Supplemental Information

Rational design of small molecule inhibitors targeting the Ras GEF, SOS1

Chris R. Evelyn, Xin Duan, Jacek Biesiada, William L. Seibel, Jaroslaw Meller, Yi Zheng

SI Table and Figures

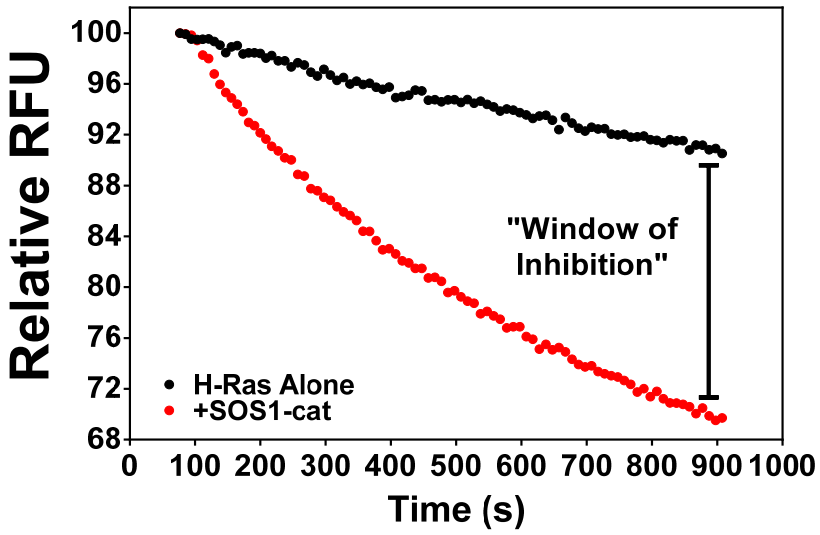
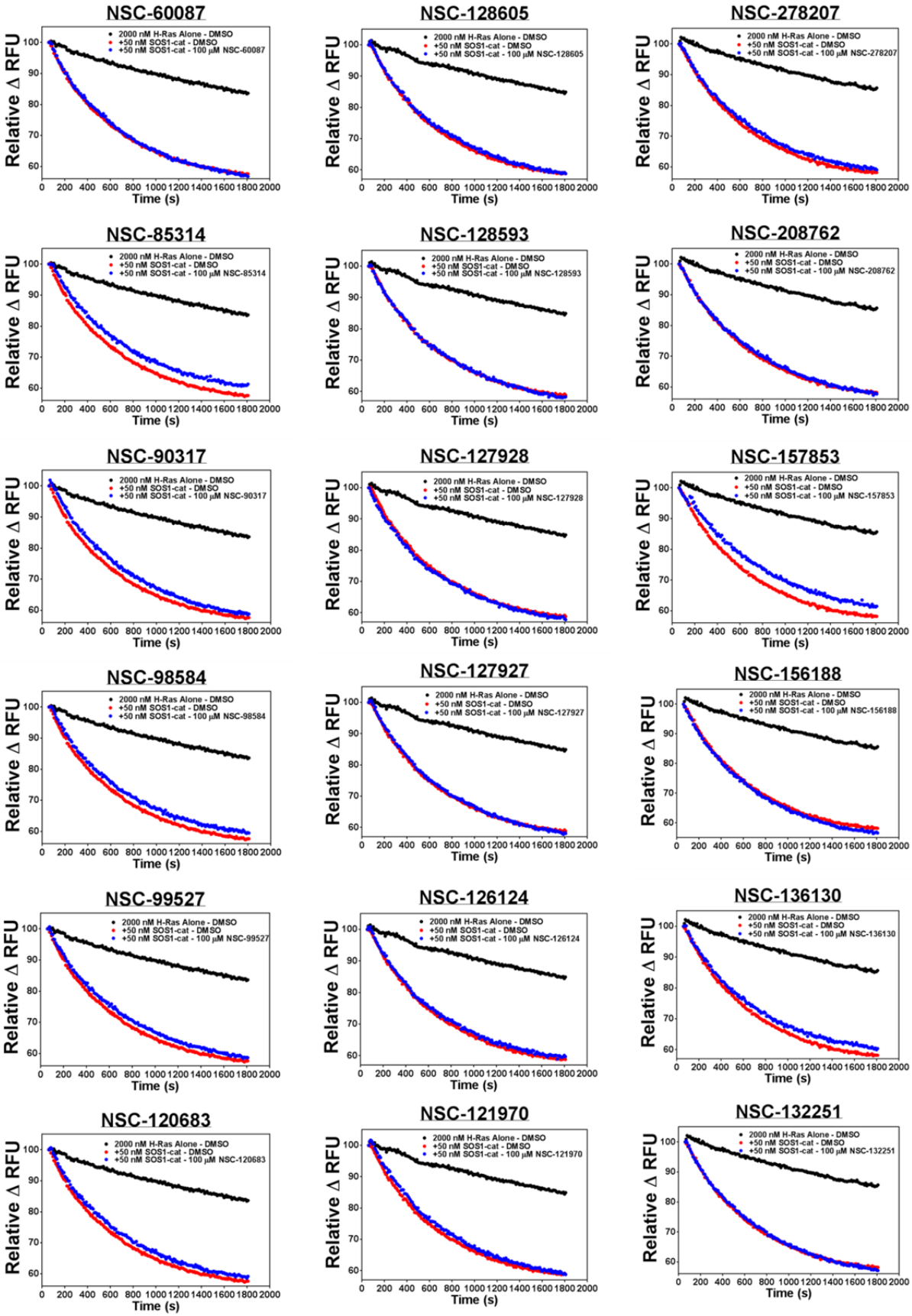
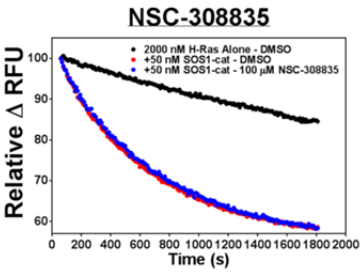
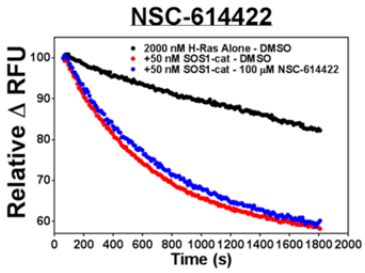
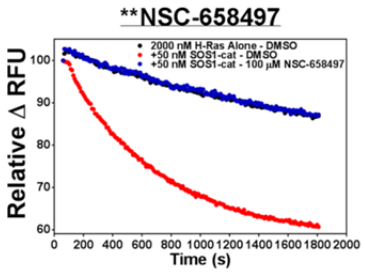


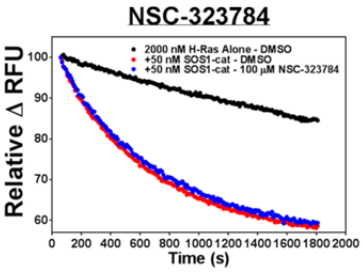
Figure S1, related to Figure 1. Characterization of the SOS1-cat and H-Ras Guanine Nucleotide Dissociation Screening Assay. 50 nM SOS1-cat (●) mediated GDP/GTP nucleotide exchange upon 2 µM H-Ras (aa. 1-166) (●) in the BODIPY-FL-GDP dissociation assay. Purified his_6-SOS1-cat stimulated the exchange of BODIPY-FL labeled GDP for unlabeled GTP on purified his_6-H-Ras (aa. 1-166). The kinetic reaction was carried out as described in the Experimental Procedures. The data is expressed as percent change of relative fluorescence units normalized to the initial time point over 15 minutes. Data were measured in triplicate and represent the mean of N = 3 experiments.

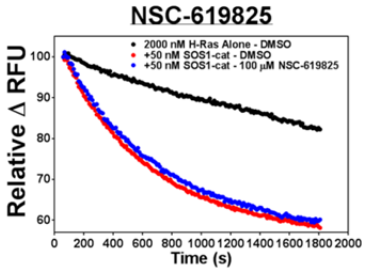

NSC-60087
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-60087
NSC-128605
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-128605
NSC-278207
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-278207
NSC-85314
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-85314
NSC-128593
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-128593
NSC-208762
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-208762
NSC-90317
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-90317
NSC-127928
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-127928
NSC-157853
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-157853
NSC-98584
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-98584
NSC-127927
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-127927
NSC-156188
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-156188
NSC-99527
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-99527
NSC-126124
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-126124
NSC-136130
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-136130
NSC-120683
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-120683
NSC-121970
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-121970
NSC-132251
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-132251
Relative Δ RFU
Time (s)

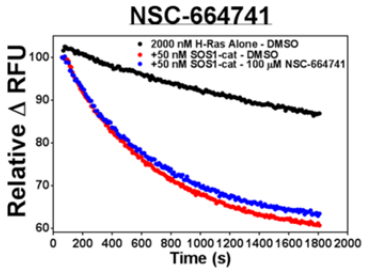

NSC-308835
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-308835
Relative Δ RFU
Time (s)

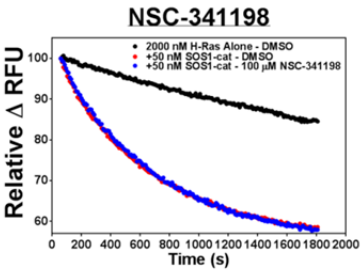

NSC-614422
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-614422
Relative Δ RFU
Time (s)

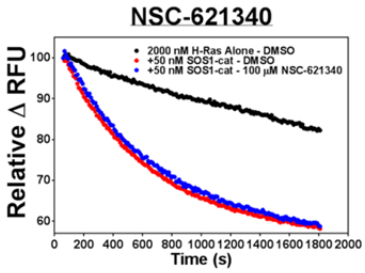

**NSC-658497
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-658497
Relative Δ RFU
Time (s)

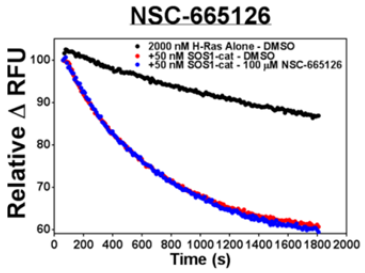

NSC-323784
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-323784
Relative Δ RFU
Time (s)

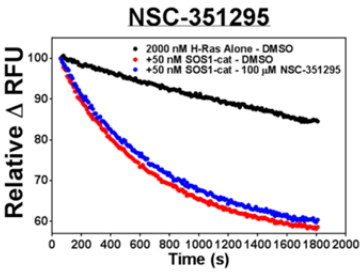

NSC-619825
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-619825
Relative Δ RFU
Time (s)

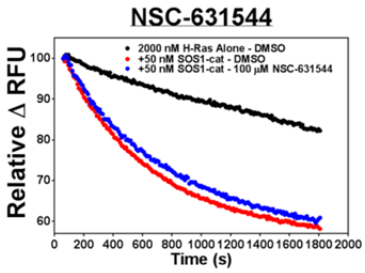

NSC-664741
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-664741
Relative Δ RFU
Time (s)

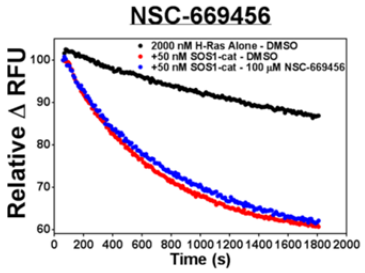

NSC-341198
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-341198
Relative Δ RFU
Time (s)

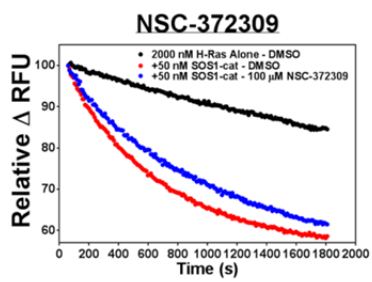

NSC-621340
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-621340
Relative Δ RFU
Time (s)

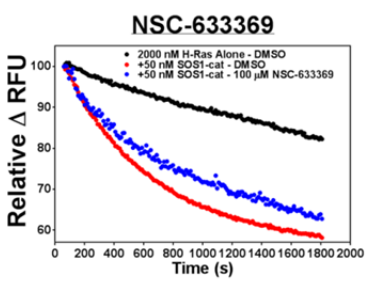

NSC-665126
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-665126
Relative Δ RFU
Time (s)

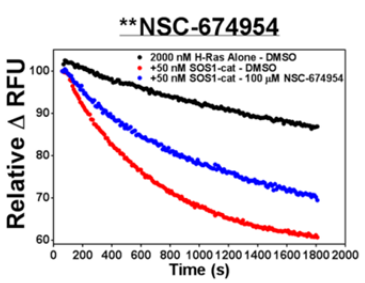

NSC-351295
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-351295
Relative Δ RFU
Time (s)

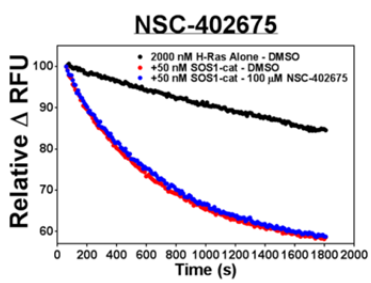

NSC-631544
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-631544
Relative Δ RFU
Time (s)

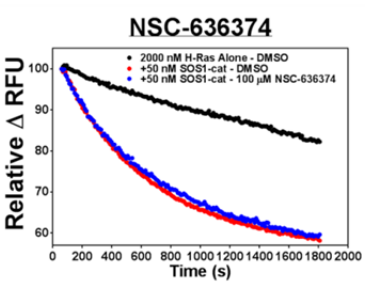

NSC-669456
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-669456
Relative Δ RFU
Time (s)


NSC-372309
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-372309
Relative Δ RFU
Time (s)


NSC-633369
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-633369
Relative Δ RFU
Time (s)


**NSC-674954
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-674954
Relative Δ RFU
Time (s)


NSC-402675
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-402675
Relative Δ RFU
Time (s)


NSC-636374
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-636374
Relative Δ RFU
Time (s)

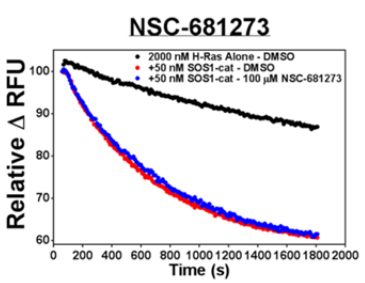

NSC-681273
2000 nM H-Ras Alone - DMSO
+50 nM SOS1-cat - DMSO
+50 nM SOS1-cat - 100 μM NSC-681273
Relative Δ RFU
Time (s)

Figure S2, related to Figure 1. Experimental screen of the 36 NCI/DTP candidate compounds. Thirty-six candidate compounds were tested at a concentration of 100 µM (●) for their ability to inhibit 50 nM SOS1-cat (●) mediated GDP/GTP nucleotide exchange upon 2 µM H-Ras (●) in the BODIPY-FL-GDP dissociation assay. ** indicates the two hit compounds identified in the screen. Data are expressed as percent change of relative fluorescence units normalized to the initial time point over 30 minutes.

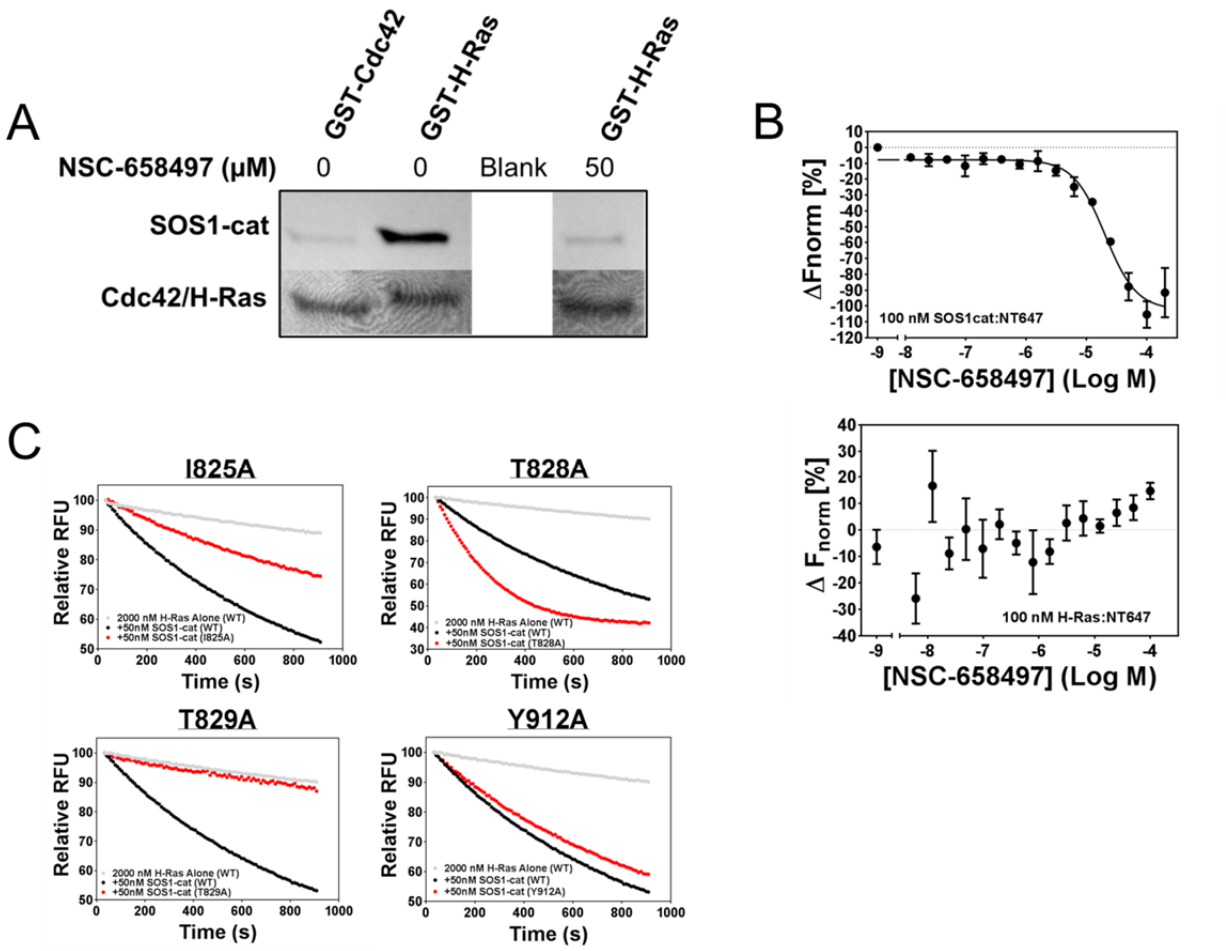


Figure S3, related to Figure 2 and Figure 3. Biochemical Characterization of NSC-658497 and SOS1 mutants. (A) NSC-658497 inhibits the interaction between SOS1-cat and H-Ras. Inhibition of the interaction between SOS1-cat and H-Ras (aa. 1-166) by NSC-658497. Purified his_6-SOS1-cat (50 nM) was incubated with GST-H-Ras (200 nM) or GST-Cdc42 (200 nM) conjugated to glutathione HiCap matrix beads in the presence of DMSO or 50 μM of NSC-658497 as indicated. The beads associated with his_6-SOS1-cat were probed by anti-his western blotting as described in the SI Experimental Procedures. (B) NSC-658497 directly binds SOS1-cat, but not H-Ras. Both purified his_6-SOS1-cat and purified his_6-H-Ras (aa. 1-166) were labeled with NT-647 as described in the Experimental Procedures. NSC-658497 was titrated between 0.01 and 200 μM or

between 0.006 and 100 µM to a constant amount of NT-647 amine-labeled purified his_6-SOS1-cat (100 nM) and purified his_6-H-Ras (aa. 1-166) (100 nM), respectively. Microscale thermophoresis binding was measured and analyzed using Thermophoresis analysis as described in the Experimental Procedures. (C) Mutant SOS1-cat-mediated guanine nucleotide exchange upon H-Ras. 50 nM wild-type or mutant SOS1-cat (black or red circles) mediated GDP/GTP nucleotide exchange upon 2 µM H-Ras (aa. 1-166) (gray circle) in the BODIPY-FL-GDP dissociation assay. Purified wild-type or mutant his_6-SOS1-cat stimulated the exchange of BODIPY-FL labeled GDP for unlabeled GTP on purified his_6-H-Ras (aa. 1-166). The kinetic reaction was carried out as described in the Experimental Procedures. Data in panel A is representative data of N = 3 experiments. Data in panel B represent the mean ± SEM of N= 3 experiments. Data in panel C are expressed as percent change of relative fluorescence.

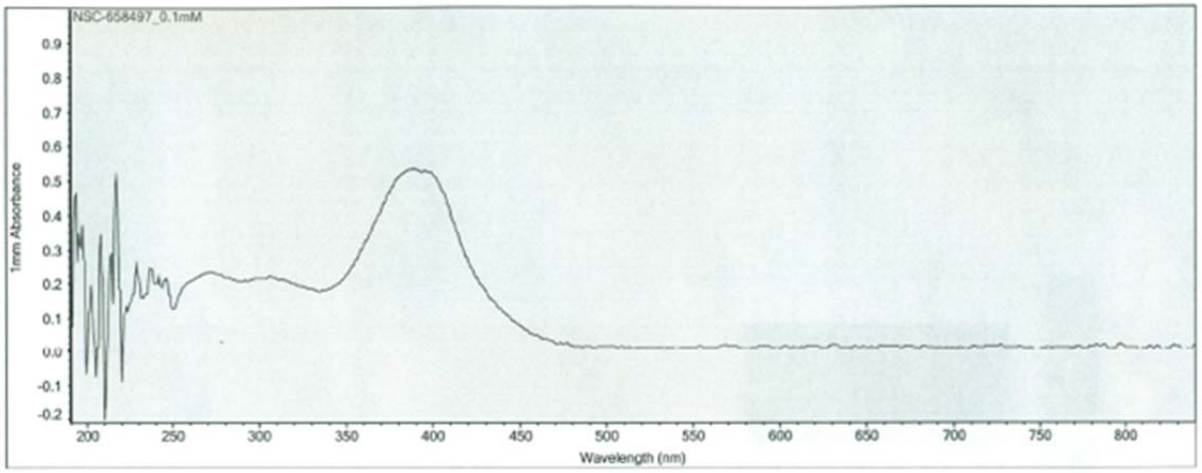


Figure S4, related to Figures 1-4. NSC-658497 NanoDrop UV-Vis absorbance spectrum. The UV-Vis absorbance spectrum for 100 µM of NSC-658497 measured by a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA).

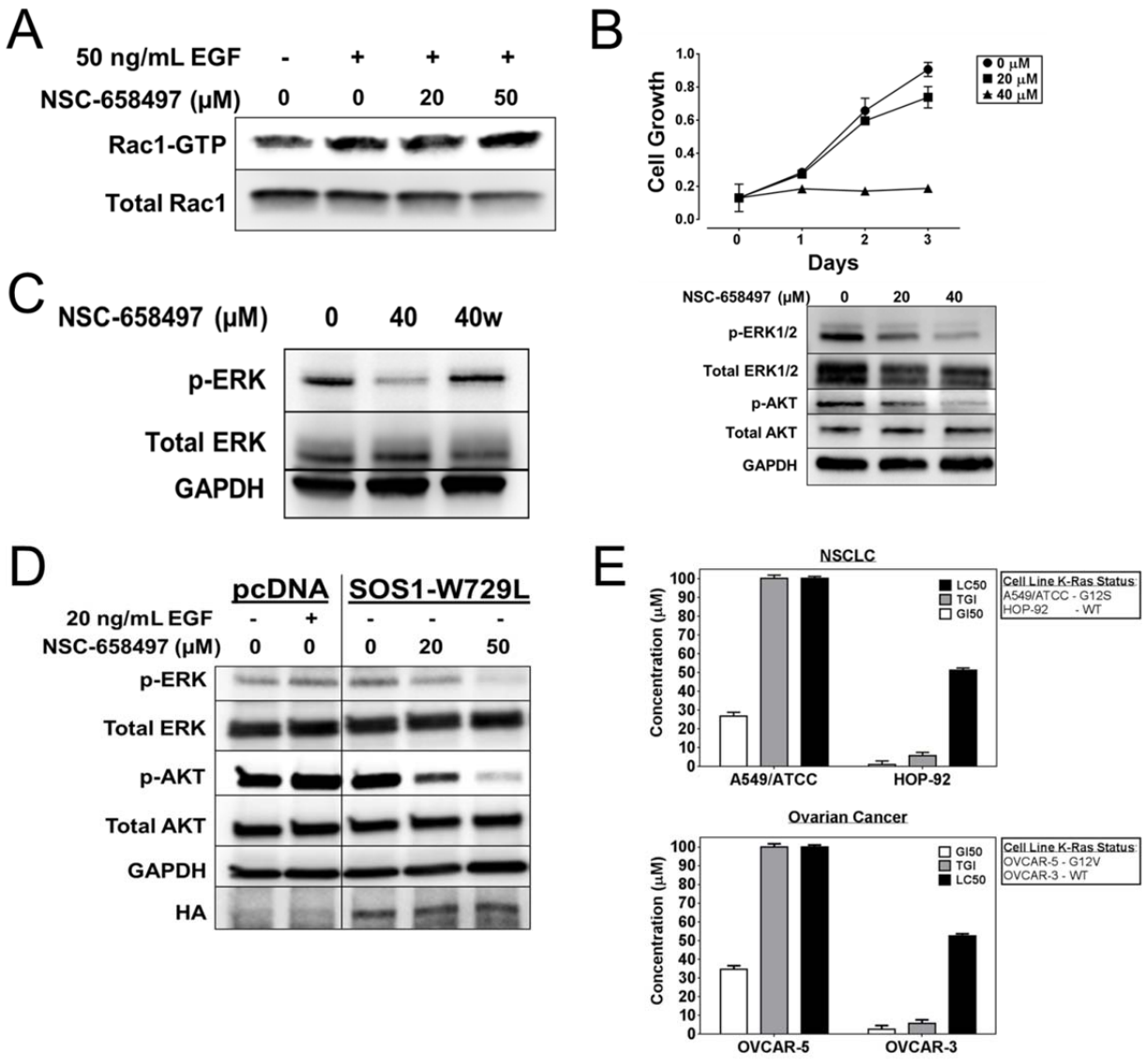


Figure S5, related to Figure 5 and Figure 6. Cellular Characterization of the SOS1 inhibitor, NSC-658497. (A) NSC-658497 does not inhibit Rac1 activation in mouse fibroblasts. NIH/3T3 cells were serum-starved overnight and pre-treated with the indicated concentrations of NSC-658497 for 2 hours. Subsequently, the cells were stimulated with 50 ng/mL of EGF for 5 minutes and were subjected to GST-Pak1 effector domain pull-down experiments. The activation of was analyzed by western blotting as described in the SI Experimental Procedures.

(B) NSC-658497 inhibits Ras signaling and proliferation of HeLa cells. *(Upper Panel)* HeLa cells were subjected to a 3 day proliferation assay. HeLa cells were grown in the presence of DMSO or the indicated concentrations of NSC-658497 for 3 days. Every 24 hours proliferation was measured by MTS assay as described in the Experimental Procedures. *(Bottom Panel)* HeLa cells were serum-starved overnight, and then treated with the indicated concentrations of NSC-658497. Cells were probed for inhibition of downstream Ras signaling for phospho-ERK1/2 and phospho-AKT activation. (C) NSC-658497 reversibly inhibits p-ERK activation in HeLa cells. HeLa cells were serum-starved overnight, and then treated with the indicated concentrations of NSC-658497 for 2 hours. Subsequently, cells in the lane labeled 40w were rinsed three times with PBS, and serum-starved for an additional 2 hours. Cells were probed for inhibition of downstream Ras signaling for phospho-ERK activation. (D) NSC-658497 inhibits SOS1(W729L)-mediated ERK and AKT activation in HEK293T cells. HEK293T cells were transfected with empty vector or mutant HA-tagged SOS1 (W729L) as described in the SI Experimental Procedures, then subsequently serum-starved overnight. Cells were pre-treated with the indicated concentrations of NSC-658497 for 2 hours, and then indicated cells were stimulated with 20 ng/mL of EGF for 15 minutes. Cells were subjected to downstream Ras signaling western blotting experiments for the activation of phospho-ERK and phospho-AKT. (E) NSC-658497 selectively inhibits cell proliferation of wild-type K-Ras versus oncogenic K-Ras non-small cell lung cancer and ovarian cancer cell lines. A subset of the non-small cell lung cancer *(Upper Panel)* and ovarian cancer *(Bottom Panel)* cell NSC-658497 proliferation data from the NCI-60 DTP human tumor cell line screen. NSC-658497 was exposed to the cells for 48 hours and tested at 5 different compound concentrations. GI50 (open bars) indicates the concentration

for 50% inhibition of growth. TGI (gray bars) indicates the concentration required for complete inhibition of growth. LC50 (black bars) indicates the concentration required for 50% reduction in cell number. The activation of Ras signaling in panels A-D were analyzed by western blotting as described in the SI Experimental Procedures. Data in panel B *(Upper Panel)* was performed in triplicate and represent the mean ± SEM. Data in panel B *(Upper Panel)* is plotted as relative growth by normalizing the data to Day 0. Data in panel B *(Bottom Panel)* is representative of N = 3 experiments.

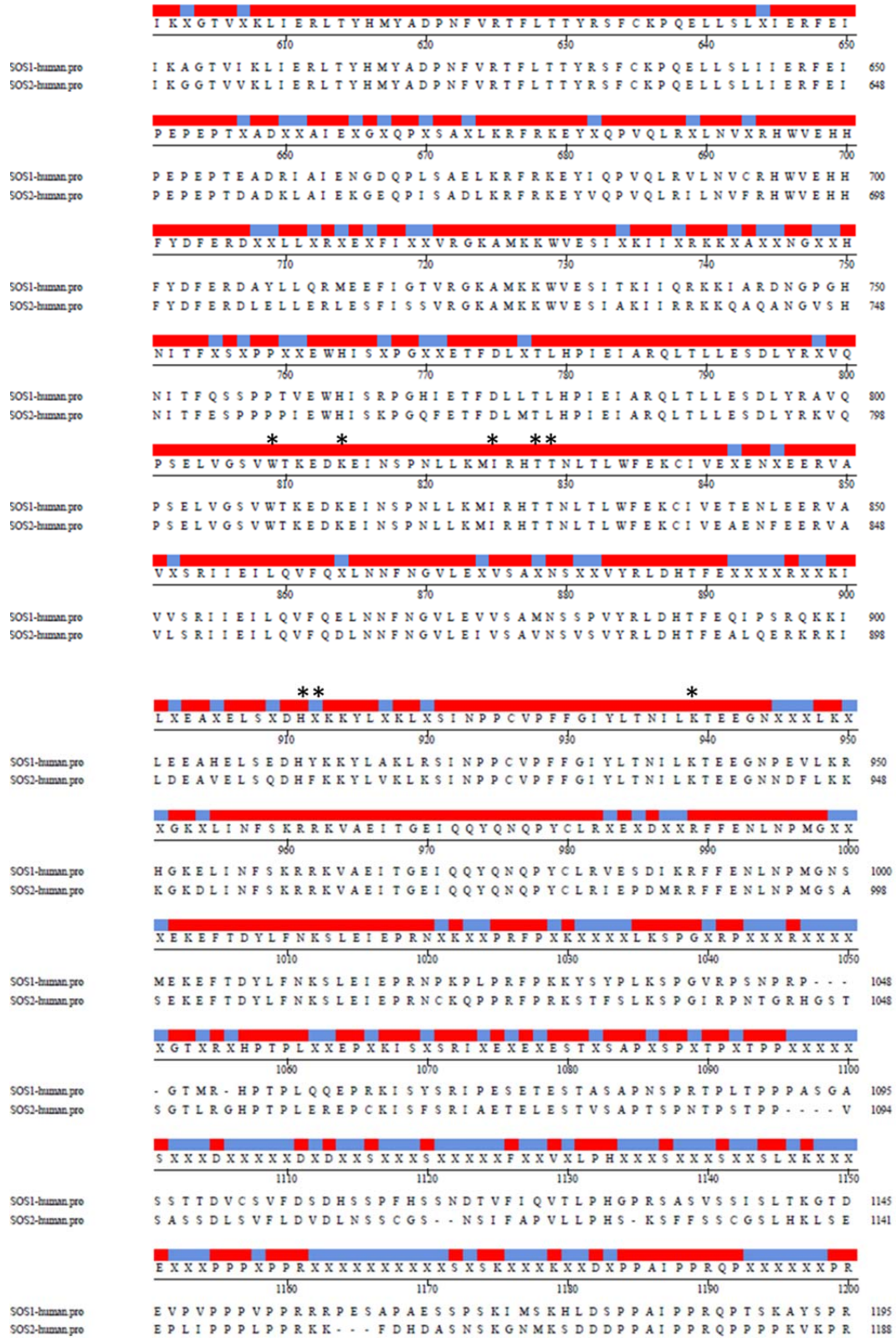

SOS1-human.pro
SOS2-human.pro

Figure S6A, related to Figure 5. Alignment of primary amino acid sequences of human SOS1 and SOS2. The amino acid sequences of human SOS1 and SOS2 were aligned using MegAlign and the Clustal W method (DNAStar Lasergene 8 Suite; Madison, WI). The portion of the sequence involved in the NSC-658497 interaction is displayed. Key residues involved in the NSC-658497 interaction with SOS1 are indicated by the asterisks (*).

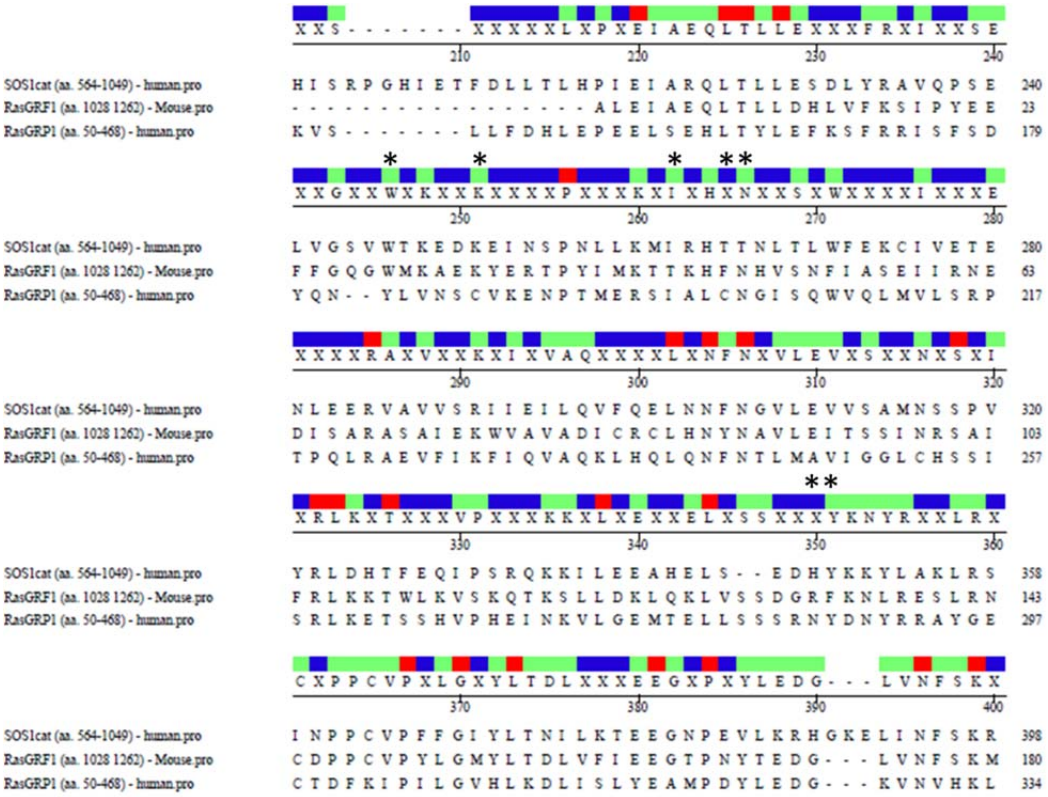


Figure S6B, related to Figure 5. Alignment of primary amino acid sequences of the Cdc25 domain of RasGEFs. The amino acid sequences of the Cdc25 domains of human SOS1 (aa. 564-1049; Acc.#: BC140215), mouse RasGRF1 (aa. 1028-1262; Acc.#: NM_011245), and human RasGRP1 (aa. 50-468; Acc.#: NM_005739) were aligned using MegAlign and the Clustal W method (DNAStar Lasergene 8 Suite; Madison, WI). The portion of the sequence involved in the NSC-658497 interaction is displayed. Key residues involved in the NSC-658497 interaction with SOS1 are indicated by the asterisks (*).

```
1024:S          1040:S              1060:S              1080:S
|     |    .    |    .    |    .    |    .    |    .    |    .    |
PLGSALEIAEQLTLLDHLVFKSIPYEEFFGQGWMKAEKYERTPYIMKTTKHFNHVSNFIASEIIRNEDIS
. ...:|||.|||||:...::.:...|..|..|.|.:|....|.::|..:|..........|:..|.:.
LTLHPIEIARQLTLLESDLYRAVQPSELVGSVWTKEDKEINSPNLLKMIRHTTNLTLWFEKCIVETENLE
|  |    .    |    .    |    .    |    .    |    .    |    .    |    .
777:A        790:A               810:A               830:A
```

```
      1100:S              1120:S              1140:S              1160:S
 .    |    .    |    .    |    .    |    .    |    .    |    .    |
ARASAIEKWVAVADICRCLHNYNAVLEITSSINRSAIFRLKKTWLKVSKQTKSLLDKLQKLVSSDGRFKN
.|...:.:.:.:..:...|.|:|.|||:.|..|.|.::||..|...:....|.:|:....|..  ..:|.
ERVAVVSRIIEILQVFQELNNFNGVLEVVSAMNSSPVYRLDHTFEQIPSRQKKILEEAHELSE--DHYKK
   |    .    |    .    |    .    |    .    |    .    |    .      |
   850:A               870:A               890:A                910:A
```

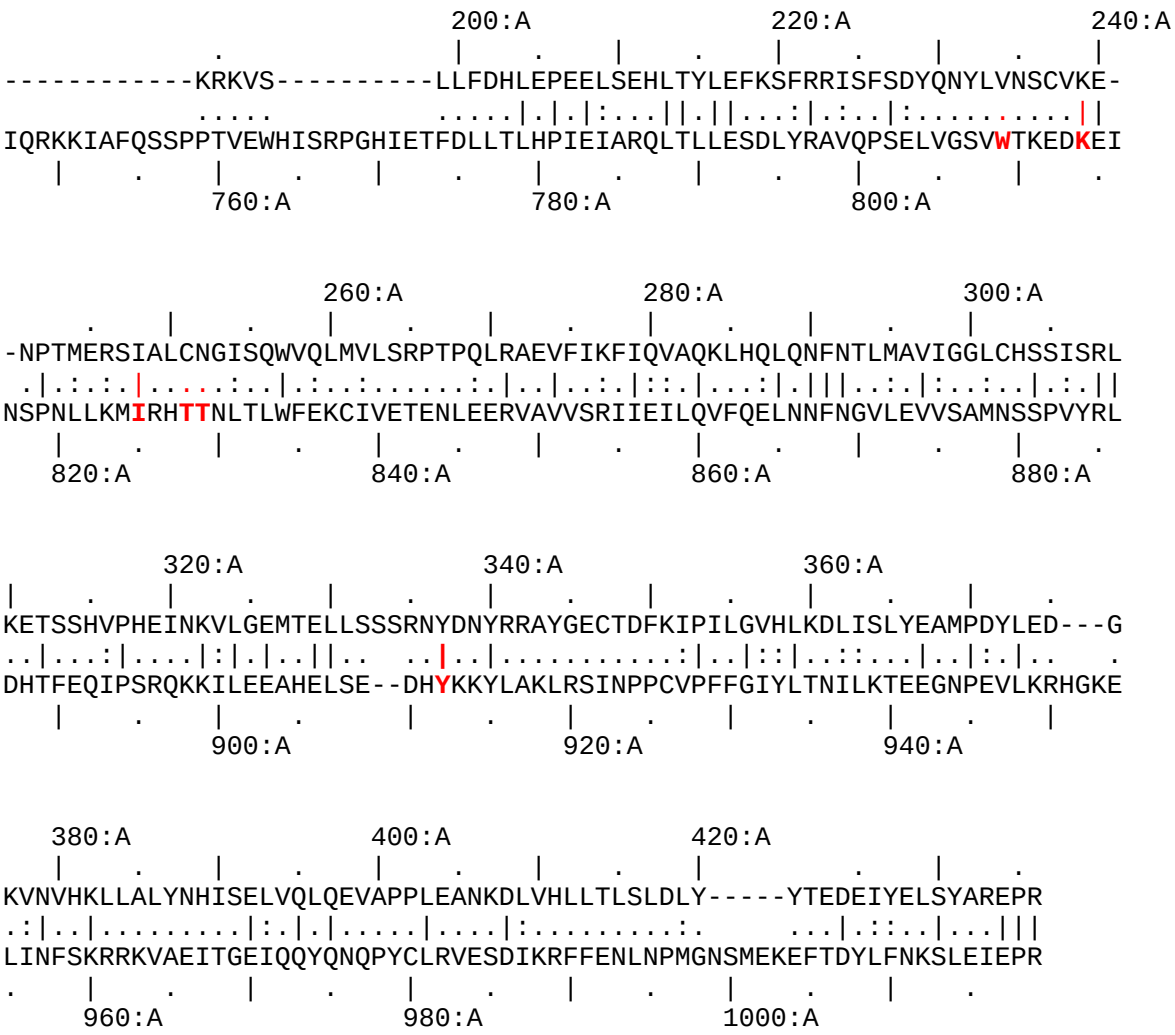
```
                1180:S                1200:S              1220:S
 .    |    .    |    .    |    .       |    .    |    .    |    .    |
LRESLRNCDPPCVPYLGMYLTDLVFIEEGTPNYTED---GLVNFSKMRMISHIIREIRQFQQTTYKIDPQ
....||...|||||:.|.|||.:...|||.|.....   .|:||||.|.:..|..||.|:|...|.:..:
YLAKLRSINPPCVPFFGIYLTNILKTEEGNPEVLKRHGKELINFSKRRKVAEITGEIQQYQNQPYCLRVE
.    |    .    |    .    |    .    |    .    |    .    |    .    |
               930:A               950:A               970:A
```

```
        1240:S
    .    |          .    |    .
PKVIQYLLDESFML------DEESLYESSLLIEPK
..:..:......|.      ..:.|:..||.|||:
SDIKRFFENLNPMGNSMEKEFTDYLFNKSLEIEPR
.    |    .    |    .    |    .
     990:A               1010:A
```

P-value: 3.63e-12
Equ: 233
RMSD: 2.43
Score: 614.96
Align-len: 245
Gaps: 12 (4.90%)

Identity: 28.16%
Similarity: 54.69%

Figure S6C, related to Figure 5. Structure alignment of SOS1 versus RasGRF1. The structure of the catalytic domain of unbound SOS1 (PDB ID: 2II0) (bottom strand) was aligned with the structure of the catalytic domain of unbound RasGRF1 (PDB ID: 2IJE) (top strand) using the method FATCAT. Structurally equivalent and identical residues are indicated by (|). Structurally equivalent and similar residues are indicated by (:). Structurally equivalent, but not similar residues are indicated by (.) Additional results of the FATCAT analysis are indicated below the alignment.



P-value: 1.45e-10
Equ: 253
RMSD: 2.77
Score: 572.27
Align-len: 415
Gaps: 162 (39.04%)

Identity: 13.01%
Similarity: 23.86%

Figure S6D, related to Figure 5. Structure alignment of SOS1 versus RasGRP1. The structure of the catalytic domain of unbound SOS1 (PDB ID: 2II0) (bottom strand) was aligned with the structure of the catalytic domain of unbound RasGRP1 (PDB ID: 2L9M) (top strand) using the method FATCAT. Structurally equivalent and identical residues are indicated by (|). Structurally equivalent and similar residues are indicated by (:). Structurally equivalent, but not similar residues are indicated by (.) Additional results of the FATCAT analysis are indicated below the alignment.

SI Experimental Procedures

Plasmids, Cell Lines, and Reagents **-** BODIPY® Fluorescein(FL)-GDP, BODIPY® Texas Red(TR)-GTP, One Shot® BL21(DE3) chemically competent *e.coli*, and DTT were obtained from Life Technologies (Grand Island, NY). Guanosine 5'-triphosphate sodium salt solution, Tween-20, and sodium deoxycholate (SDC) were obtained from Sigma-Aldrich (St. Louis, MO). Guanosine diphosphate (GDP), dimethyl sulfoxide (DMSO), and NP-40 Alternative were obtained from Calbiochem®-EMD Millipore (Billerica, MA).

Human epidermal growth factor (EGF) was obtained from Roche Applied Science (Indianapolis, IN). Full-length human SOS1 (residues 1-1343) was sub-cloned from the full-length human SOS1 MGC Clone (Accession#: BC140215) (Life Technologies, Grand Island, NY) into the mammalian pcDNA3.1+ vector (Life Technologies) with a N-terminal (1X) HA epitope tag using the HindIII and XbaI restriction enzyme sites. Human SOS1-cat encoding the REM and Cdc25 domains (residues 564-1049) was sub-cloned from the full-length human SOS1 MGC Clone (Accession#: BC140215) (Life Technologies) into the Escherichia coli pET15b vector (Novagen®-EMD Millipore) using the XhoI and HindIII restriction enzyme sites. C-terminal truncated human H-Ras (residues 1-166) was sub-cloned from the full-length human H-Ras/pcDNA3.1+ vector (Missouri S&T cDNA Resource Center, Rolla, MO) into the pET15b vector using the XhoI and HindIII restriction enzyme sites. Full-length human H-Ras (residues 1-189) was sub-cloned from the full-length human H-Ras/pcDNA3.1+ vector (Missouri S&T cDNA Resource Center, Rolla, MO) into the pGEX-2T vector using the BamHI and EcoRI restriction enzyme sites. A tyrosine (Y) to alanine (A) mutant at residue 64 was inserted into both H-Ras constructs using the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies,

Santa Clara, CA). In addition, SOS1 single mutants (K939A, T935A, F929A, Y912A, H911A, E902A, N879A, T829A, T828A, I825A, M824A, K814A, T810A, and W809A) or mutant W729L were inserted into the human SOS1-cat/pET15b vector or HA-SOS1/pcDNA3.1+ vector, respectively, using the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies, Santa Clara, CA).

Human PC-3 and DU-145 prostate cancer cells were kind gifts from Dr. Zhongyun Dong, MD, Ph.D. (University of Cincinnati, Department of Internal Medicine). Human HeLa cells were kind gifts from Dr. James Mulloy, Ph.D. (Cincinnati Children's Hospital Medical Center, Division of Experimental Hematology and Cancer Biology). The generation and maintenance of the immortalized parental $Cdc42^{f/f}$ (wild-type) and retrovirally transduced MIEG3-H-RasG12V mouse embryonic fibroblasts (MEFs) were described previously (Stengel and Zheng, 2012). Mouse NIH/3T3 fibroblasts were maintained in Dulbecco's Modified Eagle Medium (DMEM) containing 10% heat-inactivated calf serum, 100 units/mL penicillin, and 100 µg/mL streptomycin at 37°C in 5% CO_2. Human HeLa and DU-145 cells were maintained in DMEM containing 10% heat-inactivated fetal bovine serum, 100 units/mL penicillin, and 100 ug/mL streptomycin at 37^oC in 5% CO_2. Human PC-3 cells were maintained in Eagle's Minimum Essential Medium (EMEM) containing 10% heat-inactivated fetal bovine serum, 100 units/mL penicillin, and 100 ug/mL streptomycin at 37^oC in 5% CO_2.

Computational Virtual Screening and SAR Analysis **-** Virtual Screening was performed to identify candidate molecules that could disrupt the interaction between SOS1 and H-Ras, by targeting the SOS1 interaction interface ("catalytic site") with H-Ras. The docking simulations

for the virtual screening were performed using rigid body docking, as implemented in AutoDock ver. 4.2 (Morris et al., 2009), in conjunction with the Cincinnati Children's Hospital Medical Center (CCHMC) Center for Protein Informatics' computational pipelines on a Linux cluster with upwards of 512 CPUs. The crystal structure of the SOS1 complex with H-Ras (PDB ID: 1XD2) (Sondermann et al., 2004), as well as the unbound SOS1 structure (PDB ID: 1XD4) (Sondermann et al., 2004) were used to build alternative models of the SOS1 enzyme for docking simulations, using ADT graphical interface. Sppider (http://sppider.cchmc.org) and Polyview (http://polyview.cchmc.org) were used to analyze the protein-protein complexes, and guide the choice of simulation boxes.

A subset of 118,500 drug-like synthetic compounds from the NCI/DTP Open Chemical Repository (http://dtp.cancer.gov) was used for virtual screening. These compounds were derived from the NCI Plated 2007 set of 139,735 compounds deposited in the Zinc library (http://zinc.docking.org/catalogs/ncip) by using chemoinformatic filters as described in (Irwin and Shoichet, 2005; Irwin et al., 2012; OpenEye Scientific Software). 3D structures for the resulting subset of 118,500 compounds were downloaded from ZINC. Gesteiger partial charges were used for both the SOS1 enzyme and chemical compounds. Screening was performed in three stages, using increasingly stringent parameters and gradually more extensive sampling. The latter was achieved by increasing the number of energy evaluations (from 2×10^5 to 1×10^7), Genetic Algorithm runs (from 20 to 50) and population size (from 75 to 150), as discussed in detail previously (Biesiada et al., 2011). After initial fast screening, 30,000 top candidates with the highest estimated binding affinities were retained, and subsequently re-scored using

improved sampling in the refinement stage. 3,000 top hits were then re-scored using extensive sampling, and assessed further to select candidates for experimental validation.

These candidate compounds were ranked based on their estimated binding affinities (their median predicted inhibition constant) and entropy of clustering of docking poses in multiple runs of docking simulations (filtering out candidate compounds predicted to bind to SOS1 with multiple distinct poses observed with low frequency), resulting in a set of 135 hits. These top hits were subsequently clustered based on their chemical similarity using Chemmine (http://chemmine.ucr.edu) to further select candidates for experimental validation, while avoiding over-representation of some classes of chemicals, and to visually analyze candidate compounds.

From this set of top candidates, 36 compounds representing different clusters of chemicals were selected for experimental screening based upon assessment of drug-like properties, similarity to classes of compounds often identified in virtual screening as false positives, and availability of compounds from the NCI/DTP Open Chemical Repository (http://dtp.cancer.gov). The predicted docking poses for top hits were analyzed using Polyview-MM) (Porollo and Meller, 2010) to rationally predict which amino acid residues in SOS1 to mutate for site of action mutagenesis studies.

Thirteen additional structurally-related NSC-658497 (CID 54608884) NCI analogs were selected for experimental testing by identifying similar compounds from the similar compound and similar conformer lists in the National Center for Biotechnology Information (NCBI) - PubChem database (http://pubchem.ncbi.nlm.nih.gov). In addition, using the Accelrys' Pipeline Pilot (Version 8.5.0.200) software, a combination of substructural filters (including alternative

tautomers of the coumarin structure) and structure similarity searches based on NSC-658497 were run on the University of Cincinnati – Drug Discovery Center's 350,000 drug-like compound library collection. As a result of this search, nineteen structurally similar compounds were selected for experimental testing.

Protein Purifications **-** Both wild-type and mutant recombinant human SOS-cat (residues 564-1049) containing the REM-Cdc25 catalytic module and recombinant human c-terminal truncated H-Ras (residues 1-166) were expressed in BL21(DE3) *Escherichia coli* cells as N-terminal $(His)_6$-tagged fusion protein using the pET expression system (Novagen). For the N-terminal $(His)_6$-tagged fusion proteins, the *Escherichia coli* cells were grown to an OD_{600} between 0.4 and 0.8, then induced with 1 mM IPTG overnight at room temperature or for 4 hours at $37^{o}C$ for SOS1-cat and H-Ras, respectively. Recombinant $(His)_6$-tagged SOS1 fusion protein was purified using Ni^{+2}-agarose affinity chromatography followed by gel filtration with a HiLoad 16/60 Superdex 200 column and FPLC AKTA Purifier UPC 10 system (GE Healthcare Bio-Sciences, Piscataway, NJ). In addition, both wild-type and mutant full-length recombinant human H-Ras were expressed in BL21(DE3) *Escherichia coli* cells as N-terminal glutathione S-transferase (GST)-tagged fusion proteins using the pGEX expression system (GE Healthcare Bio-Sciences). The *Escherichia coli* cells for the GST-tagged fusion proteins were grown to an OD_{600} between 0.4 and 0.8, then induced with 1 mM IPTG for 3 hours at $37^{o}C$. Both the wild-type and mutant full-length recombinant GST-tagged H-Ras fusion proteins were purified using Glutathione HiCap Matrix (Qiagen, Valencia, CA) affinity chromatography.

Cell Transfection **-** 9.0 x 10^5 HEK293T cells were seeded per well in a 6-well dish. Cells were transfected with 2 ug of pcDNA3.1+ or HA-SOS1(W729L)/pcDNA3.1+ for twenty-four hours using the TransIT-293 transfection reagent (Mirus, Madison, WI) following the manufacturer's instructions.

Western Blot Analysis **-** For western blot analysis, NIH/3T3, MEF, HeLa, PC-3, or DU-145 cells were grown in 10-cm dishes to about 80% cell confluency. Lysates were clarified and the protein concentrations were normalized by Bradford assay using the Bio-Rad protein assay dye reagent (Bio-Rad). Lysates were separated in 4-15% Mini-PROTEAN® TGX™ precast polyacrylamide gels by SDS-PAGE and transferred onto PVDF membranes with the Trans-Blot® Turbo™ transfer system (Bio-Rad, Hercules, CA). The membranes were blocked with Odyssey Blocking Buffer (Li-Cor, Lincoln, NE) or 5% BSA in TBS-T (20 mM Tris-HCl, pH 7.6, 150 mM NaCl, and 0.1% Tween 20) overnight at 4°C or for 30-60 minutes at room temperature, then probed with primary antibodies and subsequently with IRDye® 800CW goat (Li-Cor) or horseradish peroxidase (HRP)-coupled (GE Healthcare Bio-Sciences) secondary antibodies for IR and chemiluminescence analysis, respectively. The anti-pan-Ras (#05-516) and anti-Rac1 (#05-389) antibodies were purchased from Millipore (Billerica, MA). The anti-R-Ras (#47536) antibody was purchased from abcam (Cambridge, MA). The anti-HA (#3724), anti-phospho-ERK1/2 (#4370), anti-ERK1/2 (#9102), anti-phospho-AKT (Ser 473) (#4060), anti-AKT (#9272), anti-phospho-EGFR (Tyr 1068) (#2234), anti-EGFR (#2232), and anti-GAPDH (#5174) antibodies were purchased from Cell Signaling Technology (Danvers, MA). All primary antibodies were diluted (1:1000) in Odyssey Blocking solution (Li-CoR) or TBS-T buffer containing 5% BSA. Membranes were

developed by either the Odyseey® CLx infrared imaging system (Li-CoR) or ECL western blot detection reagent (GE Healthcare Bio-Sciences). Images for the ECL developed membranes were taken by the Fuji Image Reader LAS1000.

In Vitro Protein-Protein Interaction Assay - Approximately, 50 nM of his_6-SOS1-cat protein was pre-incubated in the absence and presence of NSC-658497 at indicated concentrations for 5-10 minutes at room temperature in 0.5 mL of binding buffer B (20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM EDTA, 1% Triton X-100, and 1 mM DTT). Then, 25 uL of a 1:1 slurry of glutathione HiCap matrix with binding buffer A (20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM $MgCl_2$, 1% Triton X-100, and 1 mM DTT), containing about 5 ug of pre-bound GST-H-Ras protein (about 200 nM – final), was added to the SOS1-cat and compound mixture. Samples were incubated at room temperature for 1 hour with constant agitation, then washed three times with 1 mL of binding buffer A. The amount of his_6-SOS1-cat protein bound to GST-H-Ras beads was determined by anti-his western blotting. Membranes were developed by ECL western blot detection reagent (GE Healthcare Bio-Sciences), and images for the ECL developed membranes were taken by the Fuji Image Reader LAS1000. The amount of GST-H-Ras on the beads was determined by ponceau S staining. The primary His (#2365) antibody was obtained from Cell Signaling Technology.