (100%)	0.82 (100%)	1.04 (100%)

APS, Advanced Photon Source.
*Numbers in parentheses correspond to the last resolution shell.

Table S2. Complete chemical proteomics profiling results

Family	GTPase	Labeling site	Probe inhibition, %
Small GTPase, Rho	CDC42	GTPase domain	16.6
	CDC42	GTPase domain	−0.7
	CDC42	GTP P site	−1.1
	RhoA, RhoC	GTPase domain	3.7
	RhoA, RhoC	GTP P site	−30.1
	RHOG	GTP P site	20.6
	RHOG	GTPase domain	12.5
Small GTPase, Ras	HRAS, KRAS, NRAS	GTP P site	−20.6
	KRAS G12C	GTP P site	82.9
	RABL4	GTP P site	4.9
	RAP1A, RAP1B	GTP P site	2.7
	RAP1A, RAP1B	GTPase domain	−2.4
	RAP1A, RAP1B	GTP P site	−6.9
	RAP2B	GTP P site	5.4
Small GTPase, Ran	RAN	GTPase domain	8.7
	RAN	GTPase domain	2.5
	RAN	GTP P site	−5.2
Small GTPase, Rab	RAB11A, RAB11B	GTP P site	−63.6
	RAB13	GTP P site	−38.3
	RAB14	GTP P site	6.3
	RAB18	GTPase domain	11.4
	RAB18	GTP P site	−20.1
	RAB1A, RAB1B, RAB1C	GTP P site	−108.7
	RAB1A, RAB1B, RAB1C, RAB3A, RAB3C, RAB3D, RAB8A, RAB8B, RAB9B, RAB10, RAB14	GTPase domain	−0.9
	RAB21	GTP P site	11.5
	RAB22A	GTP P site	19.3
	RAB27A	GTP P site	−49.9
	RAB27B	GTP P site	20.9
	RAB28	GTP P site	14
	RAB2A, RAB2B	GTP P site	−37
	RAB31	GTP P site	−13.5
	RAB32	GTP P site	20.4
	RAB33B	GTPase domain	−1.1
	RAB33B	GTP P site	−26.1
	RAB35	GTP P site	2.4
	RAB3A	GTP P site	−24.1
	RAB3B, RAB3C	GTP P site	2.4
	RAB3D	GTP P site	−4.3
	RAB43	GTP P site	−19.7
	RAB4A	GTP P site	6.4
	RAB4B	GTP P site	−38.1
	RAB5A, RAB5B, RAB5C	GTP P site	−11.5
	RAB6A, RAB6B	GTP P site	16.8
	RAB7A	GTP P site	−2.8
	RAB7A	GTPase domain	−25.4
	RAB7L	GTP P site	−14.7
	RAB8A, RAB10	GTP P site	30
	RAB8B	GTP P site	−5.1
	RABL2A, RABL2B	GTP P site	5.6
	RABL3	GTP P site	−18.2
Small GTPase, Arf	ARF1, ARF3, ARF4, ARF5	GTP P site	3.5
	ARF6	GTP P site	−14
	ARL1	GTP P site	21.1
	ARL11	GTPase domain	20.1
	ARL2	GTP P site	0
	ARL3	GTP P site	65.1
Small GTPase, SAR1	SAR1A, SAR1B	GTP P site	−51.6
MMR1/HSR1 GTP binding	NOG2	GTP P site	15.4
	NOG2	GTPase domain	−4.3
Mitochondrial Rho GTPase	RHOT2	GTP P site domain 2	0.2
IF-2	IF2P	GTP P site	27.9
GTPBP1 GTP-binding protein	GTPBP1	GTP P site	12.6
	GTPBP2	GTP P site	−4.5
GTP-binding elongation factor	EEF1A1, EEF1A2, EEF1A3	GTP P site	15.3

Table S2. Cont.

Family	GTPase	Labeling site	Probe inhibition, %
	EFTUD1	GTP P site	73.3
	EIF2S3, EIF2S3L	GTP P site	–7
	GSPT1	GTP P site	18.9
	GSPT2	GTP P site	6.9
	GUF1	GTP P site	72.6
	HBS1L	GTP P site	20.9
	SELB	GTP P site	10.1
GTP1/OBG	DRG1	GTP P site	25.7
	DRG1	GTPase domain	17.5
	GTPBP10	GTP P site	1.6
G-alpha	GNA11, GNA14, GNAQ	GTP P site	–14.8
	GNAI1, GNAI2, GNAI3, GNAL, GNAO, GNAS1, GNAS2, GNAT1, GNAT2, GNAT3	GTP P site	–23.9
	GNAI2-like	GTP P site	–23
Eukaryotic initiation factor 4E	eIF4	7mG mRNA cap binding	3.1
	eIF4	7mG mRNA cap binding	–3.5
Era/mnmE GTP-binding protein	ERAL1	GTP P site	–12.9
	GTPBP3	GTP P site	1.7
eIF-2B/eIF-5	EIF5	GTPase domain	13.6
Dynamin	OPA-1	GTP P site	15.8
DNA-dependent protein kinase catalytic subunit	DNAPK	ATP	–5.4
	DNAPK	ATP	–10.6
WD repeat G protein beta	GNB2L1	G beta	6.1
	GNB2L1	G beta	–11.8

GTP-binding proteins that exhibited statistically significant inhibition of GDP-desthiobiotin labeling by SML ($P < 0.04$ by Student t test) are highlighted in orange or red. All others (green) show no statistically significant change.

Table S3. GTPases containing a cysteine residue near the GN-binding site

Switch II pocket				GN-binding pocket				
1S Residues 8–12	2S Residues 60–63	3S Residues 68–72	4S Residues 96–100	1 Residues 11–18	2 Residues 28–34	3 Residues 58–59	4 Residues 116–120	5 Residues 145–147
ArfRP2		Arl1	Rab33A	Arl4	Gem	FLJ22655	Di-Ras1	Arf1
Arl6		Arl10B	Rab33B	Cdc42 i.1	NKIRas1	RabL5	E-Ras	Arf3
NKIRas1		Arl10C	RabL5	Cdc42 i.2	Rab14		H-Ras i.1	Arf4
NKIRas2		E-Ras		Rab10	Rab39A		H-Ras i.2	Arf5
RabL2A		FLJ22595		Rab13	RasL10B		K-Ras2A	Arf6
RabL2B		Gem		Rab14			K-Ras2B	Arl4
RabL2B		NKIRas2		Rab15			NKIRas2	Arl5 i.1
RabL5		Rab17		Rab1A			N-Ras	Arl5 i.2
SRPRB		Rab20		Rab1B			Rab10	Arl7
		Rad		Rab21			Rab13	Arl8
		RasL10B		Rab26			Rab1A	FLJ22595
		RasL11A		Rab2A			Rab22A	Rerg
		RasL11B		Rab2B			Rab22B	Ris/RasL12
		Rem1		Rab30			Rab32	
		Rem2		Rab33A			Rab33A	
		Ris/RasL12		Rab33B			Rab33B	
		Rnd1		Rab34			Rab38	
		Wrch-1		Rab39A			Rab39A	
		Wrch-2		Rab39B			Rab39B	
				Rab41			Rab3A	
				Rab4A			Rab3B	
				Rab4B			Rab3C	
				Rab8A			Rab3D	
				Rab8B			Rab7L1	
				RabL5			Rab8A	
				Rac1			Rab8B	
				Rac1b			Rap1A	
				Rac2			Rap1B	
				Rac3			RasD2	
				Rheb2			RhoBTB1	
				RhoA			RhoBTB2	
				RhoB			RhoD	
				RhoC			Rif	
				RhoD			Rnd1	
				RhoG			Rnd2	
				Rif			Rnd3	
				Rnd1				
				Rnd2				
				Rnd3				
				SRPRB				
				TC10				
				TCL				

Based on sequence alignments, numerous GTPases are predicted to have a cysteine within regions surrounding the GN-binding site as designated in this table.