Massively parallel profiling of intracellular protein variant abundance, activity, interactions, and druggability with LABEL-seq

Jessica Simon

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

Dustin James Maly, Chair

Michael Gelb

Lauren J Rajakovich

Program Authorized to Offer Degree:

Chemistry

University of Washington

Abstract

Massively parallel profiling of intracellular protein variant abundance, activity, interactions, and druggability with LABEL-seq

Jessica Simon

Chair of the Supervisory Committee:

Dustin James Maly

Department of Chemistry

Here we describe labeling with barcodes and enrichment for biochemical analysis by sequencing (LABEL-seq), an assay for massively parallel profiling of pooled protein variants in human cells. By leveraging the intracellular self-assembly of an RNA-binding domain (RBD) with a stable, variant-encoding RNA barcode, LABEL-seq facilitates the direct measurement of protein properties and functions using simple affinity enrichments of RBD protein fusions, followed by high-throughput sequencing of co-enriched barcodes. Measurement of ~20,000 variant effects for ~1,600 BRaf variants revealed that variation at positions frequently mutated in cancer minimally impacted intracellular abundance but could dramatically alter activity, protein–protein interactions and druggability. Integrative analysis identified networks of positions with similar biochemical roles and enabled modeling of variant effects on cell proliferation and small molecule-promoted degradation. Thus, LABEL-seq enables direct measurement of multiple biochemical properties in a native cellular context, providing insights into protein function, disease mechanisms and druggability.

Table of Contents

Introduction

A protein's sequence defines its properties, including biophysical characteristics, intracellular function, pathogenicity, and druggability. Protein variant effect measurements have provided valuable insight into mechanisms of protein allostery[1,2] and regulation[3], drug resistance and sensitivity[4,5], molecular evolution[6,7], the effects of germline and somatic mutations[8-10] and guided design and engineering efforts[11]. Multiplexed assays of variant effect (MAVEs) are powerful tools for characterizing the consequences of protein sequence variation because they leverage high-throughput DNA sequencing to measure thousands of protein variants in a single experiment. However, while first-generation MAVEs have enabled the measurement of sequence variation effects in diverse proteins[12], they can be difficult to scale, may divorce the protein from its cellular context, and often measure protein function indirectly; for example, using cell proliferation as a proxy for enzymatic activity. Furthermore, most MAVEs were developed to measure a single protein effect, making it difficult for them to be reconfigured to assess additional effects. Thus, a general, scalable platform for measuring protein sequence variation effects that can be readily reconfigured to directly read out diverse protein properties and functions in a human intracellular environment is urgently needed.

The multi-domain kinase BRaf is emblematic of the need for such a platform. BRaf plays a central role in oncogenic signaling and >250 somatic mutations have been identified[13-15]. While numerous biochemical and cellular studies have been performed, they have not provided a comprehensive view of sequence variation effects due to their focus on a limited subset of BRaf variants [16,17]. Thus, our understanding of the mechanistic basis of BRaf regulation and signaling, the role that BRaf wild type (WT) and specific variants play in disease-relevant signaling, and the druggability of clinically-observed variants remains limited. Like many signaling proteins, BRaf's

activity is tuned by conformational changes and protein-protein interactions. Consequently, BRaf sequence variation can affect multiple coupled intra- and intermolecular interactions. Therefore, to fully understand BRaf's function, multiple biochemical properties must be measured in a native human cellular environment where all regulatory components and interaction partners are present.

Here, we present LAbeling with Barcodes and Enrichment for biochemicaL analysis by sequencing (LABEL-seq), a platform that enables massively parallel biochemical profiling of thousands of protein variants in cultured human cells (Fig. 1a). LABEL-seq leverages the intracellular self-assembly of an RNA binding domain (RBD) with a single circular RNA that contains a hairpin with high affinity for the RBD and a short barcode sequence that provides a nucleic acid-based identifier for a co-expressed protein variant. Intracellular protein variant properties or functions can be quantified after cell lysis by performing a simple affinity enrichment of RBD-tagged protein fusions, followed by high-throughput sequencing of co-enriched barcodes. The ability to vary the type of affinity enrichment performed enables the measurement of diverse protein properties or functions.

Applying LABEL-seq to ~1,600 BRaf variants revealed that substitutions at positions where somatic mutations are frequently observed minimally affected intracellular abundance. However, mutations at many of these same positions had dramatic effects on activity, interactions with multiple signaling partners, and druggability. Integrative analysis of our variant effect data revealed groups of positions that play similar mechanistic roles in modulating BRaf's regulation and function and illustrated how LABEL-seq-derived biochemical measurements can enable predictive modeling of proliferation and sensitivity to small molecule-promoted protein degradation, shedding light on the factors influencing druggability. Applying LABEL-seq to ~1,600 BRaf variants generated ~20,000 variant effect measurements of abundance, activity,

protein-protein interactions, and druggability. Thus, LABEL-seq is a platform technology for measuring multiple biochemical properties and functions of protein variants, facilitating a deeper understanding of the complex cellular behavior of proteins.

Results

LABEL-seq, a platform for multiplexed protein profiling

To implement LABEL-seq, an RNA barcoding component that forms a stable complex with an RBD is required. As the LABEL-seq barcoding component, we developed RNAs that (1) are stable to cellular and biochemical enrichment conditions, (2) tolerate the addition of a 15-20-nucleotide barcode and, most importantly, (3) contain the MS2 hairpin, which self-assembles with the MS2 coat protein (MCP)[18] RBD. Due to RNA's general instability, we pursued a highly stable ribozyme-assisted circRNA architecture[19,20]. We used RNAfold to guide the placement of a 16-nucleotide barcode and replacement of a stem loop with the MS2 hairpin within two different architectures (Fig. 1b and Extended Data Fig. 1a,b) and identified seven MS2-circRNA designs that were predicted to possess minimally altered folding of the terminal ribozymes required for circularization. We found that most of these designs expressed at comparably high levels to the parent circRNAs and were minimally affected by barcode identity (Fig. 1c and Extended Data Fig. 1c). MS2-circRNAs 1 and 2 were selected for further investigation.

To identify a suitable version of the MCP RBD, we fused four different MCP variants (Extended Data Fig. 1d) that contain an affinity-enhancing mutation (V29I)[21,22] to the N-terminus of the tyrosine kinase Lck, transiently expressed the fusions in HEK293T cells with MS2-circRNA 2, and assessed the ability of each fusion to enrich barcodes. We found that all four MCP-Lck fusions efficiently and specifically co-enriched MS2-circRNA 2 when captured with an

immobilized kinase inhibitor (Extended Data Fig. 1e). We selected a tandem MCP dimer (tdMCP) that contains a deletion that prevents higher-order oligomerization[21-23] for all subsequent experiments based on its ability to efficiently co-enrich MS2-circRNAs 1 and 2 (Fig. 1d).

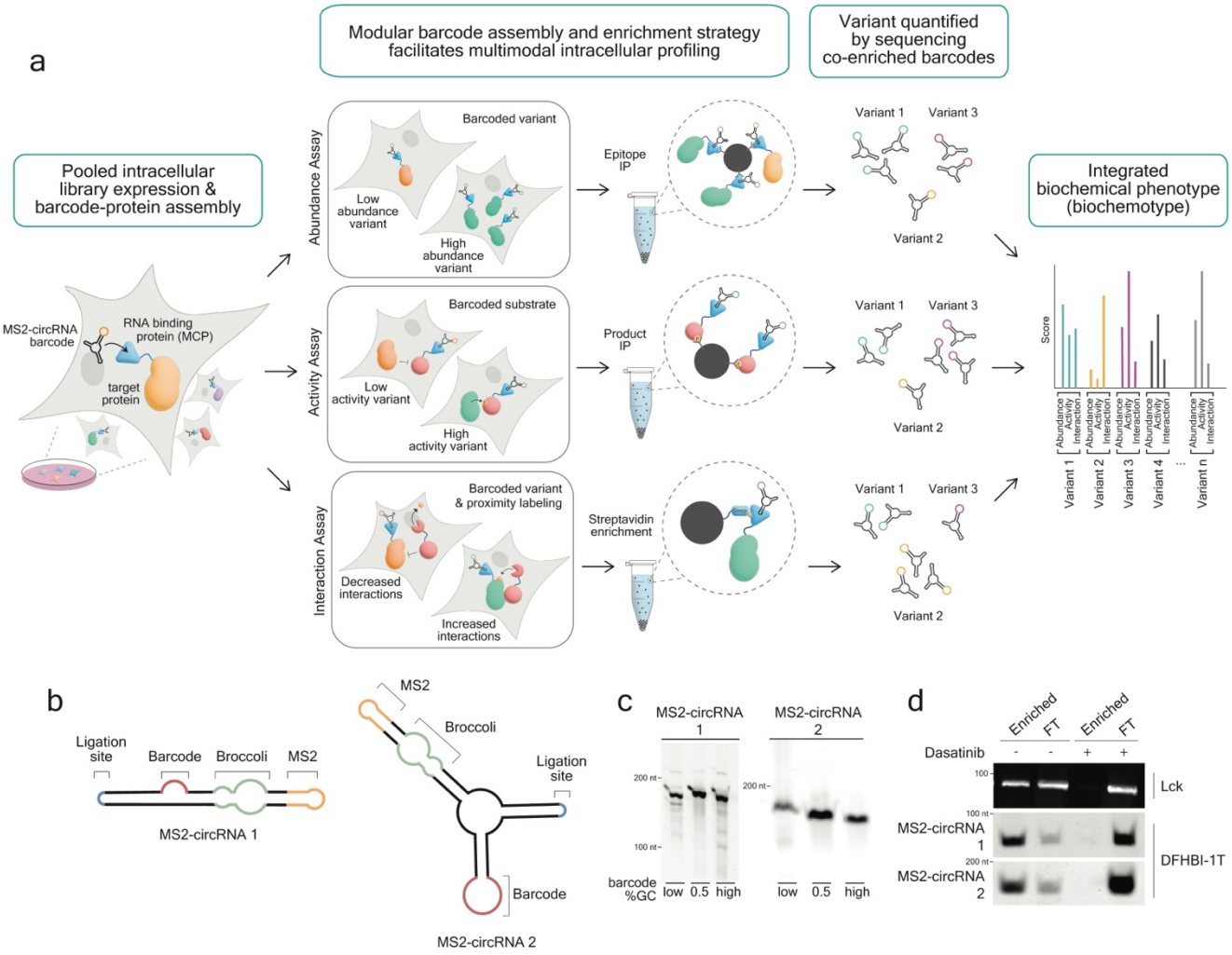


Fig. 1 | Overview of LABEL-seq and development of the RNA and protein components. a, Schematic of the LABEL-seq platform. Intracellular self-assembly of tdMCP (an RBD with high affinity for the MS2 stem loop)-tagged proteins with a single circular RNA containing the MS2 stem loop and a short barcode sequence facilitates the identification of a co-expressed protein variant. Intracellular variant properties or functions such as abundance (top), activity (middle), or protein-protein interactions (bottom) are quantified after cell lysis by performing an appropriate affinity enrichment of tdMCP-tagged protein fusions, followed by quantification of co-enriched MS2-circRNAs using high-throughput sequencing. Immunoprecipitation = IP. **b,** Two MS2-

circRNA architectures containing MS2 RNA hairpins and degenerate 16 nucleotide barcodes. **c,** In-gel fluorescence of total RNA extracted from cells expressing MS2-circRNAs 1 or 2 and stained with DFHBI-1T (n=1). Molecular weight markers have units of nucleotides. **d**, Enrichment of tdMCP-Lck protein fusions complexed with MS2-circRNAs 1 or 2 with the immobilized ATP-competitive inhibitor dasatinib. FT = flow through of enrichment. + = presence of free dasatinib competitor (n=3 independent experiments). Molecular weight markers have units of kDa and nucleotides for protein and RNA gels, respectively.

High fidelity protein variant-barcode associations

A key requirement of LABEL-seq is the intracellular assembly of tdMCP-tagged protein fusions with their cognate MS2-circRNAs. To ensure the correct association of protein variants with their prescribed barcode sequences, we engineered a HEK293 landing pad line[24], circRNA-Protein Landing Pad (RPLP), that facilitates efficient integration of a single protein variant and a single MS2-circRNA barcode sequence per cell in a pooled format (Extended Data Fig. 2a). The RPLP contains a Bxb1 recombination site at a single genomic locus that is flanked by Tet-inducible and U6 promoters (Extended Data Fig. 2b), thereby limiting integration to a single transgenic cassette containing a protein variant and MS2-circRNA barcode sequence pair that are only expressed after recombination. We next recombined the RPLP with plasmids containing a MS2-circRNA 1b library and a Flag-tagged tdMCP-BRaf fusion and found barcode sequence frequencies were robustly correlated between two replicate immunoprecipitations (Pearson's R^2 = 0.93, Fig. 2a). Measurements from immunoprecipitations of a different protein kinase, tdMCP-BTK, co-expressed with a library of MS2-circRNA barcode sequences were also highly reproducible (Pearson's R^2 = 0.98, Extended Data Fig 2d). Immunoprecipitations from as few as 25 RPLP cells per barcode sequence were well-correlated (Pearson's R^2 = 0.80; Extended Data Fig. 2c), suggesting that a 15-cm dish can provide reproducible measurement of >800,000 barcode sequences.

Barcode sequences must remain stably associated with their cognate protein following lysis and during enrichment because dissociation and reassociation of protein:MS2-circRNA complexes would diminish measurement precision. Therefore, we measured the extent of MS2-circRNA barcode sequence misassociation following enrichment (Fig. 2b). We separately integrated a MS2-circRNA 1b (Extended Data Fig. 1b) barcode sequence library encoding a Flag-tagged tdMCP-BRaf fusion and an orthogonal MS2-circRNA 1b barcode sequence library encoding a Myc-tagged tdMCP-BRaf fusion into the RPLP. We then combined the two libraries, performed parallel Myc and Flag immunoprecipitations, and sequenced the barcodes from each. Only ~1% of MS2-circRNAs enriched from the Flag immunoprecipitation were associated with Myc-tagged tdMCP-BRaf (Fig. 2c,d). Notably, the frequency of MS2-circRNAs associated with Myc-tagged tdMCP-BRaf did not increase over time, suggesting that minimal barcode exchange occurred during enrichment. Furthermore, we observed that when an excess of a barcodeless MS2-circRNA that cannot be reverse transcribed ("dummy RNA", Extended Data Fig. 2e) was added to the lysis buffer, only ~0.3% of quantified MS2-circRNAs encoded Myc-tagged tdMCP-BRaf after 105 min (Fig. 2c,d). Comparably low off-target Flag-tdMCP-BRaf barcode sequence levels were enriched following Myc immunoprecipitation (Fig. 2d and Extended Data Fig. 2f,g). Together, MS2-circRNAs efficiently assemble with co-expressed tdMCP-protein fusions and protein variant-barcode sequence associations are maintained during prolonged enrichments, enabling variant library analysis at scale.

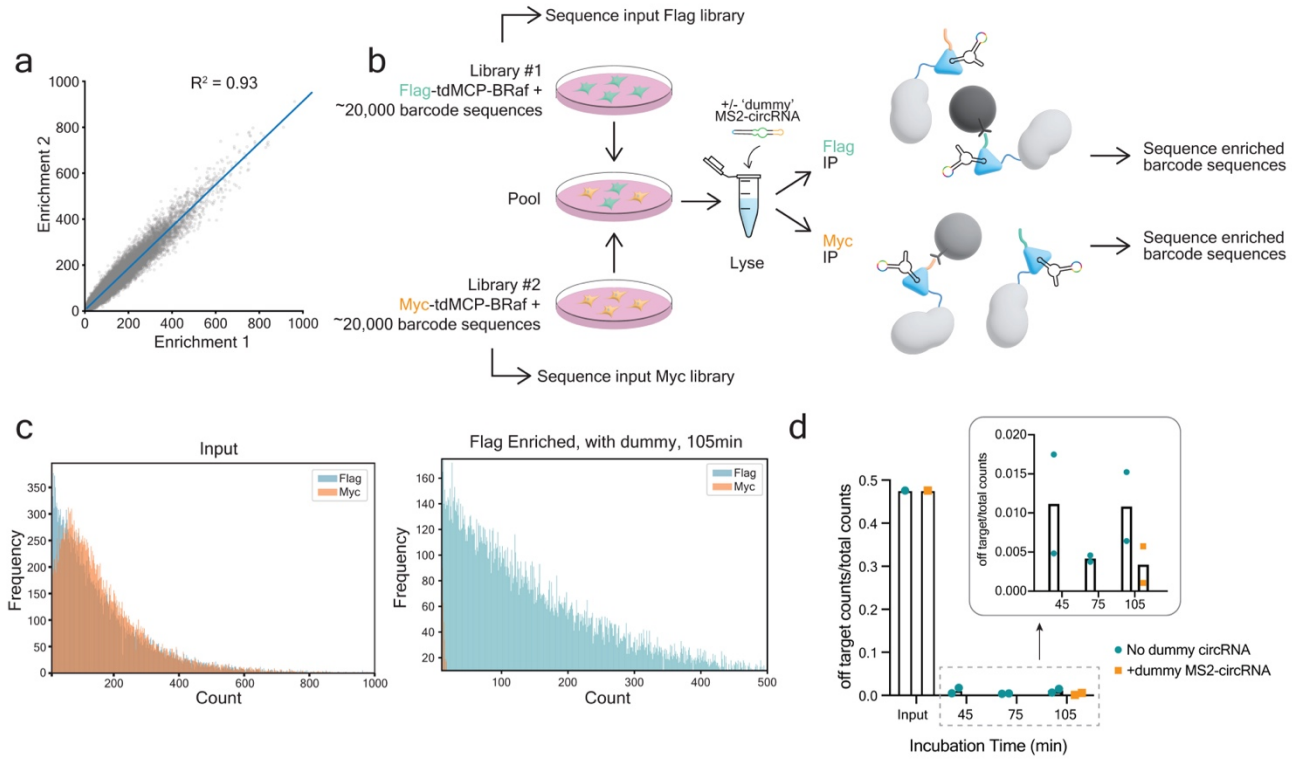


Fig. 2 | Protein:MS2-circRNA complexes are stable during biochemical enrichment. a, Sequenced barcode counts derived from two independent Flag immunoprecipitations of Flag-tdMCP-BRaf co-expressed with a library of MS2-circRNA 1b barcode sequences (Pearson's R^2 = 0.93). **b,** Experimental workflow of the barcode sequence misassociation experiment. Dummy MS2-circRNA does not contain a barcode or the priming site required for reverse transcription. **c,** Results from the experiment shown in **b**. Relative frequencies of barcode sequence counts matched to Flag-tdMCP-BRaf and Myc-tdMCP-BRaf libraries in unenriched lysate (left) and after 105 min Flag immunoprecipitation performed with lysate that was generated with dummy MS2-circRNA in the lysis buffer (right). **d,** Quantification of the data shown in **c**. The y-axis is the fraction of barcode sequence counts encoding the Myc-tagged library (n=2 expression, lysis, enrichment, and sequencing replicates).

Profiling of protein variant abundance with LABEL-seq

We selected the kinase BRaf to demonstrate the profiling capabilities of LABEL-seq due to its complex regulation, ability to interact with multiple signaling partners, and the prevalence of disease-associated mutations[25]. Confirming that an N-terminal tdMCP fusion[26,27] of BRaf behaves

like untagged BRaf, we observed that previously characterized BRaf variants promoted expected levels of pMek (Fig. 3a), co-expression of CRaf enhanced tdMCP-BRaf's activation of downstream signaling (Extended Data Fig. 3a), and tdMCP-BRaf possessed similar levels of phosphorylation at a regulatory site as untagged BRaf (Extended Data Fig. 3b). In order to understand how sites that have been observed to be mutated in cancer affect the signaling properties of BRaf, we generated a library of single amino acid variants at all positions in BRaf that contain two or more missense mutations in TCGA and COSMIC databases (Supplementary Table 1). Two positions (R509 and I666) that contain biochemically well-characterized variants were also varied, for a total of 80 positions (Fig. 3b and Extended Data Fig. 3c). Long read sequencing of the ~1,600 member tdMCP-BRaf variant library indicated each variant was represented by ~30 unique MS2-circRNA 1b barcode sequences (Extended Data Fig. 3d), which was recombined into the RPLP.

Quantitative protein variant abundance measurements are important because proteins must express at high enough levels to perform their cellular function[10,28,29] and many biochemical properties can only be fully understood with protein stability information[30]. To enable the measurement of a range of protein abundances, we developed a LABEL-seq abundance assay that involves co-expressing Flag-tdMCP-tagged protein variants with a Myc-tdMCP standard, which serves as an intracellular competitor for MS2-circRNAs (Fig. 3c), and calculation of barcode sequence frequency ratios obtained from parallel Flag and Myc immunoprecipitations. This ratiometric quantification provides a readout of abundance that does not depend on the relative intracellular expression levels of MS2-circRNAs and protein variants. We confirmed that the sequencing-derived abundance scores of cells co-expressing the Myc-tdMCP standard with tdMCP-BRaf WT or G466V, and treated with a BRaf variant-selective PrOteolysis TArgeting

Chimera (PROTAC)[31], reflected levels determined by western blot analysis (Fig. 3d and Extended Data Fig. 3e).

We next used the LABEL-seq abundance assay to profile the ~1600 tdMCP-BRaf variant library (Fig. 3e and Supplementary Table 2). We found that abundance scores were highly correlated (Pearson's R^2=0.75, Extended Data Fig. 3f) between biological replicates, and that the synonymous WT and nonsense score distributions were well separated (Fig. 3f). As expected, we observed that proline and charged amino acid substitutions most negatively impacted BRaf abundance (Extended Data Fig. 3g)[29]. LABEL-seq-calculated abundance scores of individual BRaf variants were highly correlated with variant levels assessed by western blot analysis (Pearson's R^2=0.95, Fig. 3g), demonstrating LABEL-seq can accurately and precisely measure protein variant abundance at scale.

Our abundance profiling shows that most profiled positions of BRaf were highly tolerant of substitution, with 86.8% of missense variants demonstrating WT-like abundance (Fig. 3f and Extended Data Fig. 3g). This permissiveness is notable based upon previous observations where up to 50% of substitutions are not tolerated in some proteins[28], but is consistent with their location on BRaf. Profiled regions of BRaf include the linker connecting the cysteine-rich domain (CRD) to the kinase domain, the kinase domain's phosphate-binding loop (P-loop), activation loop, and C-terminal tail, which are all flexible and structurally dynamic (Extended Data Fig. 3h,i). Many of the positions profiled also undergo large conformational transitions between regulatory states (Fig. 3h)[32, 33]. Calculation of position-averaged abundance scores showed that select positions are relatively intolerant to substitution (Fig. 3h). These positions are situated in regions that are more structured than other profiled positions, with four of the five lowest average abundance scores being constituents of β-strands despite only 12.5% of all profiled positions being classified as β-

stranded (Extended Data Fig. 3i). The two least tolerant positions to substitution, I582 and I592, are located on adjacent β-strands, make hydrophobic contacts, and are structurally invariant in the autoinhibited and active forms of BRaf (Fig. 3h and Extended Data Fig. 3j). Our results are consistent with systematic analyses showing that disease-relevant mutations commonly alter the biochemical properties of a protein without dramatically affecting its folding and stability,[34,35] and suggest that most sites of somatic cancer mutations in BRaf can readily accumulate gain-of-function activities because substitutions at these positions minimally affect abundance.

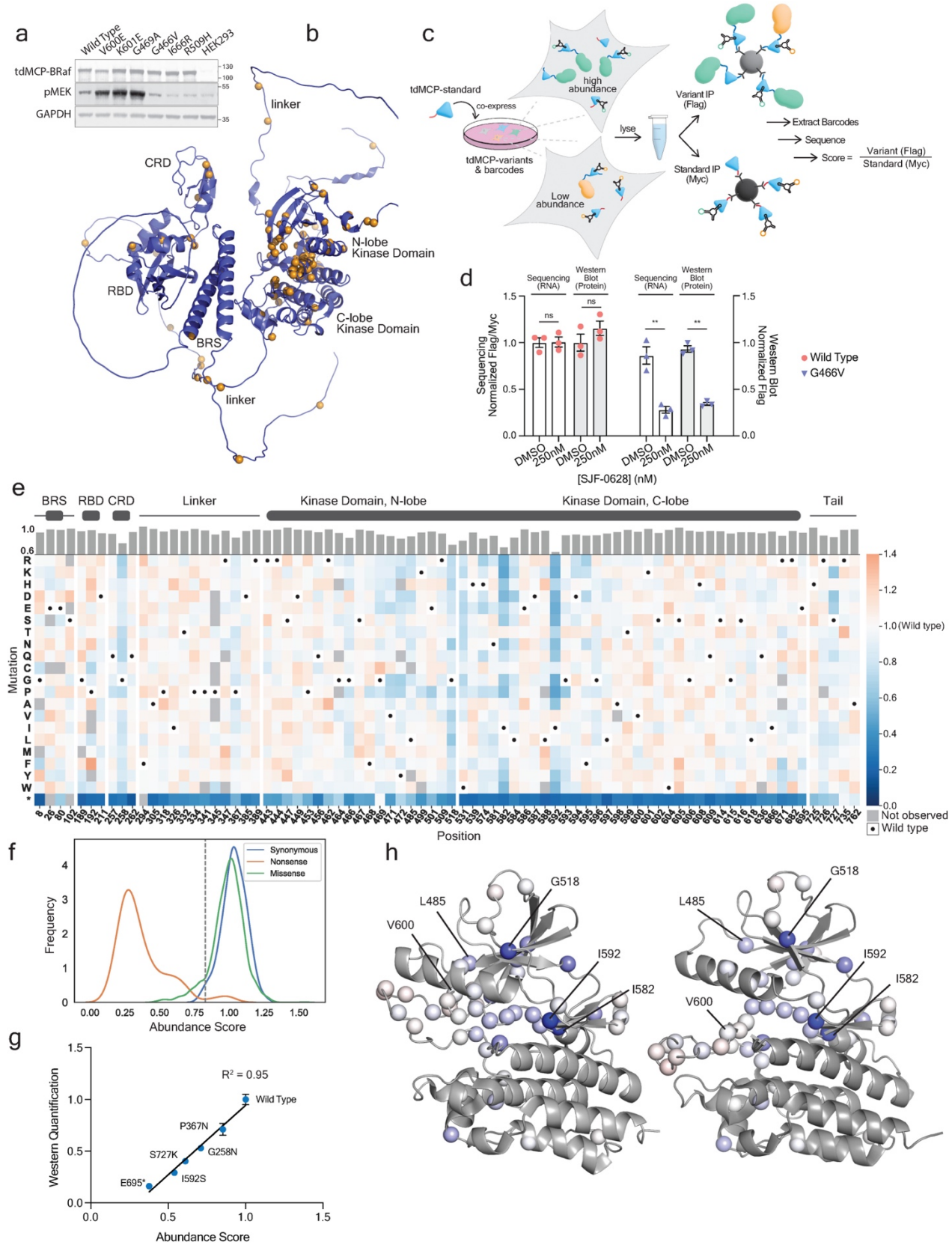

a
Wild Type
V600E
K601E
G469A
G466V
I666R
R509H
HEK293
tdMCP-BRaf
pMEK
GAPDH
130
100
55
35
b
linker
CRD
N-lobe
Kinase Domain
RBD
C-lobe
Kinase Domain
BRS
linker
c
tdMCP-standard
co-express
high abundance
tdMCP-variants & barcodes
lyse
Low abundance
Variant IP (Flag)
Standard IP (Myc)
Extract Barcodes
Sequence
Score = Variant (Flag) / Standard (Myc)
d
Sequencing (RNA)
Western Blot (Protein)
ns
**
Sequencing Normalized Flag/Myc
Western Blot Normalized Flag
Wild Type
G466V
DMSO
250nM
[SJF-0628] (nM)
e
BRS
RBD
CRD
Linker
Kinase Domain, N-lobe
Kinase Domain, C-lobe
Tail
Mutation
Position
Not observed
Wild type
(Wild type)
f
Synonymous
Nonsense
Missense
Frequency
Abundance Score
g
R² = 0.95
Wild Type
P367N
G258N
S727K
I592S
E695*
Western Quantification
Abundance Score
h
G518
L485
V600
I592
I582

Fig. 3 | Multiplexed abundance measurements reveal BRaf's tolerance of substitution at positions where somatic mutations are frequently observed. a, Western blot of RPLP cells expressing different tdMCP-BRaf variants (n=1). **b,** AlphaFold2 model (AF-15056-F1) of full-length human BRaf. Orange spheres represent varied positions. **c,** Schematic depicting the LABEL-seq abundance assay where a tdMCP standard (Myc-tdMCP-SNAP) and Flag-tdMCP-protein variants are co-expressed. Abundance scores equal WT-normalized ratios of variant barcode sequence frequencies from parallel Flag and Myc immunoprecipitations (IP). **d,** Abundance measurements obtained by western blotting (right y-axis = Flag/Myc signals normalized to DMSO-treated, Flag-tdMCP-BRaf WT-expressing cells) and high-throughput sequencing of co-enriched MS2-circRNAs (left y-axis = ratio of barcode sequence frequencies (each replicate shows the mean of 4 barcode sequence ratios) from parallel Flag and Myc immunoprecipitations normalized to DMSO-treated, Flag-tdMCP-BRaf WT-expressing cells) for cells expressing Flag-tdMCP-BRaf WT or G466V and treated with DMSO or SJF0628 for 24 h.Bars equal the mean of n=3 replicates and error bars represent SEM. Unpaired two-sided t-test without multiple comparison adjustment, P values (left to right): 0.922, 0.263, 0.0033, 0.001. Western blots shown in Extended Data 3e. **e,** Sequence-abundance map for tdMCP-BRaf library variants. Abundance scores equal the mean of two replicates, each involving an independent Myc-tdMCP-SNAP standard transduction, immunoprecipitations, and barcode sequence quantification. Black dots indicate the WT amino acid and gray tiles indicate missing data. Bar graphs show position-averaged abundance scores. **f,** LABEL-seq abundance score distributions for all synonymous WT and nonsynonymous BRaf variants. The gray dashed line equals the abundance score we defined as decreased abundance (>2 standard deviations below the mean of the synonymous WT distribution). **g,** Individually assessed variant abundances measured by western blotting (values shown equal the mean of n=3 replicates and error bars represent mean +/- SEM) compared to abundance scores determined with the LABEL-seq abundance assay (Pearson's R^2 = 0.95). **h,** Position-averaged abundance scores for each position in the tdMCP-BRaf variant library (represented as spheres) projected onto the structures of (left) autoinhibited (PDB ID: 6NYB) and (right) active BRaf (PDB ID: 4MNE). The shade of blue indicates position-averaged abundance score: white = 1, darkest blue = 0.72.

LABEL-seq profiling of intracellular BRaf variant activity

We next configured LABEL-seq to measure intracellular enzymatic activity by tagging a protein substrate with tdMCP. Co-expression of the tdMCP-tagged substrate, untagged protein variants and variant-encoding MS2-circRNAs in the RPLP followed by selective capture of tdMCP-tagged products enriches barcode sequences based on the activity of the encoded variant (Fig. 4a). To provide a Raf-specific activity readout, we used Flag-tdMCP-Erk2 (pErk reporter), a

substrate that is phosphorylated as a direct result of Raf activity (Fig. 4b), and phospho-Erk immunoprecipitation to capture the MS2-circRNA-bound pErk reporter. We first confirmed that the pErk reporter was efficiently captured (Extended Data Fig. 4a) and that cells expressing hyperactive BRaf V600E yielded higher co-enriched barcode frequencies compared to BRaf WT (Fig. 4c). Next, we profiled an untagged BRaf variant library with the LABEL-seq activity assay (Extended Data Fig. 3d) in RPLP cells (Fig. 4d,e). Because BRaf variants are expressed at higher levels than endogenous BRaf in the RPLP cell line and many BRaf variants, including WT, signal as heterodimers with CRaf[34], CRaf was co-overexpressed to ensure appreciable levels of both isoforms. Activity scores were calculated using the ratio of WT-normalized barcode sequence frequencies derived from parallel pErk and Flag (total tdMCP-Erk2) immunoprecipitations (Fig. 4d and Supplementary Table 2). Biological replicates provided well correlated activity scores (Pearson's $R^2 = 0.80$, Fig. 4f and Extended Data Fig. 4b,c) and BRaf variants previously shown to promote elevated pErk[34-36] had significantly higher activity scores than WT (Extended Data Fig. 4c). Activity scores of individual BRaf variants were highly correlated with western blot-derived pErk levels (Pearson's $R^2 = 0.95$, Fig. 4g). We classified variants as activating or inactivating (Extended Data Fig. 4d,e), and found that 32.8% and 2.3% were activating and inactivating, respectively (Extended Data Fig. 4e-g). The small percentage of inactivating variants is consistent with BRaf WT existing primarily as an autoinhibited monomer in unstimulated HEK293 cells. Observed activating variants are broadly distributed across BRaf and a number of positions contained multiple activating mutations.

Because BRaf WT mainly exists as an inactive monomer stabilized by a network of interactions in unstimulated cells, we reasoned that multiple activating substitutions should be observed at positions involved in autoinhibition. We identified 17 positions where >75% of

substitutions were classified as activating (Supplementary Table 3). Two of these positions–G258 and P367–are in BRaf's N-terminal CRD and RBD domains, respectively, and substitutions at these positions likely disrupt autoinhibitory protein-protein interactions (Extended Data Fig. 4h,i). Nine other positions form a spatial cluster in the kinase domain of autoinhibited but not active BRaf (Fig. 4h), and the side chains of four of these positions within the activation loop (L597, T599, V600 and K601) form interactions with three P-loop positions (S467, F468, and G469) and L485. While W604 lacks obvious contacts with these positions, its proximity and high percentage of activating substitutions suggests it is part of this autoinhibitory network. Our data are consistent with structural studies suggesting that eight of these nine positions stabilize an "autoinhibitory turn"[36-38], and confirm that this region of BRaf's kinase domain forms a finely tuned regulatory network that mutations can readily disrupt to achieve increased activity.

To understand how heterodimerization affects BRaf variant activity, we repeated the LABEL-seq activity assay in the absence of overexpressed CRaf and calculated differences in activity scores with and without CRaf (Fig. 4i and Extended Data Fig. 5a-e). We observed that seven of the ten positions with the highest average CRaf-dependent activity scores are highly conserved across the human kinome[39] and are largely clustered in and around the ATP-binding pocket (Fig. 4j and Extended Data Fig. 5f). Among these positions, mutations at G466, N581, D594, and G596 have previously been demonstrated to impair BRaf's kinase activity[34,36] while increasing cellular pErk levels by promoting transactivation of CRaf through heterodimerization[34]. Two other positions with the highest average CRaf-dependent activity, G615 and S616, are located at the BRaf:Mek interface (Fig. 4j). Substitutions at these positions likely disrupt BRaf's interactions with its substrate but not its ability to transactivate CRaf. Together, our data show that

disruption of BRaf's catalytic machinery through mutation is a general mechanism for promoting downstream signaling through heterodimerization with CRaf.

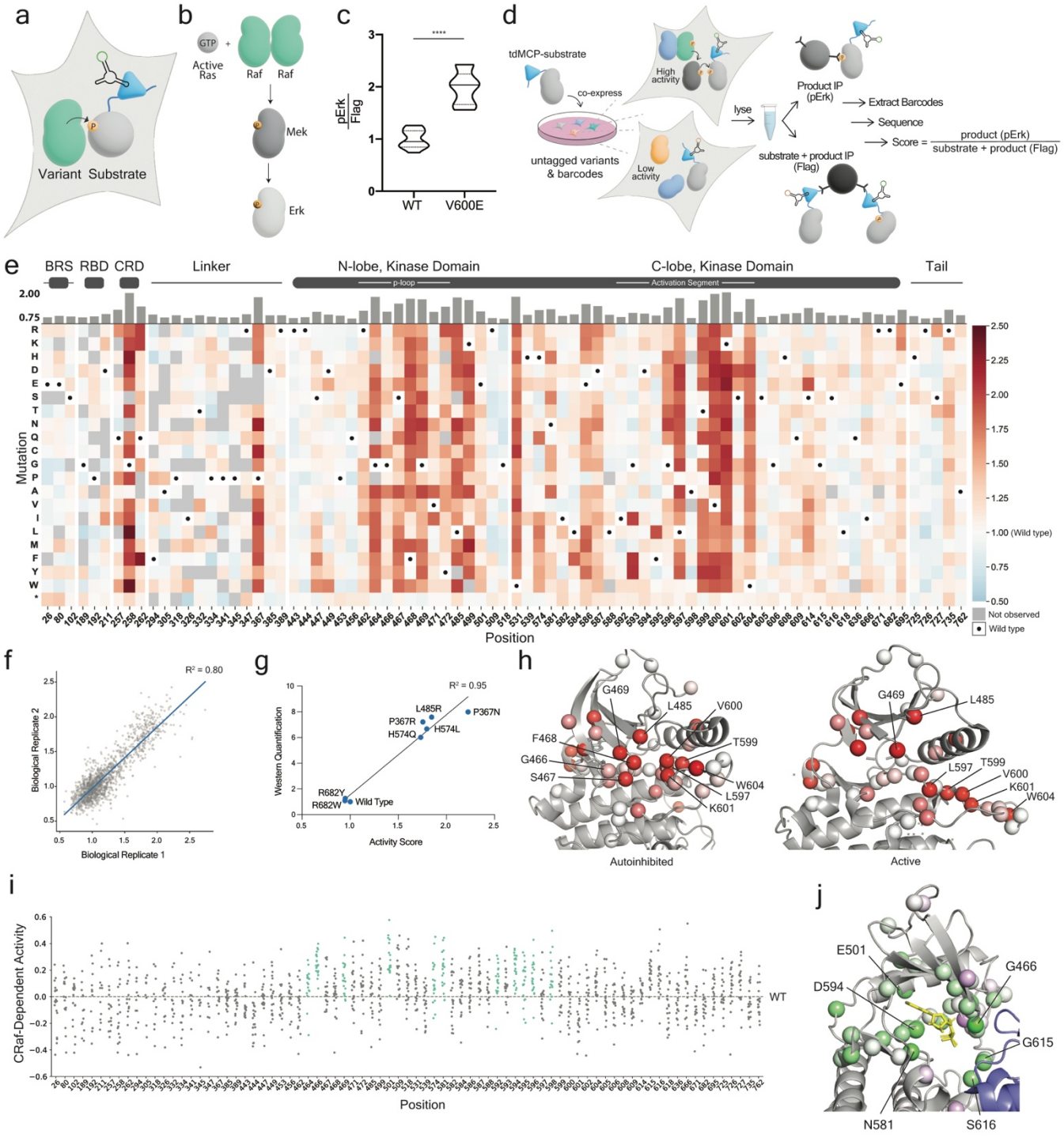


Fig. 4 | Profiling of intracellular BRaf variant activity with a pErk reporter. a, LABEL-seq activity assay. **b,** Raf-Mek-Erk signaling pathway. **c,** Activities (represented as pErk/Flag barcode sequence frequencies) of BRaf WT and V600E in HEK293T cells co-expressing the pErk reporter

and MS2-circRNAs. pErk/Flag frequencies for each co-enriched MS2-circRNA barcode sequence was calculated following parallel pErk and Flag immunoprecipitations. Unpaired two-sided t-test, P value < 0.0001. Center line represents the mean and dotted lines represent the 25 and 75 percentiles. n=10-11 barcodes/variant. **d,** LABEL-seq assay for the multiplexed measurement of untagged BRaf variant activity with the pErk reporter. **e,** Sequence-activity map for the untagged BRaf variant library co-expressed with CRaf. Activity scores are the average of two replicates, each involving independent cell culturing, lysis, immunoprecipitations, and high-throughput sequencing quantification. Black dots indicate the WT amino acid and gray tiles indicate missing data. Bar graphs show position-averaged activity scores. **f,** Correlation of activity score replicates represented in **e** (Pearson's R^2 = 0.80). **g,** Correlation between LABEL-seq activity scores and pErk levels measured by western blotting (values shown equal mean of n=3 replicates, Pearson's R^2 = 0.95). **h,** Number of activating variants at each position (spheres) in the untagged BRaf variant library projected onto the (left) inactive (PDB ID: 6NYB) and (right) active conformations (PDB ID: 4MNE) of BRaf's kinase domain. The shade of red represents the number of activating substitutions: white = 0, brightest red = 19. **i,** CRaf-dependent activity of BRaf variants. Each dot represents the difference in activity score for a BRaf variant with and without CRaf overexpression. Difference scores calculated by determining the difference in BRaf variant activity score with and without CRaf overexpression, normalized to BRaf WT under the same condition. Positions highly conserved amongst protein kinases are shown in green. Gray dashed line indicates the WT score. Positive values indicate CRaf-promoted activity. **j,** Position-averaged CRaf-dependent activity scores projected onto the autoinhibited BRaf:Mek complex (PDB ID: 6NYB) bound to ATP-γ-S. Purple spheres are scores <0 (decreased activity with CRaf) and green spheres are scores >0 (increased activity with CRaf).

LABEL-seq profiling of BRaf's intracellular interactions

Multiple intracellular protein-protein interactions modulate BRaf's activity. In the absence of active Ras, BRaf WT mainly exists as an autoinhibited monomer within a complex containing its substrate, Mek, and two 14-3-3 proteins. Ras activation recruits active homodimeric and heterodimeric BRaf complexes. To study how sequence variation affects BRaf's intracellular interactions, we developed a LABEL-seq interaction assay using TurboID proximity labeling (Fig. 5a and Extended Data Fig. 6a-e). In this interaction assay, tdMCP-tagged protein variants and variant-encoding MS2-circRNAs are co-expressed with a TurboID-interaction partner fusion. Intracellular biotin labeling of tdMCP-tagged protein variants by the TurboID-interactor fusion

followed by post-lysis streptavidin enrichment facilitates quantification of intracellular interactions. We first confirmed that TurboID labeling could reliably measure BRaf's interactions by co-expressing TurboID-Mek1 with tdMCP-BRaf WT or I666R, a BRaf:Mek-disrupting variant, and treating cells with biotin. We observed that that streptavidin enrichment followed by barcode sequence quantification recapitulated I666R's reduced biotinylation levels relative to WT (Fig. 5b,c). Both tdMCP-BRaf variants showed markedly higher signal than an unrelated cytosolic kinase (tdMCP-BTK), which we used to measure non-specific labeling (Fig. 5c). We further found that the LABEL-seq interaction assay provided expected increases in tdMCP-BRaf WT labeling by TurboID-CRaf or TurboID-BRaf in cells treated with a dimer-promoting inhibitor (Fig. 5d-f).

We first used our LABEL-seq interaction assay to profile BRaf:CRaf heterodimerization by co-expressing TurboID-CRaf with the Flag-tdMCP-BRaf variant library, labeling cells with biotin, and calculating CRaf interaction scores using the ratio of WT-normalized barcode sequence frequencies derived from parallel streptavidin and Flag (total tdMCP-BRaf variant) enrichments (Fig. 5g,h and Supplementary Table 1). Interaction scores were well correlated between biological replicates (Pearson's $R^2 = 0.96$, Extended Data 7b) and biotin labeling times (Pearson's $R^2 = 0.86$, Extended Data Fig. 7a). We further found that BRaf WT's interaction score was markedly higher than the non-specific labeling control, tdMCP-BTK, demonstrating that LABEL-seq's highly sensitive readout allows the measurement of low basal interactions that are difficult to capture with other approaches such as bioluminescence resonance energy transfer[26] and co-immunoprecipitation[34]. BRaf variants previously shown to exhibit elevated heterodimerization with CRaf[34] had interaction scores significantly above BRaf WT (Extended Data Fig. 7c). We next classified variants as forming increased or decreased CRaf interactions (Extended Data Fig. 7d,e) and observed that there are a number of positions distributed throughout BRaf where a majority of

variants were classified as increased CRaf interaction (Fig. 5i), suggesting that they participate in specific interactions that stabilize monomeric BRaf. Notably, nonsense mutations that truncate BRaf C-terminal to the RBD and CRD provided some of the highest interaction scores, confirming that BRaf's CRD can act in trans with CRaf's kinase domain[40]. Given that BRaf WT signals as a dimer, we were curious whether mutations that promote CRaf heterodimerization are generally activating. Although we did not find a strong correlation between activity and CRaf interaction scores (Extended Data Fig. 7f), 82% of the missense variants classified as activating were also classified as increased CRaf interaction. Together, our CRaf interaction data support a model where BRaf WT exists primarily as an autoinhibited monomer, which numerous somatic mutations are capable of releasing to allow CRaf heterodimerization.

We next profiled the interactions of BRaf with its substrate by co-expressing TurboID-Mek1 with the tdMCP-BRaf variant library, labeling cells with biotin (Fig. 5j and Supplementary Table 1), and calculating Mek1 interaction scores (Pearson's $R^2 = 0.80$, Extended Data Fig. 7g-i). BRaf I666R's and the non-specific labeling control's interaction scores were both significantly lower than BRaf WT (Extended Data Fig. 7j), and Mek1 interaction scores were consistent with levels of Mek1 co-immunoprecipitated with individual BRaf variants (Extended Data Fig. 7k). Unlike CRaf, a larger percentage of BRaf variants demonstrated decreased rather than increased interaction with Mek1 (Extended Data Fig. 7h), consistent with autoinhibited BRaf WT forming a stable complex with Mek. The lowest interaction scores were observed at positions directly lining the BRaf:Mek interface and nonsense variants that truncate BRaf's kinase domain (Fig. 5j and Extended Data Fig. 7i). Notably, 84% of variants at positions that don't line the BRaf:Mek interface were classified as both decreased Mek and increased CRaf interaction, suggesting that weakening Mek interactions allows heterodimerization with CRaf.

To obtain a structural understanding of which positions contribute to BRaf's interactions with CRaf and Mek1, we projected position-averaged interaction scores onto a structure of autoinhibited BRaf bound to an ATP analogue (Fig. 5k,l). Consistent with ATP interactions stabilizing monomeric BRaf[37], we observed that most positions that line the ATP-binding pocket possessed high position-averaged CRaf interaction scores. Intriguingly, positions with low average Mek1 interaction scores also line the ATP-binding pocket, suggesting that disruption of BRaf's interactions with ATP weakens a BRaf:Mek autoinhibitory complex (Fig. 5l and Extended Data Fig. 7l,m). Together, our data highlight a unique feature of BRaf regulation where ATP, Mek, and monomeric BRaf form a stable autoinhibitory complex that can readily be disrupted through mutation to allow BRaf:CRaf heterodimerization.

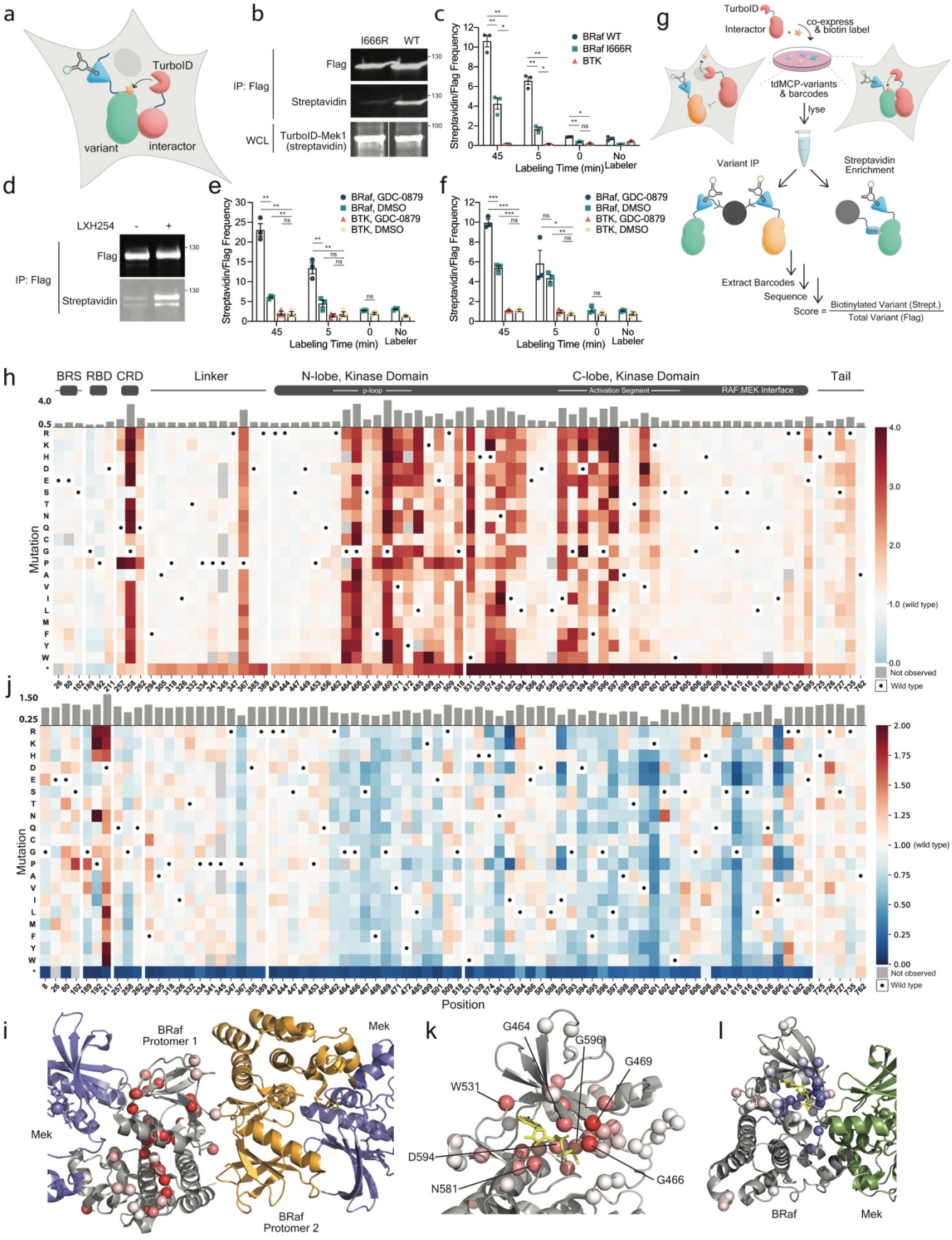

a
TurboID
variant
interactor
b
I666R
WT
IP: Flag
Flag
Streptavidin
WCL
TurboID-Mek1 (streptavidin)
c
BRaf WT
BRaf I666R
BTK
Streptavidin/Flag Frequency
Labeling Time (min)
No Labeler
d
LXH254
IP: Flag
Flag
Streptavidin
e
BRaf, GDC-0879
BRaf, DMSO
BTK, GDC-0879
BTK, DMSO
f
g
TurboID
Interactor
co-express & biotin label
tdMCP-variants & barcodes
lyse
Variant IP
Streptavidin Enrichment
Extract Barcodes
Sequence
Score = Biotinylated Variant (Strept.) / Total Variant (Flag)
h
BRS RBD CRD
Linker
N-lobe, Kinase Domain
C-lobe, Kinase Domain
Tail
p-loop
Activation Segment
RAF:MEK Interface
Mutation
Not observed
Wild type
j
Position
i
BRaf Protomer 1
BRaf Protomer 2
Mek
k
G464
G596
G469
W531
D594
N581
G466
l
BRaf
Mek

Fig. 5 | LABEL-seq measurements of BRaf's intracellular interactions with CRaf and Mek1. **a,** Proximity-based, LABEL-seq interaction assay. **b,** Flag-tdMCP-BRaf I666R and WT biotinylation levels in HEK293 cells co-expressing TurboID-Mek1 after biotin labeling. **c**, Flag-tdMCP-BRaf WT, I666R and Flag-tdMCP-BTK biotinylation levels in HEK293T cells co-expressing TurboID-Mek1 and MS2-circRNAs. Following biotinylation, cells were lysed, pooled, parallel enrichments were performed, and co-enriched barcodes were quantified (barcode sequence frequency ratios) by high-throughput sequencing. Bar = mean of n=3 replicates (n=3 barcodes/replicate), error bars represent SEM. Two-sided unpaired Welch's t-test without multiple comparison correction, P values (left to right) = 0.0031, 0.0014, 0.019, 0.0037, 0.0019, 0.013, 0.010, 0.0026, 0.176. **d**, Flag-tdMCP-BRaf WT biotinylation levels in DMSO- or LXH-254-treated HEK293T cells co-expressing TurboID-CRaf and labeled with biotin. **e,f** Flag-tdMCP-BRaf WT and Flag-tdMCP-BTK biotinylation levels in DMSO- or GDC-0879-treated HEK293T cells co-expressing TurboID-CRaf (**e**) or TurboID-BRaf (**f**) and MS2-circRNAs after biotin labeling. Quantification and analysis as in **c**. P values (left to right): (**e**) 0.0078, 0.0031, 0.0065, 0.879, 0.0089, 0.0094, 0.0598, 0.623, 0.138, (**f**) 0.0004, 0.0004, 0.0009, 0.989, 0.381, 0.059, 0.0045, 0.235, 0.189. **G**, LABEL-seq measurement of intracellular BRaf variant interactions. **h,** Sequence-CRaf interaction map for tdMCP-tagged BRaf variants co-expressed with TurboID-CRaf. Interaction scores are the mean of two replicates, each involving independent cell culturing, 5 min biotin labelings, enrichments, and quantifications. Black dots indicate the WT amino acid and gray tiles indicate missing data. Bar graphs show position-averaged scores. **i,** Number of variants at each position classified as increased CRaf interaction in **h** projected onto active BRaf's kinase domain (PDB: 4MNE). White = 0, brightest red = 19. **j,** Sequence-Mek1 interaction map for tdMCP-tagged BRaf variants co-expressed with TurboID-Mek1. Interaction scores are the average of two replicates, each involving independent transfections, 5 min biotin labelings, enrichments, and quantifications. Black dots indicate the WT amino acid and gray tiles indicate missing data. Bar graphs show position-averaged scores. **k,l** Position-averaged CRaf (**k**) and Mek1 (**l**) interaction scores projected onto autoinhibited BRaf bound to ATP-γ-S (PDB ID: 6NYB). The shade of red (**k**) indicates the position-averaged CRaf interaction score (white = 1, red >3.2) and the shade of blue (**l**) indicates the position-averaged Mek1 interaction score (white = 1, blue <0.5).

Identification of BRaf positions with similar biochemotypes

Our profiling data indicate substitutions at certain positions in BRaf result in dramatic changes to multiple biochemical properties, but we were unable to identify positions with similar functions or gain a mechanistic understanding of the positional effects of substitution using pairwise analyses (Extended Data Fig. 8a). Therefore, we performed hierarchical clustering based on the six biochemical properties measured with LABEL-seq to group positions into nine distinct

biochemical phenotypes (biochemotypes) (Fig. 6a,b). Clusters 7–9, encompassing 64% of profiled positions, represent a biochemotype in which most substitutions exhibited modest effects on all measured biochemical properties. In contrast, cluster 1 positions are defined by a biochemotype where substitutions resulted in some of the largest decreases in Mek interactions and largest increases in activity and CRaf heterodimerization, with minimal CRaf-dependent activity (Fig. 6a,c). One cluster 1 position–P367–is proximal to the autoinhibitory phosphorylation site, S365[41], which recruits 14-3-3 proteins (Extended Data Fig. 4i). Substitutions at P367 likely disrupt the autoinhibited BRaf complex and allow formation of active Raf dimers by disrupting 14-3-3 protein interactions[42]. The six other cluster 1 positions are located in two different regions of BRaf's ATP-binding site (Fig. 6c). Four of these positions (F468, L485, L597 and V600) are components of the autoinhibitory turn (Fig. 4h) and are spatially clustered in autoinhibited but not active BRaf. Despite their disparate locations, cluster 1 positions appear to share a common functional role in stabilizing the autoinhibited state, which substitutions can disrupt without compromising the ability of BRaf to form catalytically active Raf dimers.

Cluster 2 positions, proximal to cluster 1 autoinhibitory turn positions (Extended Data Fig. 8b), exhibited increased activity and decreased Mek interactions upon substitution like cluster 1, but with smaller increases in CRaf interactions. The ability of substitutions in cluster 2 to increase activity and decrease Mek interactions suggest that these positions also play a role in stabilizing autoinhibited BRaf but with a compromised ability to heterodimerize with CRaf once autoinhibition is released. Positions in clusters 4 and 5 demonstrated a biochemotype where substitutions strongly promoted CRaf heterodimerization and decreased Mek interactions, but, in contrast to clusters 1 and 2, only modestly increased activity, which was highly CRaf dependent. Cluster 4 and 5 positions are located deep within BRaf's active site (Fig. 6b,d) and form a

contiguous spatial network in BRaf's active form (Fig. 6d). Additionally, nearly all positions in clusters 4 and 5 are highly conserved amongst protein kinases[39,43] and many make ATP contacts[37] or are catalytic residues (Extended Data Fig. 5e). The biochemotype of clusters 4 and 5 suggest that these positions participate in the stabilization of autoinhibited BRaf. However, substitutions at these positions likely impair BRaf's kinase activity, rendering cluster 4 and 5 variants dependent on their ability to form strong CRaf heterodimers for Mek activation. Together, we find that a number of positions in BRaf that are frequently mutated in cancer participate in finely tuned autoinhibitory interactions that many substitutions are capable of destabilizing. The signaling properties of these BRaf variants can be classified into several distinct biochemotypes once autoinhibition is disrupted.

LABEL-seq-guided prediction of a complex cellular phenotype

We next determined whether LABEL-seq data could be used to predict which biochemical properties of BRaf variants are capable of driving complex phenotypes such as cell proliferation. Here, we trained logistic models using previously collected data that classified 65 BRaf variants as activating, inactivating, or neutral based on their ability to promote the growth factor-independent proliferation of Ba/F3 and MCF10A cells[44]. Models that used only one biochemical property had modest classification accuracy (balanced accuracy: 0.33-0.57; Fig. 6e). However, a model that used all six biochemical properties measured with LABEL-seq provided the highest accuracy (balanced accuracy = 0.68) and, critically, was the only one capable of correctly identifying all three BRaf variant classifications (Fig. 6e). We applied this six-property classifier to the 239 somatic mutations cataloged in TCGA, COSMIC, and ClinVar and for which we obtained LABEL-seq data, and found that 40.6% and 14.2% were predicted to be activating and

inactivating, respectively (Supplementary Table 1). Notably, 92% of somatic variants previously determined to be activating based on increased intracellular pErk levels or growth factor-independent cell proliferation in separate analyses were correctly classified by our model (Supplementary Table 1). Thus, LABEL-seq's capability of measuring multiple biochemical properties for the same variant enabled prediction of a complex cellular phenotype.

Measurement and modeling of BRaf degradation

We next explored whether LABEL-seq could be used to provide insight into another complex cellular process: small molecule-promoted protein degradation[45]. Specifically, we investigated the Vemurafenib-based PROTAC SJF0628 (Extended Data Fig. 9a), which promotes selective degradation of several BRaf variants[31]. We measured the abundances of all ~1,600 BRaf variants in cells treated with six different SJF0628 concentrations, which were converted into degradation scores (Fig. 6f, Extended Data Fig. 9b,c, and Supplementary Table 2). As expected, BRaf WT was minimally degraded by SJF0628 but a number of variants, including previously characterized SJF0628-sensitive variants[31] (Extended Data Fig. 9d), showed dose-dependent decreases in abundance. Additionally, LABEL-seq-derived degradation scores were consistent with protein levels quantified by western blot (Extended Data Fig. 9c). We classified variants as sensitized or resistant to degradation at the highest concentration of SJF0628 tested and observed that 22.9% of variants were sensitized and 13.5% were resistant (Extended Data Fig. 9e and Supplementary Table 1). Among the 239 somatic mutations for which we measured degradation scores and that are listed in TCGA, COSMIC, or ClinVar, we found that 32.3% were sensitized to degradation relative to BRaf WT and 9.2% were resistant. Notably, variants at the same position

were generally either all sensitized or all resistant, suggesting that susceptibility to SJF0628 degradation is largely defined by the position of a mutation rather than its identity.

We explored the structural basis of BRaf PROTAC sensitivity by projecting position-averaged degradation scores onto the structure of Vemurafenib-bound BRaf (Extended Data Fig. 9f). Several positions within the Vemurafenib-binding pocket generally conferred resistance to degradation when substituted but, surprisingly, some of the most sensitive positions were located on nearby positions without an obvious structural rationale for either property. Although projecting position-averaged degradation scores onto the autoinhibited structure of BRaf bound to an ATP analogue revealed that substitutions at ATP contact residues were largely sensitizing (Fig. 6g), none of the biochemical properties we measured correlated strongly with variant sensitivity (Extended Data Fig. 9g) and positions with the lowest average degradation scores did not fit within any of the biochemotypes we identified (Fig. 6a).

We performed logistic modeling to identify which biochemical properties were most predictive of PROTAC sensitivity. To provide additional properties relevant to ATP-competitive inhibitor binding, we measured CRaf heterodimerization levels promoted by the type 1.5 inhibitor Vemurafenib and the type 2 inhibitor LXH254[46] using the LABEL-seq interaction assay (Extended Data Fig. 10a,b). The resulting eight biochemical properties were used to train a logistic model to predict whether a BRaf variant was WT-like, sensitized, or resistant to the highest concentration of SJF0628. Models using any one property yielded modest classification accuracy (balanced accuracy = 0.33-0.43), but the model using all eight properties was most accurate (balanced accuracy = 0.67, Fig. 6h). Only the model trained using all eight properties could accurately distinguish BRaf variants across all three classifications (Fig. 6h). The properties most heavily weighted in predicting sensitivity to SJF0628 were the CRaf interaction and Type 2 inhibitor

(LXH254)-promoted heterodimerization scores (Fig. 6i), suggesting that dimerization is an important component of degradation. Low basal abundance and Type 2 inhibitor (LXH254)-promoted heterodimerization scores were the most predictive of variant resistance. Thus, the ability of LABEL-seq to profile diverse biochemical properties facilitates the analysis of complex cellular processes that are otherwise difficult to understand using single measurements.

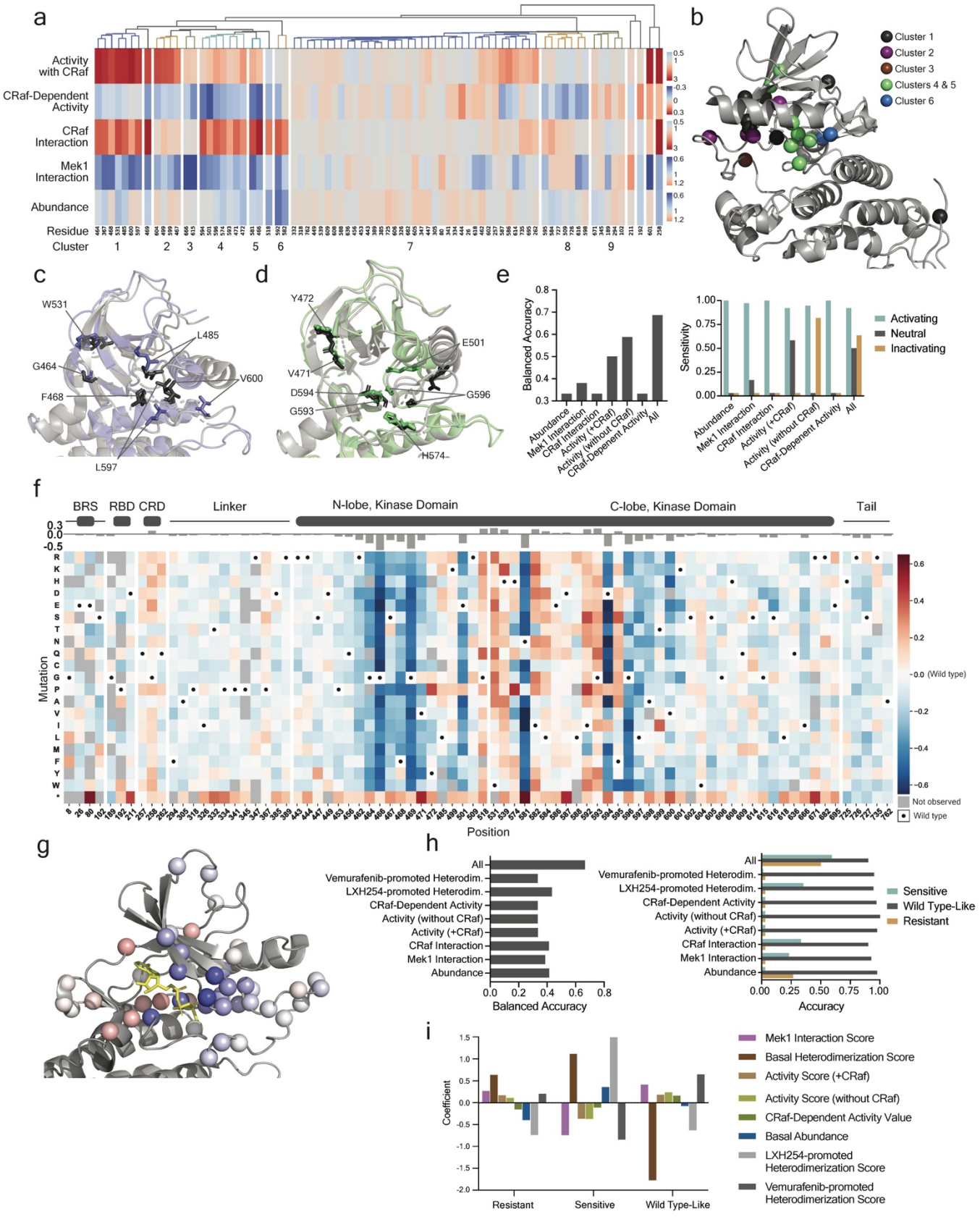

a
Activity with CRaf
CRaf-Dependent Activity
CRaf Interaction
Mek1 Interaction
Abundance
Residue
Cluster
1
2
3
4
5
6
7
8
9
b
Cluster 1
Cluster 2
Cluster 3
Clusters 4 & 5
Cluster 6
c
W531
L485
G464
V600
F468
L597
d
Y472
E501
V471
D594
G596
G593
H574
e
Balanced Accuracy
Abundance
Mek1 Interaction
CRaf Interaction
Activity (+CRaf)
Activity (without CRaf)
CRaf-Depenent Activity
All
Sensitivity
Activating
Neutral
Inactivating
f
BRS
RBD
CRD
Linker
N-lobe, Kinase Domain
C-lobe, Kinase Domain
Tail
Mutation
Position
(Wild type)
Not observed
Wild type
g
h
All
Vemurafenib-promoted Heterodim.
LXH254-promoted Heterodim.
CRaf-Dependent Activity
Activity (without CRaf)
Activity (+CRaf)
CRaf Interaction
Mek1 Interaction
Abundance
Balanced Accuracy
Accuracy
Sensitive
Wild Type-Like
Resistant
i
Coefficient
Resistant
Sensitive
Wild Type-Like
Mek1 Interaction Score
Basal Heterodimerization Score
Activity Score (+CRaf)
Activity Score (without CRaf)
CRaf-Dependent Activity Value
Basal Abundance
LXH254-promoted Heterodimerization Score
Vemurafenib-promoted Heterodimerization Score

Fig. 6 | Integration of diverse biochemical properties. a, Hierarchical clustering of BRaf positions based on six LABEL-seq-derived biochemical properties. Clustering was performed using the average distance metric, which grouped positions based on similarities in response to specific substitutions. Clusters of positions, which were selected based on the dendrogram, are numbered. The heatmap color (scale on the right) indicates the average score for each biochemical property at each position. **b,** Positions within clusters 1-6 mapped onto the kinase domain of the autoinhibited BRaf complex (PDB ID: 6NYB). **c,** Alignment of BRaf's ATP-binding site in the active (purple, PDB ID: 4MNE) and autoinhibited (gray, PDB ID: 6NYB) conformations, with cluster 1 positions shown as sticks. F468 is not resolved in the active conformation. **d,** Alignment of BRaf's ATP-binding site in the active (green, PDB ID: 4MNE) and autoinhibited (gray, PDB ID: 6NYB) conformations, with cluster 4 positions shown as sticks. **e,** Performance of our logistic model in predicting growth factor-independent proliferation of Ba/F3 and MCF10A cells by BRaf variant. (left) Balanced accuracy for models that used a single property versus all properties. (right) Model accuracy for each classification. **f,** Sequence-degradation map for BRaf variants in the tdMCP-BRaf library treated with 2.5 μM SJF0628. Degradation score = [variant abundance score (SJF0628) - variant abundance score (DMSO)]/variant abundance score (DMSO). Degradation scores shown are the average of two replicates, each involving independent SJF0628 treatments, immunoprecipitations, and quantifications. Black dots indicate the WT amino acid and gray tiles indicate missing data. Bar graphs show position-averaged degradation scores. **g,** Position-averaged degradation scores for each position in the tdMCP-BRaf variant library (spheres) projected onto autoinhibited BRaf bound to ATP-γ-S (PDB ID: 6NYB). Color indicates the position-averaged degradation score: white = 0, darkest blue <-0.5, brightest red >0.18. **h,** Performance of our logistic model predicting SJF0628 sensitivity. (left) Balanced accuracy for models that used a single property versus all properties. (right) Model accuracy for each classification. **i,** Property coefficients in the most accurate model.

Discussion

At-scale measurements of protein variant effects have provided valuable insights into protein structure-function[4,8,12]. However, existing MAVEs are generally specific to a single protein property or function, often use indirect readouts, and can be challenging to reconfigure to measure different properties and functions. While DNA, phage, or yeast surface display methods enable measurement of different properties[47,48], they divorce the protein from its cellular context and possess additional constraints such as limited post-translational modifications and protein size. Therefore, we developed LABEL-seq, which enables the direct measurement of diverse

biochemical and functional properties of protein variants in human cells. The simplicity and scalability of LABEL-seq represents a considerable advance compared to existing methods, allowing us to rapidly and inexpensively profile eight biochemical properties of ~1,600 BRaf variants under different cellular conditions for a total of ~20,000 variant effect measurements.

A key strength of LABEL-seq is its modularity, which allows the direct measurement of diverse properties through simple reconfiguration of the protein tagged with tdMCP and/or the affinity enrichment. This enabled profiling of the abundance, activity, protein-protein interactions, and drug sensitivity of BRaf variants. Because any affinity enrichment can be integrated into the LABEL-seq workflow, numerous additional protein properties and functions can be profiled, including subcellular localization, drug binding, and post-translational modification levels, among others. This is in contrast to existing MAVEs which were designed to measure a single property (e.g. protein-protein interactions or abundance). As LABEL-seq only necessitates the fusion of an RBD to a target protein, it can likely be applied to a wide range of proteins beyond the three (BRaf, BTK, and Lck) described in this study.

The sensitivity and dynamic range of LABEL-seq assays is dependent upon multiple factors, including the amount of enriched MS2-circRNA, the selectivity of the affinity enrichment, the number of replicates performed, and the configuration of the assay. We configured the abundance assay to be independent of the amount of tdMCP-tagged protein variant relative to MS2-circRNAs by co-expressing a tdMCP-tagged standard, which establishes an intracellular competition for MS2-circRNAs, and performing affinity enrichments and quantifications of both tdMCP-tagged proteins. This allows the ratiometric measurement of a broad range of protein abundances, including minimally expressing protein variants, because encoding MS2-circRNAs are always captured by the tdMCP-tagged standard. A tdMCP-tagged standard was not included

in the LABEL-seq activity and interaction assays because their lower levels of detection and dynamic range do not depend on the relative stoichiometry of MS2-circRNAs to tdMCP-tagged protein variants. The sensitivity and dynamic range of all LABEL-seq assays is dependent upon multiple factors, including the amount of enriched MS2-circRNA, the number of barcodes per variant, the selectivity of the affinity enrichment, and the number of replicates performed. While the abundance, activity and interaction assays we performed provided highly replicable data with a reasonable dynamic range, additional optimization of assay-specific parameters could improve the sensitivity and dynamic range that can be achieved.

Our ability to use the LABEL-seq abundance assay to measure SJF0628's degradation of BRaf variants allowed us to conduct the first comprehensive analysis of how sequence variation affects a neosubstrate's degradation by a PROTAC. We observed that substitutions at multiple positions in BRaf provided strong sensitization to degradation, with SJF0628 selectively degrading ~30% of cataloged somatic variants relative to BRaf WT. By integrating multiple biochemical measurements obtained with LABEL-seq, we obtained insight into which properties sensitize a BRaf variant to degradation. Although no single property was dominant, the ability of a BRaf variant to be recruited into heterodimers with CRaf was the most predictive of sensitivity to SJF0628-mediated degradation. Integration of all LABEL-seq scores also facilitated prediction of growth factor-independent proliferation. Notably, the accuracy of both predictive models required multiple LABEL-seq measurements, underlining the importance of measuring a range of properties.

Overall, we envision LABEL-seq will find widespread use for the profiling of diverse proteins under different cellular conditions. The multi-dimensional profiling data obtained with this platform will be a valuable resource for understanding intracellular protein function and

druggability, and, when integrated with MAVEs that assess composite behaviors such as transcriptional outputs and proliferation, should provide invaluable insight into protein properties that drive complex cellular processes.

Acknowledgements

We thank David Shechner (UW, Seattle) for guidance on RNA barcode design and Madeline Walsh (UW, Seattle) for feedback on the manuscript. This work was supported by the NIH, grant no. R01GM145011 (NIGMS) and R01GM086858 (NIGMS) to D.J.M. and RM1HG010461 (NHGRI) to D.M.F.

Contribution Statement

Jessica J. Simon, Douglas M. Fowler, and Dustin J. Maly conceived of the work and wrote the manuscript. Jessica J. Simon developed the LABEL-seq method, performed all profiling experiments and data analysis.

Initial Publication

The material in this work was originally published in: Simon, J.J., Fowler, D.M. & Maly, D.J. Multiplexed profiling of intracellular protein abundance, activity, interactions and druggability with LABEL-seq. *Nat Methods* **21**, 2094–2106 (2024).

References

1. Faure, A.J. et al. Mapping the energetic and allosteric landscapes of protein binding domains. *Nature* **604** , 175-183 (2022).
2. Bandaru, P. et al. Deconstruction of the Ras switching cycle through saturation mutagenesis. *Elife* **6** (2017).

3. Ahler, E. et al. A Combined Approach Reveals a Regulatory Mechanism Coupling Src's Kinase Activity, Localization, and Phosphotransferase-Independent Functions. *Mol Cell* **74** , 393-408 e320 (2019).
4. Chakraborty, S. et al. Profiling of drug resistance in Src kinase at scale uncovers a regulatory network coupling autoinhibition and catalytic domain dynamics. *Cell Chem Biol* **31** , 207-220 e211 (2024).
5. Hayes, T.K. et al. Comprehensive mutational scanning of EGFR reveals TKI sensitivities of extracellular domain mutants. *Nat Commun* **15** , 2742 (2024).
6. Macdonald, C.B. et al. DIMPLE: deep insertion, deletion, and missense mutation libraries for exploring protein variation in evolution, disease, and biology. *Genome Biol* **24** , 36 (2023).
7. Tsuboyama, K. et al. Mega-scale experimental analysis of protein folding stability in biology and design. *Nature* **620** , 434-444 (2023).
8. Kwon, J.J. et al. Structure-function analysis of the SHOC2-MRAS-PP1C holophosphatase complex. *Nature* **609** , 408-415 (2022).
9. Estevam, G.O. et al. Conserved regulatory motifs in the juxtamembrane domain and kinase N-lobe revealed through deep mutational scanning of the MET receptor tyrosine kinase domain. *bioRxiv* (2023).
10. Matreyek, K.A. et al. Multiplex assessment of protein variant abundance by massively parallel sequencing. *Nat Genet* **50** , 874-882 (2018).
11. Laroche, A. et al. Deep mutational engineering of broadly-neutralizing nanobodies accommodating SARS-CoV-1 and 2 antigenic drift. *MAbs* **14** , 2076775 (2022).
12. Esposito, D. et al. MaveDB: an open-source platform to distribute and interpret data from multiplexed assays of variant effect. *Genome Biol* **20** , 223 (2019).
13. Landrum, M.J. et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* **42** , D980-985 (2014).
14. Cancer Genome Atlas Research, N. et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* **45** , 1113-1120 (2013).
15. Tate, J.G. et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res* **47** , D941-D947 (2019).
16. Jeon, H., Tkacik, E. & Eck, M.J. Signaling from RAS to RAF: The Molecules and Their Mechanisms. *Annu Rev Biochem* (2024).
17. Simanshu, D.K. & Morrison, D.K. A Structure is Worth a Thousand Words: New Insights for RAS and RAF Regulation. *Cancer Discov* **12** , 899-912 (2022).
18. Wu, J. et al. Self-Assembly of Intracellular Multivalent RNA Complexes Using Dimeric Corn and Beetroot Aptamers. *J Am Chem Soc* **144** , 5471-5477 (2022).
19. Litke, J.L. & Jaffrey, S.R. Highly efficient expression of circular RNA aptamers in cells using autocatalytic transcripts. *Nat Biotechnol* **37** , 667-675 (2019).
20. Jeck, W.R. et al. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **19** , 141-157 (2013).

21. Peabody, D.S. & Ely, K.R. Control of translational repression by protein-protein interactions. *Nucleic Acids Res* **20** , 1649-1655 (1992).
22. Lim, F. & Peabody, D.S. Mutations that increase the affinity of a translational repressor for RNA. *Nucleic Acids Res* **22** , 3748-3752 (1994).
23. Wu, B., Chao, J.A. & Singer, R.H. Fluorescence fluctuation spectroscopy enables quantitative imaging of single mRNAs in living cells. *Biophys J* **102** , 2936-2944 (2012).
24. Matreyek, K.A., Stephany, J.J., Chiasson, M.A., Hasle, N. & Fowler, D.M. An improved platform for functional assessment of large protein libraries in mammalian cells. *Nucleic Acids Res* **48** , e1 (2020).
25. Lavoie, H. & Therrien, M. Regulation of RAF protein kinases in ERK signalling. *Nat Rev Mol Cell Biol* **16** , 281-298 (2015).
26. Lavoie, H. et al. Inhibitors that stabilize a closed RAF kinase domain conformation induce dimerization. *Nat Chem Biol* **9** , 428-436 (2013).
27. Jin, T. et al. RAF inhibitors promote RAS-RAF interaction by allosterically disrupting RAF autoinhibition. *Nat Commun* **8** , 1211 (2017).
28. Amorosi, C.J. et al. Massively parallel characterization of CYP2C9 variant enzyme activity and abundance. *Am J Hum Genet* **108** , 1735-1751 (2021).
29. Chiasson, M.A. et al. Multiplexed measurement of variant abundance and activity reveals VKOR topology, active site and human variant impact. *Elife* **9** (2020).
30. Li, X. & Lehner, B. Biophysical ambiguities prevent accurate genetic prediction. *Nat Commun* **11** , 4923 (2020).
31. Alabi, S. et al. Mutant-selective degradation by BRAF-targeting PROTACs. *Nat Commun* **12** , 920 (2021).
32. Kondo, Y. et al. Cryo-EM structure of a dimeric B-Raf:14-3-3 complex reveals asymmetry in the active sites of B-Raf kinases. *Science* **366** , 109-115 (2019).
33. Park, E. et al. Architecture of autoinhibited and active BRAF-MEK1-14-3-3 complexes. *Nature* **575** , 545-550 (2019).
34. Yao, Z. et al. Tumours with class 3 BRAF mutants are sensitive to the inhibition of activated RAS. *Nature* **548** , 234-238 (2017).
35. Davies, H. et al. Mutations of the BRAF gene in human cancer. *Nature* **417** , 949-954 (2002).
36. Wan, P.T. et al. Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* **116** , 855-867 (2004).
37. Liau, N.P.D. et al. Negative regulation of RAF kinase activity by ATP is overcome by 14-3-3-induced dimerization. *Nat Struct Mol Biol* **27** , 134-141 (2020).
38. Park, E. et al. Cryo-EM structure of a RAS/RAF recruitment complex. *Nat Commun* **14** , 4580 (2023).
39. Modi, V. & Dunbrack, R.L., Jr. A Structurally-Validated Multiple Sequence Alignment of 497 Human Protein Kinase Domains. *Sci Rep* **9** , 19790 (2019).

40. Spencer-Smith, R. et al. RASopathy mutations provide functional insight into the BRAF cysteine-rich domain and reveal the importance of autoinhibition in BRAF regulation. *Mol Cell* **82** , 4262-4276 e4265 (2022).
41. Roring, M. et al. Distinct requirement for an intact dimer interface in wild-type, V600E and kinase-dead B-Raf signalling. *EMBO J* **31** , 2629-2647 (2012).
42. Yaffe, M.B. et al. The structural basis for 14-3-3:phosphopeptide binding specificity. *Cell* **91** , 961-971 (1997).
43. Hu, J. et al. Kinase regulation by hydrophobic spine assembly in cancer. *Mol Cell Biol* **35** , 264-276 (2015).
44. Koga, T., Suda, K. & Mitsudomi, T. Utility of the Ba/F3 cell system for exploring on-target mechanisms of resistance to targeted therapies for lung cancer. *Cancer Sci* **113** , 815-827 (2022).
45. Bekes, M., Langley, D.R. & Crews, C.M. PROTAC targeted protein degraders: the past is prologue. *Nat Rev Drug Discov* **21** , 181-200 (2022).
46. Roskoski, R., Jr. Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. *Pharmacol Res* **103** , 26-48 (2016).
47. Cantor, A.J., Shah, N.H. & Kuriyan, J. Deep mutational analysis reveals functional trade-offs in the sequences of EGFR autophosphorylation sites. *Proc Natl Acad Sci U S A* **115** , E7303-E7312 (2018).
48. Starita, L.M. et al. Activity-enhancing mutations in an E3 ubiquitin ligase identified by high-throughput mutagenesis. *Proc Natl Acad Sci U S A* **110** , E1263-1272 (2013).
49. Yao, Z. et al. BRAF Mutants Evade ERK-Dependent Feedback by Different Mechanisms that Determine Their Sensitivity to Pharmacologic Inhibition. *Cancer Cell* **28** , 370-383 (2015).
50. Simon, J.J., Fowler, D.M. & Maly, D.J. Multiplexed profiling of intracellular protein abundance, activity, interactions and druggability with LABEL-seq. *Nat Methods* **21**, 2094–2106 (2024).

Methods

A list of oligo sequences is provided in Supplementary Table 5. Protein fusion and MS2-circRNA sequences are provided in Supplementary Table 6. A list of reagents and supplier catalog numbers is provided in Supplementary Table 7.

Cloning

All plasmids, except those encoding libraries, were constructed using gibson assembly and standard protocols then validated through Sanger or whole-plasmid sequencing. Plasmids can be found on addgene (#226162-226170).

Expression of MS2-circRNA architectures

HEK293T cells were transfected in 12-well plates using 0.7 μg of plasmid encoding MS2-circRNA barcode + 0.1 μg pMAX-GFP (to estimate transfection efficiency) + 2.4 μL PEI (1 mg/mL). 48 hr after transfection, media was removed, cells were resuspended directly in TRIzol, and total RNA was extracted according to the manufacturer's protocol. Total RNA concentrations were normalized using a NanoDrop 2000 (Thermo Scientific).

Pulldowns with gel-based visualization of MS2-circRNAs

To prepare dasatinib-conjugated sepharose, 1.25 mL of NHS-activated sepharose slurry was washed twice with 50% DMF/50% EtOH then incubated with 64 μL of dasatinib-amine (50mM) + 240 μL 1M EDC + 1.6 mL 50% DMF/50% EtOH overnight at room temperature rotating end-over-end. After coupling, beads were washed three times with 50% DMF/50% EtOH then capped with ethanolamine by incubating with 544 μL 1M ethanolamine + 240 μL 1M EDC + 1060 μL 50% DMF/50% EtOH overnight at room temperature rotating end-over-end. Capped beads were washed three times with 50% DMF/50% EtOH, two times with 0.5M NaCl, then one time with 20% ethanol in water. Washed beads were resuspended in 20% ethanol to make a 50% bead/50% solvent slurry and stored at 4°C.

For pulldowns, HEK293T cells were transfected in 12-well plates using 0.75 μg tdMCP-protein fusion plasmid + 0.75 μg MS2-circRNA plasmid + 0.15 μg pMAX-GFP (to estimate transfection efficiency) + 5 μL PEI (1 mg/mL). 48 hr after transfection, cells were resuspended in 160 μL lysis buffer (50mM Tris pH=7.6, 150mM NaCl, 10mM NaF, 4mM $MgCl_2$, 5% glycerol, Halt protease inhibitor tablet (1 tablet/50mL), 2mM PMSF, RiboLock RNase Inhibitor (4 μL/mL)) and lysates were cleared at 17,000g for 10 min at 4°C. Prior to enrichment, 40 μL of supernatant was incubated with 15 μM free dasatinib (1% DMSO) or 1% DMSO for 30 min at 4°C. To enrich, 10 μL of dasatinib-conjugated sepharose (pre-washed with lysis buffer) was added to each sample and incubated for 2.5 h at 4°C, rotating end-over-end. Protein-bound beads were washed three times with 100 μL of lysis buffer. For western blot analysis, washed beads were resuspended in 3xSDS. For analysis of RNA enrichment, beads were resuspended in 100 μL TRIzol then extracted and purified using IBI RNA cleanup columns, eluted into 25 μL of water and immediately visualized by gel (see "Gel-based visualization of MS2-circRNA").

Gel-based visualization of MS2-circRNA

To visualize the broccoli aptamer in MS2-circRNA, total RNA containing MS2-circRNA (1 μg for test expressions or 7 μL for pulldown samples) was mixed 1:1 with 2x TBE-urea sample buffer, heated to 75°C for 5 min, then briefly cooled on ice. Samples were loaded onto precast 6% TBE or 10% TBE-Urea gels and run at 270V in TBE buffer to completion. Gels were washed three times for 5 min in milliQ water then stained for 20 min at room temperature in DFHBI-1T buffer (10 μM DFHBI-1T, 40 mM HEPES, pH 7.4, 100 mM KCl, and 1 mM $MgCl_2$), both steps on an orbital shaker. Broccoli was imaged on a Typhoon FLA 9000 using preset settings for Alexa Fluor 488 (495nm excitation; 519nm emission).

Generating the RNA-Protein Landing Pad (RPLP)

To stably generate a high-efficiency Bxb1 landing pad containing Pol II- and Pol III-driven promoters, Flp-In™ T-REx™ 293 cells were plated into a 6-well dish and transfected with 1 μg of landing pad plasmid containing an FRT site, 1 μg Flp recombinase expression plasmid (pOG44), 6 μL PEI (1 mg/mL), and 100 μL serum-free DMEM. 48 hr after transfection, cells were re-seeded to ~20% confluency. After an additional 24 hr, media was supplemented with 200 μg/mL hygromycin to initiate selection. Cells were expanded and maintained in 200 μg/mL hygromycin until >95% of cells were GFP positive indicating near-complete selection, and stocks were cryo-preserved until needed.

Generating stable cell lines

The generation of cell lines stably expressing protein variants and MS2-circRNA is described in full in Supplementary Note 1. Briefly, stable cell lines expressing both protein and MS2-circRNA were generated in the HEK293 RPLP. Stable lines not expressing MS2-circRNA were generated in the HEK293T lenti landing pad (LLP-iCasp9-Blast[24]). The following protocol applies to both cell lines.

Both cell lines were maintained in 2 μg/mL doxycycline (DOX) at all times unless otherwise noted. From transfection up until 24-48 h prior to cell lysis, cultures were also maintained in 100 nM SCH772984. To introduce variants into the empty landing pad, cells were co-transfected with Bxb1 attB-containing plasmids and a plasmid encoding the Bxb1 recombinase. Selection for successful

recombinants was initiated >3 days after transfection by treating cells with AP1903 and puromycin.

Barcode sequence missassociation assay (Fig. 2)

Two barcode libraries were cloned as described in the 'Generating single protein variant barcode libraries' section of Supplementary Note 1, each containing 20,000-30,000 unique barcode sequences and either 3xFlag- or 2xMyc-tdMCP-BRaf wild type. Plasmid libraries were individually recombined into the RPLP as described in the 'Generating stable cell lines' section.

The day before enrichment, equal numbers of cells from both libraries were combined and plated into 15 cm dishes in +DOX media. On the day of enrichment, cells were trypsinized and all plates pooled. The pool aliquots were then divided into 5 enrichment samples: 2 no-dummy samples ($30x10^6$ cells each), 2 +dummy samples ($10x10^6$ cells each), and 1 input sample ($10x10^6$ cells each). Each aliquot was washed once with DPBS, then fully resuspended in 1,650 µL (no dummy samples) or 550 µL (+dummy and input) modRIPA without detergent (resuspension buffer: 50mM Tris pH=7.6, 150mM NaCl, 10mM NaF, 4mM $MgCl_2$, 5% glycerol, Halt protease inhibitor tablet (1 tablet/50mL), 2mM PMSF, RiboLock RNase Inhibitor (4 µL/mL)). For +dummy samples, the resuspension buffer also contained 400ng/µL total RNA previously extracted from HEK293T cells that were transiently transfected with dummy MS2-circRNA. Immediately after resuspending cells in the resuspension buffer, cells were lysed by adding an equal volume of modRIPA containing 0.1% NP40. Addition of buffer with detergent marked the beginning of the incubation time. Lysates were cleared by centrifugation at 17,000g for 10 min at 4°C. Each enrichment sample was split into two aliquots (1.5 mL ea. for no dummy, 0.5 mL ea. for +dummy) and transferred to low-

bind eppendorf tubes. To enrich tdMCP-protein fusions, 150 µL (no dummy) or 50 µL (+dummy) of anti-Flag or anti-Myc magnetic beads (volumes correspond to bead slurry volumes) was added to the lysates and rotated end-over-end at 4°C for the indicated time. 500 µL samples of bead suspension were collected at each timepoint and washed 3 times with 500 µL modRIPA (including 0.1% NP40). After washing, all buffer was removed before resuspending beads in 75 µL nuclease-free water and 300 µL TRIzol, then stored at -20. The input sample was resuspended in 1.5 mL TRIzol and stored at -20. RNA was extracted using IBI scientific TRI-Isolate RNA Pure Kit, concentrated using Zymo RNA Clean & Concentrator, eluted into 10 µL nuclease-free water, and stored at -80°C.

Variant abundance assay, proof-of-concept experiment (Fig. 3d)

HEK293T LLP cells stably expressing 3xFlag-tdMCP-BRaf wild type (WT) or G466V were transfected in 24-well plates, each well receiving 300 ng pcDNA5 2xMyc-tdMCP-SNAP + 200 ng of an equal mix of 4 plasmids that each contain a unique circRNA barcode + 50 ng pMAX GFP + 50 µL DMEM + 1.65 µL PEI (1 mg/mL). 6 transfections per variant encompassing a total of 48 unique barcodes were performed. 24 hr after transfection cells were treated with 250 nM SJF0628 or DMSO for an additional 24 hr. Cells were next lysed in 120 µL modRIPA (0.1% NP40, 50mM Tris pH=7.6, 150mM NaCl, 10mM NaF, 4mM $MgCl_2$, 5% glycerol, Halt protease inhibitor tablet (1 tablet/50mL), 2mM PMSF, RiboLock RNase Inhibitor (4uL/mL)) and lysate cleared by centrifugation at 17,000g for 10 min at 4°C. Samples were taken for western blot analysis before pooling 40 µL lysate from all 12 wells. Pooled lysate was subsequently split into 2x200 µL aliquots and protein was enriched by adding 20 µL (slurry volume) separately to anti-Flag or anti-Myc magnetic beads and rotating end-over-end at 4°C for 1 hr. Beads were washed 3 times with

modRIPA, resuspended in 50 µL nuclease-free water and 200 µL TRIzol, then stored at -20°C. Enriched RNA was extracted and sequenced as described in the 'Preparation of RNA for Genewiz Amplicon-EZ sequencing' section.

Sequencing ratios were calculated by dividing the sequencing counts in the Flag enrichment by counts in the Myc enrichment then normalizing to the mean Flag/Myc ratio for WT cells treated with DMSO. The mean ratio of all barcodes in a well was reported. Similarly, for western blot ratios the intensity of the Flag signal was divided by the intensity of the Myc signal. Ratios were normalized to the average ratio of WT cells treated with DMSO. The three data points correspond to the three transfections.

Variant activity assay, proof-of-concept experiment (Fig. 4c)

HEK293T LLP cells stably expressing untagged BRaf WT or V600E were transfected in 12-well plates with 0.5 µg 3xFlag-tdMCP-Erk2 + 0.4 µg of an equal mix of 4 plasmids that each contain a unique MS2-circRNA barcode sequence + 0.1 µg pMAX-GFP + 50 µL DMEM + 3 µL PEI (1 mg/mL). Transfections of each variant were performed in triplicate, encompassing a total of 24 unique barcodes. 48 hr after transfection, cells were lysed in 150 µL modRIPA (0.1% NP40, 50mM Tris pH=7.6, 150mM NaCl, 10mM NaF, 4mM $MgCl_2$, 5% glycerol, Halt protease inhibitor tablet (1 tablet/50mL), 2mM PMSF, RiboLock RNase Inhibitor (4 µL/mL), 100 µL each phosphatase inhibitor cocktails 2 and 3) and lysate cleared by centrifugation at 17,000g for 10 min at 4°C. 30 min after lysis all samples were pooled. The pooled lysate was divided into two 350 µL aliquots and protein was enriched by adding 35 µL (slurry volume) anti-Flag or anti-pErk magnetic beads (see 'Activity Assay' for instructions on generating pErk beads), rotating end-over-end at 4°C for

2 hr. Beads were washed 3 times with modRIPA, resuspended in 50 µL nuclease-free water and 200 µL TRIzol and then stored at -20°C. Enriched RNA was extracted and sequenced as described in the 'Preparation of RNA for Genewiz Amplicon-EZ sequencing' section.

pErk/Flag ratios were calculated by dividing the sequencing counts in the Flag enrichment by counts in the pErk enrichment then normalizing to the mean Flag/Myc ratio for WT cells. The mean ratio of all 24 barcodes for each variant was reported.

Proximity labeling with western blot readout (Fig. 5b,d and ED Fig. 6a-b)

For TurboID- or miniTurbo-MEK1 experiments, HEK293 RPLP cells expressing MS2-circRNA 2 and Flag-tdMCP-BRaf WT or I666R were transfected in 10 cm plates using 6.75 µg pCDNA5 TurboID/miniTurbo plasmid + 0.75 µg pMAX GFP + 22.5 µL PEI + 375 µL DMEM. 24 h after transfection, cells were re-seeded into 12-well plates. 48 h after transfection cells were treated with DMSO or 10 µM cobimetinib for 4 h followed by a co-treatment with drug and biotin (DMSO, 200 µM, or 500 µM) for 90 or 180 min. After biotin labeling, cells were washed 1x with DPBS and lysed in modRIPA (50 mM Tris pH=7.6, 150 mM NaCl, 10 mM NaF, 4 mM $MgCl_2$, 5% glycerol, Halt protease inhibitor tablet (1 tablet/50 mL), 2 mM PMSF, RiboLock RNase Inhibitor (4 µL/mL), 0.1% NP40). Lysates were cleared by centrifugation at 17,000 g (4°C), 10 min. To capture tdMCP-fusions, 7 µL (slurry volume) anti-Flag beads were incubated with 70 µL lystate at 4°C, rotating for 1 h. Beads were washed 3x with 100 µL modRIPA, resuspended in 30 µL 3xSDS and boiled for 5 min.

For TurboID- or miniTurbo-CRaf experiments, HEK293 RPLP cells expressing MS2-circRNA 2 and Flag-tdMCP-BRaf WT or R509H were transfected and re-seeded as above. 48 h after transfection, cells were treated with DMSO or 10 μM LXH254 for 2 h followed by a co-treatment with drug and biotin (DMSO, 200 μM, or 500 μM) for 90 or 180 min. tdMCP-fusions were processed as described above.

Proximity labeling time courses with sequencing readout (Fig. 5c,e,f and ED Fig. 6c-e)

Prior to the experiment, 48 unique barcode sequences of MS2-circRNA 1b were expressed individually in HEK293T's. 48 hr after transfection, cells were collected and total RNA was extracted using TRIzol and isopropanol precipitation following the manufacturer's protocol. The extracted total RNA was diluted to ~1μg/μL in nuclease-free water, aliquoted, and stored at -80.

HEK293T LLP cells stably expressing tdMCP fusions were seeded into plates using DMEM supplemented with regular or dialyzed FBS. Cells were carried through the entire experiment in the same type of FBS (regular or dialyzed). After seeding, cells were transfected with TurboID- or miniTurbo- fusions. Transfection mixes per $1x10^6$ cells included 1.8 μg pCDNA5 TurboID/miniTurbo plasmid + 0.2 μg pMAX GFP + 6 μL Turbofectin + 100 μL DMEM. 24 hr after transfection, cells were re-seeded to ~80% confluency into 96-well poly-D-lysine coated plates. ~24 hr after moving to 96-well plates, cells were treated with DMSO or 10 μM GDC-0879 (TurboID-/miniTurbo- CRaf or BRaf experiments only). After 2 h pre-treatment with drug, biotin treatment was initiated by supplementing media with biotin to a final concentration of 150 μM biotin and 10 μM GDC-0879 (TurboID-/miniTurbo- CRaf or BRaf experiments only). Subsequent biotin labelings were initiated to achieve 5, 15, or 45 min labeling. Labeling was quenched for all

samples simultaneously by removing media, washing 3x with warm DPBS, and immediately adding 40 µL cold modRIPA (50 mM Tris pH=7.6, 150 mM NaCl, 10 mM NaF, 4 mM $MgCl_2$, 5% glycerol, Halt protease inhibitor tablet (1 tablet/50 mL), 2 mM PMSF, RiboLock RNase Inhibitor (4 µL/mL), 0.1% NP40) to each well. Cells were incubated at 4°C for 20 min on an orbital shaker to ensure cells were fully lysed. During the 4°C incubation, total RNA stocks containing MS2-circRNAs were thawed and diluted 1/75 into cold modRIPA. 10 µL of a different diluted MS2-circRNA sequence was added to each lysate-containing well. Lysates were then incubated with RNA at 4°C for 2 h on an orbital shaker. After incubation, 50 µL cold modRIPA supplemented with 1mg/mL BSA was spiked into each well and 80 µL from each well were pooled. Lysates were cleared by centrifugation at 17,000g for 10 min at 4°C. 1.9 mL of supernatant was added separately to 95 µL (slurry volume) anti-Flag bead or streptavidin-conjugated agarose and rotated end-over-end at 4°C for 1hr. After incubation, the beads were washed 3x1 mL cold modRIPA then resuspended in 75 µL water and 300 µL TRIzol and stored at -20°C. MS2-circRNA was extracted and sequenced as described in 'Preparation of RNA for Genewiz Amplicon-EZ sequencing'.

Streptavidin/Flag ratios were calculated by dividing the sequencing counts in the Streptavidin enrichment by counts in the Flag enrichment. Each barcode ratio is reported.

Preparation of RNA for Genewiz Amplicon-EZ sequencing

A detailed description is provided in Supplementary Note 1. Briefly, enriched RNA was extracted from TRIzol and reverse transcribed using SuperScript™ IV Reverse Transcriptase and a gene-specific custom primer (JJS1). Two subsequent rounds of PCR were performed to amplify only

circular RNA and add Genewiz custom adapters. Reactions were cleaned, pooled, and sent to Genewiz for their Amplicon-EZ sequencing service.

Western blotting of individual variants (Fig. 3g and 4g)

HEK293T LLP cells stably expressing Flag-tagged BRaf variants were lysed by resuspending in cold modRIPA (50 mM Tris pH=7.6, 150 mM NaCl, 10 mM NaF, 4 mM $MgCl_2$, 5% glycerol, Halt protease inhibitor tablet (1 tablet/50 mL), 2 mM PMSF, 0.1% NP40, 1% each phosphatase inhibitor cocktails 2 and 3). Lysates were cleared at 17,000g (4°C) for 10 min and total protein concentrations were measured then normalized using Pierce™ 660nm Protein Assay Reagent. Samples were mixed 2:1 with 3xSDS loading buffer (188 mM Tris-Cl (pH 6.8), 3% SDS, 30% glycerol, 0.01% bromophenol blue, 15% β-mercaptoethanol) and heated to 95°C for 5 min. 3-10 µg total protein was loaded and run on AnykD™ Criterion™ TGX™ Precast acrylamide protein gels then transferred to nitrocellulose. Nitrocellulose was incubated at room temp for one hour in 5% dry milk in TBST then treated overnight with primary antibody diluted in 5% dry milk in TBST at 4°C. The next day, nitrocellulose was washed 3 x 5 min in TBST then incubated for >1 h with secondary antibody diluted 1:10,000 in 5% dry milk in TBST. Westerns were scanned on a LI-COR Odyssey® and quantified using ImageStudio Lite or ImageJ software.

pErk was probed using CST #4370L, total Erk CST #9107S, BRaf CST #14814S, Flag Sigma # F1804, GAPDH Santa Cruz Biotechnology #365062, streptavidin fluorophore LI-COR #926-32230, pMek CST#9154S, HA CST #3724, BRaf pS729 Abcam #ab124794, Vinculin CST #13901.

Mek1 co-immunoprecipitations (Extended Data Fig. 7k)

HEK293T LLP cells stably expressing Flag-tagged BRaf variants (WT, P367R, K601E, L485S, P318*) were transfected in 6-well plates in triplicate with 1.8 µg pcDNA5 HA-Mek1 + 0.2 µg pMAX GFP + 6 µL PEI (1 mg/mL) + 100 µL DMEM. 48 hr after transfection, cells were lysed using 200 µL high-glycerol modRIPA (0.1% NP40, 50mM Tris pH=7.6, 150mM NaCl, 10mM NaF, 4mM $MgCl_2$, 10% glycerol, Halt protease inhibitor tablet (1 tablet/10mL), phosphatase inhibitor cocktails 2 and 3 (100uL/10mL)) and lysates were cleared by centrifugation at 17,000g for 10 min at 4°C. To enrich Mek-Raf complexes, 150 µL of supernatant was combined with 12 µL (slurry volume) anti-Flag magnetic beads and rotated end-over-end at 4°C for 1 h then quickly washed three times with 100 µL modRIPA. Enriched complexes were eluted from the bead by resuspending in 30 µL 3xSDS and heating to 95°C for 5 min.

Cloning tdMCP-BRaf variant library

A detailed description of library cloning can be found in Supplementary Note 1. Briefly, A site saturation mutagenesis library of full length BRaf consisting of variants at 80 chosen positions was purchased from Twist Bioscience. The double stranded DNA pool contained the entire 3xFlag-tdMCP-BRaf open reading frame with flanking 5' and 3' sequences complementary to the vector. This gene fragment library was cloned into a linearized attB vector using Gibson assembly, transformed into NEB® 5-alpha Competent *E. coli*, and grown in liquid LB overnight. DNA was harvested from the saturated overnight culture. The resulting BRaf variant library was barcoded and bottlenecked to the desired number of transformants as described in the 'Generating single protein variant barcode libraries' section of Supplementary Note 1. All oligonucleotides used to generate libraries and amplicons are reported in Supplementary Table 5.

Cloning untagged BRaf variant library

A detailed description of library cloning can be found in Supplementary Note 1. Briefly, to clone a BRaf variant library without the 3xFlag-tdMCP fusion, the attB 3xFlag-tdMCP-BRaf variant library (prior to barcoding) was digested to remove 3xFlag-tMCP. A DNA fragment encoding an HA-tag was subsequently inserted at the N-terminus of BRaf using gibson assembly, transformed into *E. coli*, grown overnight in liquid LB, and DNA harvested from the saturated culture. This variant library was barcoded and bottlenecked to the desired number of transformants as described in the 'Generating single protein variant barcode libraries' section of Supplementary Note 1. All oligonucleotides used to generate libraries and amplicons are reported in Supplementary Table 5.

Generating barcode sequence:protein variant map

Barcoded variant libraries were digested overnight with SpeI-HF and NheI-HF then submitted to University of Washington PacBio Sequencing Services for further processing. Briefly, PacBio libraries were generated with the SMRTbell Library Prep kit (Pacific Biosciences) and sequenced on a Sequel II SMRT Cell 8M (Pacific Biosciences). Reads were processed using a custom analysis pipeline available at https://github.com/shendurelab/AssemblyByPacBio to identify and link genes with associated barcodes (Extended Data Fig. 3d). Enrich2 was used to convert DNA variants into protein coding mutations.

Abundance profiling of tdMCP-tagged BRaf variants (Figure 3e,f and Extended Data 3f,g)

The day before transfection, tdMCP-tagged BRaf variant library cells were plated into 4 x 15 cm dishes at ~45% confluency (+2 μg/mL DOX, +100 nM SCH772984). The next day, each plate of

cells was transfected with 17 μg (pCDNA5) Myc-tdMCP-SNAP + 1.7 μg (pMAX) GFP + 70 μL Fugene HD + 935 μL serum-free DMEM. 48 h after transfection, cells were trypsinized, pooled, and divided into two aliquots. Each aliquot of cells was fully resuspend in 2 mL cold resuspension buffer, which is modRIPA (without detergent) supplemented with dummy barcode to minimize barcode sequence missassociation (50mM Tris pH=7.6, 150mM NaCl, 10mM NaF, 4mM $MgCl_2$, 5% glycerol, Halt protease inhibitor tablet (1 tablet/50 mL), 2 mM PMSF, RiboLock RNase Inhibitor (4 uL/mL), 100 ng/μL total RNA containing dummy barcode). To lyse the resuspended cells, an equal volume of lysis buffer (0.2% NP40, 50 mM Tris pH=7.6, 150 mM NaCl, 10 mM NaF, 4 mM $MgCl_2$, 5% glycerol, Halt protease inhibitor tablet (1 tablet/50mL), 2 mM PMSF, RiboLock RNase Inhibitor (4uL/mL)) was added. Lysates were cleared by centrifugation at 17,000g, 4°C, 10 min then each sample was split into 1.2 mL and 2.8 mL aliquots and added separately to 120 μL (slurry volume) anti-Flag beads or 140 μL (slurry volume) anti-Myc beads, respectively. Lysates were rotated end-over-end at 4°C for 3 h to capture protein. The beads were subsequently washed 3 times with cold wash buffer (50 mM Tris pH=7.6, 150 mM NaCl, 10 mM NaF, 4 mM $MgCl_2$, 5% glycerol, 0.1% NP40) then resuspended in 75 μL nuclease-free water and 300 μL TRIzol and stored at -20.

Amplicons were prepared and sequenced described in “Preparation of RNA for Illumina sequencing”.

Activity profiling of untagged BRaf variants (Fig. 4e,f, Extended Data Fig. 4b-f and Extended Data Fig. 5a-d)

Anti-pErk beads were generated by combining the following components the day before enrichment and rotating end-over-end at 4°C until ready to use: 350 µL protein A beads + 350 µL Tris buffered saline + 70 µL anti-pErk antibody.

To prepare libraries, untagged (no MCP) BRaf variant library cells were plated into two 15-cm dishes (+2 µg/mL DOX, +100 nM SCH772984) the day before transfection. On the day of transfection, the media was replaced with fresh media without SCH772984. When activity was profiled with overexpressed CRaf (Fig. 4e,f and Extended Data Fig. 4b-f), each 15 cm dish was transfected with 17 µg pcDNA5 CRaf-IRES-pErk reporter (3xFlag-tdMCP-Erk2) + 1.7 µg pMAX GFP + 70 µL Fugene HD + 950 µL serum-free DMEM. When activity was profiled without overexpressed CRaf (Extended Data Fig. 5a-d), cells were transfected with 2 µg pcDNA5 pErk reporter + 14 µg empty plasmid + 1.6 µg pMAX GFP + 66 µL FugeneHD + 880 µL serum free DMEM. 48 h after transfection, each plate of cells was trypsinized, washed once with DPBS, then fully resuspended in 1 mL cold resuspension buffer supplemented with 1% each phosphatase inhibitor cocktails 2 and 3. To lyse the resuspended cells, an equal volume of lysis buffer was added. The lysates were subsequently cleared by centrifugation at 17,000g for 10 min at 4°C, then each sample was split into 1.75 mL and 250 µL aliquots. 150 µL (slurry volume) of anti-pErk beads or 25 µL (slurry volume) anti-Flag beads were added to the 1.75 mL and 250 µL aliquots, respectively. The lysates were incubated with beads for 4 hr at 4°C, rotating end-over-end. After enrichment, the beads were washed three times with 1 mL cold wash buffer, resuspended in 75 µL nuclease-free water and 300 µL TRIzol, then stored at -20.

Amplicons were prepared and sequenced described in "Preparation of RNA for Illumina sequencing".

CRaf interaction profiling (Fig. 5h, Extended Data Fig 7a-e and Extended Data 10)

TurboID-CRaf was introduced to the tdMCP-BRaf variant library using lentivirus. To generate virus, one 10-cm dish of HEK293T cells were transfected with 6 μg transfer plasmid pLJM1 Myc-TurboID-CRaf, EGFP-IRES-BlastR + 3 μg pMDLg/pRRE-Gag & Pol + 1.5 μg pRSV-Rev + 1.5 μg VSV-G envelope + 42 μL Fugene 6 + 600 μL serum-free DMEM. 48 h after transfection the media was collected and filtered through a 0.4 μm frit. Half of the filtered media was applied to a 15 cm dish of the tdMCP-BRaf variant library (+2 μg/mL DOX, +100 nM SCH772984). Prior to transduction, the variant library was supplemented with RPLP cells stably expressing barcoded Flag-tdMCP-BTK at a frequency of 0.1% of the cell population. The transduced cell population was maintained for 3 days in complete media (+2 μg/mL DOX, +100 nM SCH772984) supplemented with regular FBS. On day 3, the cells were transferred to media containing dialyzed FBS (+DOX, +SCH772984). On day 4 the cells were plated into 2 15-cm and 2 10-cm dishes, each 50% confluent in media without SCH772984. Drug treatment, biotin labeling and enrichments were performed on day 5.

On day 5, cells were treated with DMSO (0.1% final concentration) (Fig. 5h and Extended Data Fig 7a-e), 2 μM LXH254 (Extended Data Fig. 10b), or 5 μM Vemurafenib (Extended Data Fig. 10a). After 1 h incubation, in-cell proximity labeling was performed by adding biotin to a final concentration of 200 μM. Cells in the 15-cm dishes were returned to the incubator for 5 min at which point the labeling was quenched by trypsinizing the cells, washing once with DPBS, then

proceeding immediately to cell lysis. The cells in the 10 cm dishes were not treated with biotin and were reserved for Flag enrichment.

To lyse the cells, each cell pellet was first fully resuspended in 1 mL (15-cm dish) or 500 µL (10-cm dish) cold resuspension buffer then treated with an equal volume of lysis buffer. The resulting lysate was cleared by centrifugation at 17,000 g for 10 min at 4°C. To enrich biotin-labeled variants, 2 mL of cleared lysate was combined with 100 µL (slurry volume) streptavidin-conjugated agarose. 1 mL lysate from unlabeled plates was combined with 100 µL (slurry volume) anti-Flag beads. Samples were rotated end-over-end at 4°C for 1 h then the beads were washed 3 times with 1 mL cold wash buffer. After removing the buffer, the beads were resuspended in 75 µL water and 300 µL TRIzol then stored at -20.

Amplicons were prepared and sequenced described in "Preparation of RNA for Illumina sequencing".

Mek1 interaction profiling (Fig. 5j and Extended Data Fig 7g,h)

The day prior to transfection, MCP-tagged BRaf variant library cells were plated into two 15-cm dishes (+DOX, +100 nM SCH772984, 10% regular FBS). On the day of transfection, the media was exchanged for DMEM supplemented with 10% dialyzed FBS (+DOX, +100nM SCH772984). The culture was also supplemented with RPLP cells stably expressing barcoded Flag-tdMCP-BTK at a frequency of 0.1%. The resulting mixed-population was transfected with 16 µg pcDNA5 TurboID-Mek1 + 1.6 µg pMAX GFP + 66 µL Fugene HD + 880 µL DMEM. 24 hr after transfection, the culture was expanded to two 15-cm dishes and two 10-cm dishes in DMEM

(+dialyzed FBS, +DOX, +100 nM SCH772984). 48 hr after transfection, cells in the 15-cm dishes were incubated at 37°C in media supplemented with 250 μM biotin for 5 min. To halt labeling, cells were trypsinized, washed once with DPBS and immediately lysed. The cells in the 10-cm dishes were not treated with biotin and were reserved for Flag enrichment

To lyse, the collected cells were first fully resuspended in 1 mL (15-cm dishes) or 440 μL (10-cm dishes) resuspension buffer then treated with an equal volume of cold lysis buffer. The resulting lysates were cleared by centrifugation at 17,000g for 10 min at 4°C. 100 μL streptavidin-conjugated agarose (50% slurry, washed) was added to the 2 mL of supernatant corresponding to each biotin-labeled replicate. 50 μL (slurry volume) anti-Flag beads were added to 800 μL of supernatant corresponding to the non-biotin-labeled samples. Lysates were rotated end-over-end at 4°C for 45 min then beads were washed 3x1 mL using cold wash buffer. After washing, beads were resuspended in 75 μL nuclease-free water and 300 μL TRIzol then stored at -20.

Amplicons were prepared and sequenced described in "Preparation of RNA for Illumina sequencing".

SJF0628 degradation assay (Fig. 6f and Extended Data Fig. 9d,e))

The tdMCP standard, 2xMyc-tdMCP-SNAPtag, was introduced to the tdMCP-BRaf variant library using lentivirus. To generate virus, five 10-cm dishes of HEK293T cells were each transfected with 6 μg transfer plasmid pLJM1 2xMyc-tdMCP-SNAP, EGFP-IRES-BlastR + 3 μg pMDLg/pRRE-Gag & Pol + 1.5 μg pRSV-Rev + 1.5 μg VSV-G envelope + 42 μL Fugene 6 + 600 μL serum-free DMEM. 48 h after transfection the media from each plate was collected, pooled,

and filtered through a 0.4 μm frit. 6% of the filtered media was applied to each of 9 15-cm dishes of the tdMCP-BRaf variant library (+2 μg/mL DOX, 100 nM SCH772984). 24 h after transduction, media was replaced with fresh media without SCH772984 (+DOX, no SCH772984). 48 h after transduction, cells were expanded to 12 x 15 cm dishes and treated with SJF0628 without SCH772984 (+DOX). Final concentrations of SJF0628 were 2.5 μM, 500 nM, 100 nM, 20 nM, 4 nM, DMSO only and two replicates were performed per SJF0628 concentration.

20 h after SJF0628 treatment, each 15-cm dish was lysed by trypsinizing and collecting cells, resuspending in 1 mL cold resuspension buffer followed by treatment with 1 mL cold lysis buffer. Lysates were cleared by centrifugation at 17,000g for 10 min at 4°C, then each sample was split into 540 μL and 1260 μL aliquots and added separately to 40 μL (slurry volume) anti-Flag beads or 120 μL (slurry volume) anti-Myc beads, respectively. Lysates were rotated end-over-end at 4°C for 3 h to capture protein. The beads were subsequently washed 3 times with 1 mL cold wash buffer then resuspended in 75 μL nuclease-free water and 300 μL TRIzol and stored at -20.

Amplicons were prepared and sequenced described in "Preparation of RNA for Illumina sequencing".

Preparation of RNA for Illumina sequencing

A complete description of amplicon preparation is described in Supplementary Note 1. To prepare barcodes for Illumina sequencing, RNA was extracted from the enrichment beads using TRIzol, cleaned, then reverse transcribed using SuperScript™ IV Reverse Transcriptase and a gene-specific custom primer (JJS1). Two subsequent rounds of PCR were performed using primers JJS8-

24, amplifying only circular RNA and adding Illumina adapter sequences. Subsequent reactions were pooled, cleaned, and sequenced as described in "Illumina Sequencing".

Illumina Sequencing

Amplicons were sequenced on a NextSeq 500 using a NextSeq 500/550 High Output v2.5 75 cycle kit or on a NextSeq 2000 using a NextSeq 1000/2000 P2 Reagents 100 cycle kit. 16 nt barcodes were read twice by the custom paired-end sequencing primers JJS25 and JJS26. 8 nt dual indices were sequenced using JJS27 and JJS28. Sequencing reads were converted from BCL to FASTQ format and de-multiplexed using bcl2fastq. Paired sequencing reads were joined by PANDAseq (v2.11) using BFSRKM assembly.

Calculating variant scores and classifications

Scripts used to filter barcodes and calculate scores are available on the GitHub repository (see Code Availability). Frequency values for each barcode in an experiment (F_b) were generated by calculating the sum of counts recorded for that barcode in all enrichments and dividing by the sum of counts recorded across all barcodes and enrichments.

$$F_b = \frac{C_{enrichment\ 1} + C_{enrichment\ 2} + C_{enrichment\ 3} + C_{enrichment\ 4}}{\Sigma C_{enrichment\ 1} + \Sigma C_{enrichment\ 2} + \Sigma C_{enrichment\ 3} + \Sigma C_{enrichment\ 4}}$$

For example, in pErk experiments with two replicates there were four enrichments including two selective anti-pErk immunoprecipitations and two anti-Flag immunoprecipitations. This frequency calculation was performed for each experiment.

Frequency values were used to remove low-frequency barcodes which we reasoned would contribute undue levels of counting noise. To determine a frequency cutoff for each experiment, we started with the assumption that accurate synonymous wild type variant scores should form a symmetrical distribution centered at the wild type score. Note, in all experiments wild type was overrepresented with approximately 80 barcodes and therefore the wild score was assumed to be highly accurate. In each experiment we evaluated the impact of a range of minimum F_b thresholds on the width and central tendency of the distribution of synonymous wild-type scores by excluding barcodes falling below the specified minimum. The selected threshold F_b values, which were unique to each experiment, minimized the skew of the synonymous distribution while retaining most variants.

For pErk enrichments, an additional filter was applied that removed all barcodes containing the motif 'ATAAA' or 'ATTAA' from our analyses. When compared to other barcodes mapped to the same variant, sequences containing these motifs were disproportionately enriched suggesting they were being directly enriched by the anti-pErk antibody.

To generate variant scores, we first calculated a ratio for each barcode (r_b) that passed the filters above. For activity scans, the ratio ($r_{b,act}$) was the barcode count observed in the selective pErk enrichment (C_{pErk}) normalized to the barcode count observed by enriching all pErk reporter through Flag (C_{Flag}).

$$r_{b,act} = \frac{C_{pErk}}{C_{Flag}}$$

The interaction scan barcode ratio ($r_{b,int}$) was the barcode count observed in the selective streptavidin enrichment (C_{Strept}) normalized to the barcode count observed after enriching all variants through Flag (C_{Flag}).

$$r_{b,int} = \frac{C_{Strept}}{C_{Flag}}$$

The abundance barcode ratio ($r_{b,ab}$) was the barcode count from the variant enrichment (C_{Flag}) divided by the barcode count from the tdMCP standard enrichment (C_{Myc}).

$$r_{b,ab} = \frac{C_{Flag}}{C_{Myc}}$$

In activity, interaction, and abundance assays the variant score (S_{Var}) is the mean r_b across all barcodes mapped to the variant ($\overline{r_{b,Var}}$) normalized to the mean wild type barcode ratio ($\overline{r_{WT}}$).

$$S_{Var} = \frac{\overline{r_{b,Var}}}{\overline{r_{WT}}}$$

Using these definitions, the wild type score is always set at 1. Scores greater than 1 indicate variants with increased activity, interactions, or abundance. Conversely, scores less than 1 indicate variants with decreased activity, interactions, or abundance.

For each scan the distribution of synonymous wild type variant scores was used to classify each variant. First, we calculated a cutoff value defined as two standard deviations from the mean of

the synonymous distribution ($2SD_{synon}$). Next, using the wild type score (1) we defined upper and lower threshold values as

$$upper\ threshold\ =\ 1\ +\ 2SD_{synon}$$

$$lower\ threshold\ =\ 1\ -\ 2SD_{synon}$$

Variants with scores greater than the upper threshold value were classified as 'increased abundance', 'increased activity', or 'increased interaction'. Variants with scores lower than the threshold value were classified as 'decreased abundance', 'decreased activity', or 'decreased interaction'. Variants with scores that fell between the threshold values were classified as 'Wild Type-like'.

Regression analysis

Linear least-squares regression plots of replicate LABEL-seq assays were generated using the Seaborn lmplot function. Corresponding R values were calculated using the SciPy lineregress function.

Hypothesis testing

Hypothesis testing, calculation of P values, and plotting the associated data was performed using GraphPad Prism using the analyses described in the figure legends.

Hierarchical clustering

To cluster residues by their biochemical properties, we first transformed our CRaf interaction scores, Mek interaction scores, activity scores measured with CRaf, CRaf-dependent activity scores, and abundance scores into z-scores. The euclidean distances between positions (e.g. comparing the activity score for substitution A at residue 1 to the activity score for the same substitution A at residue 2, repeated across all substitutions and scores) were calculated using sklearn.metrics.pairwise.nan_euclidean_distances which ignores coordinates with missing values and scales up the weight of the remaining coordinates accordingly. Residues were then clustered using an average distance calculation (scipy.cluster.hierarchy.linkage).

Logistic modeling

The growth factor-independent proliferation predictive model was built using the scikit-learn LogisticRegression package and liblinear solver. Model features were z-score-transformed LABEL-seq activity scores, CRaf interaction scores, Mek1 interaction scores, abundance scores, and CRaf-promoted activity.

The model built to predict variant sensitivity to SJF028 degradation was created using the scikit-learn LogisticRegression package with the liblinear solver. While training the model, 20% of the SJF0628 degradation data was reserved for model evaluation. Model features were z-score-transformed LABEL-seq activity scores, CRaf-promoted activity, CRaf interaction scores, Mek1 interaction scores, abundance scores, Vemurafenib-promoted heterodimerization scores, and LXH254-promoted heterodimerization scores. Predicted values were LABEL-seq-derived SJF0628 degradation classifications.

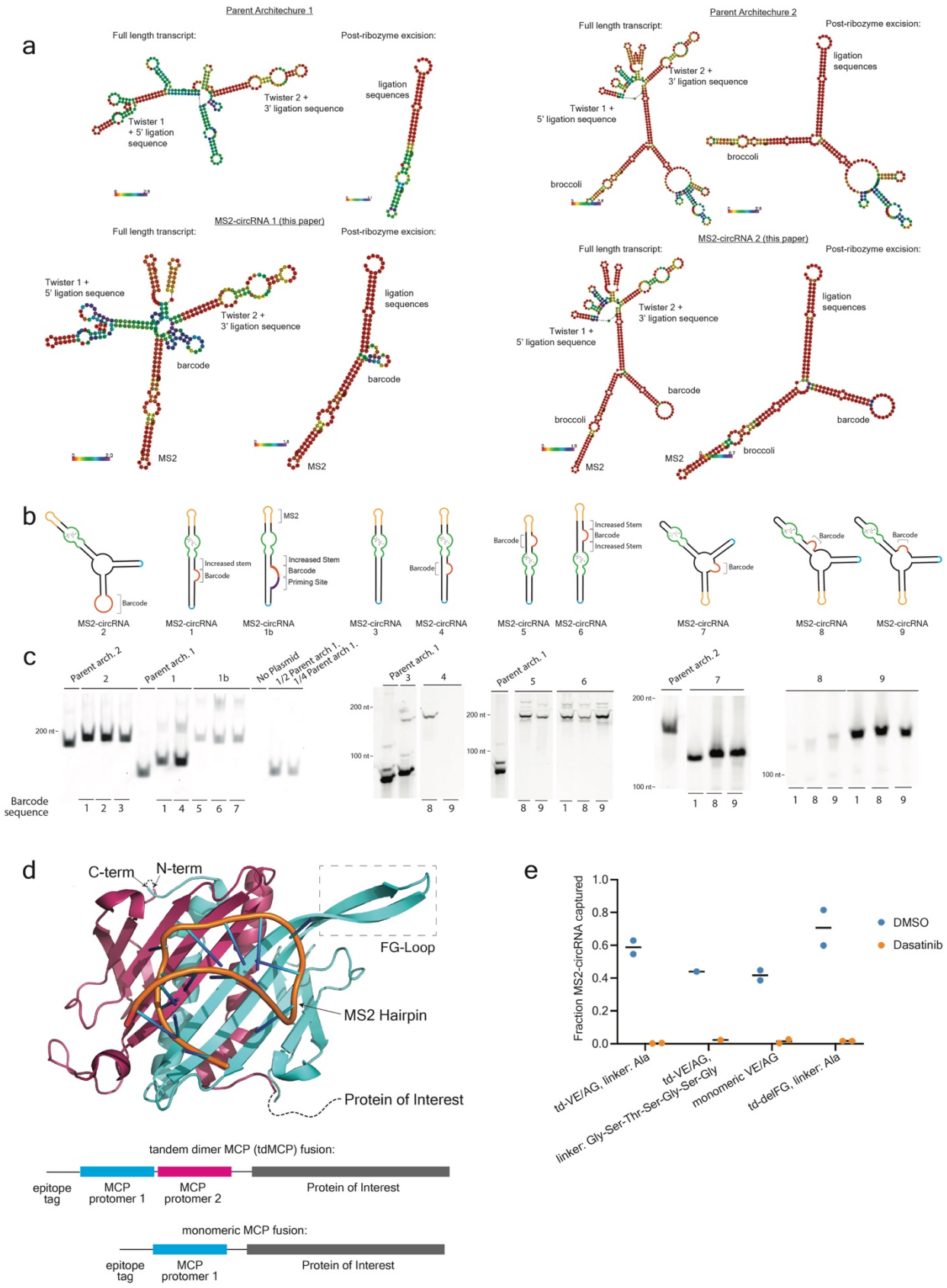
Data availability

Sequencing data is available in the Sequence Read Archive (SRA) under accession number PRJNA1110207. Variant data from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) and Catalogue of Somatic Mutations in Cancer (COSMIC, https://cancer.sanger.ac.uk/cosmic) was used to identify sites of recourring BRaf mutations. Data from TCGA, COSMIC, and ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) were used to identify somatic BRaf variants assessed using LABEL-seq predicted models.

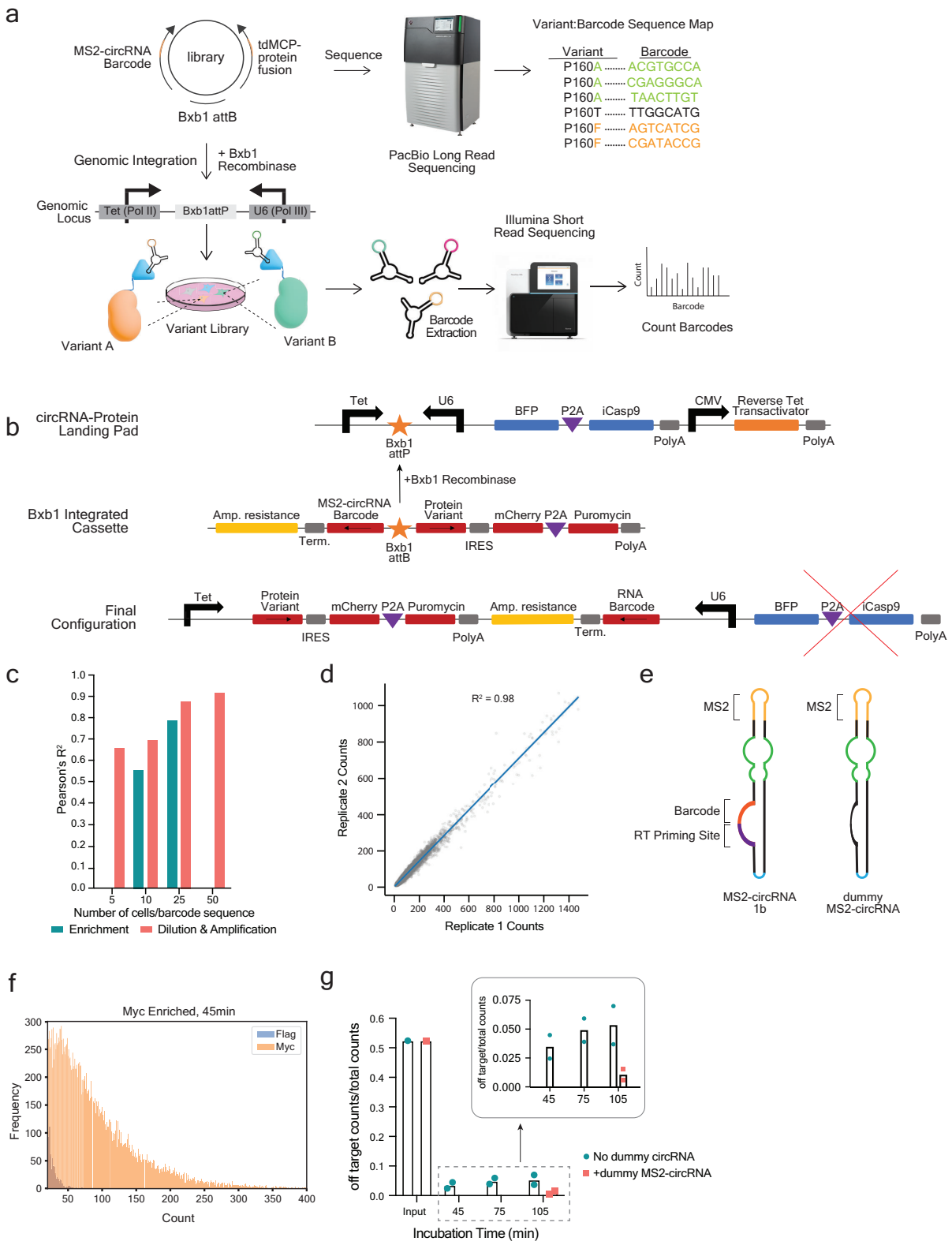
Code availability

Custom scripts written for parsing LABEL-seq data and plotting figures are available on Github (http://github.com/MalyLab/LABEL-seq-Scripts). The code used for PacBio subassembly is also available on Github (https://github.com/shendurelab/AssemblyByPacBio).

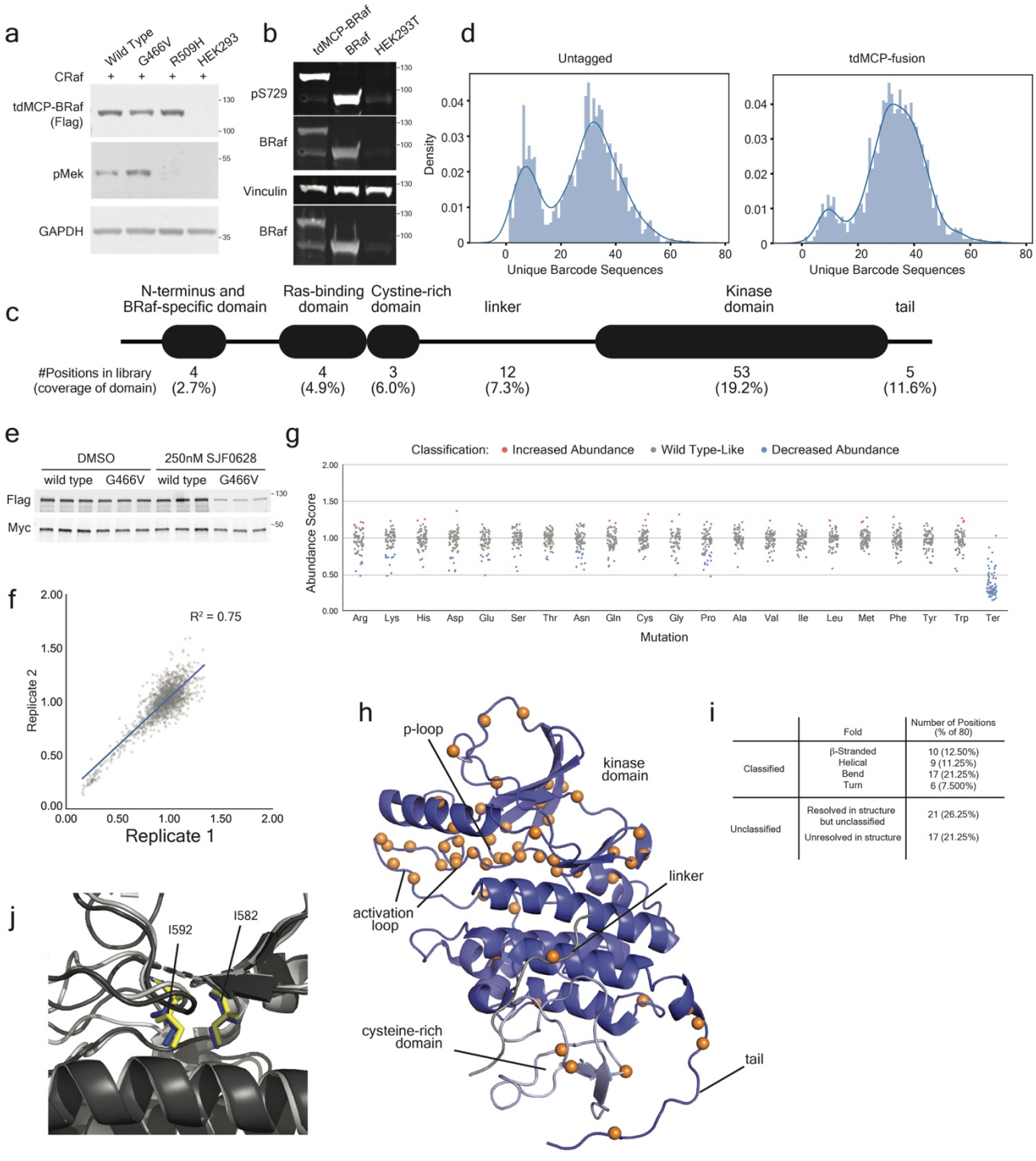
Extended Figures


a
Parent Architechure 1
Full length transcript:
Post-ribozyme excision:
Twister 1 + 5' ligation sequence
Twister 2 + 3' ligation sequence
ligation sequences
Parent Architechure 2
broccoli
MS2-circRNA 1 (this paper)
MS2-circRNA 2 (this paper)
barcode
MS2
b
Increased stem
Barcode
Priming Site
Increased Stem
MS2-circRNA 1
MS2-circRNA 1b
MS2-circRNA 2
MS2-circRNA 3
MS2-circRNA 4
MS2-circRNA 5
MS2-circRNA 6
MS2-circRNA 7
MS2-circRNA 8
MS2-circRNA 9
c
Parent arch. 1
Parent arch. 2
No Plasmid
1/2 Parent arch 1.
1/4 Parent arch 1.
200 nt
100 nt
Barcode sequence
d
C-term
N-term
FG-Loop
MS2 Hairpin
Protein of Interest
tandem dimer MCP (tdMCP) fusion:
epitope tag
MCP protomer 1
MCP protomer 2
monomeric MCP fusion:
e
Fraction MS2-circRNA captured
DMSO
Dasatinib
td-VE/AG, linker: Ala
td-VE/AG, linker: Gly-Ser-Thr-Ser-Gly-Ser-Gly
monomeric VE/AG
td-delFG, linker: Ala

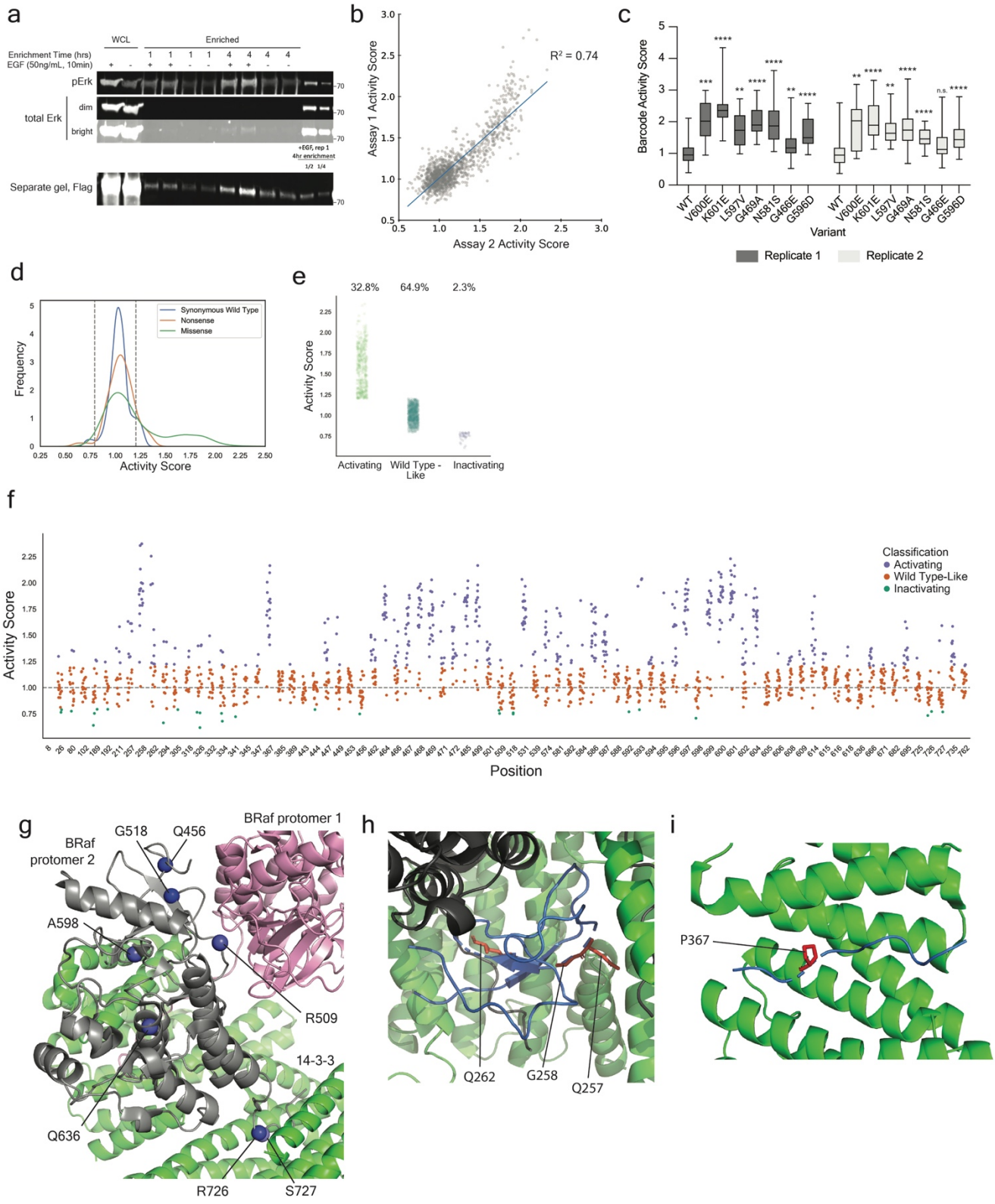
Extended Data Fig. 1 | Design, expression, and characterization of MS2-circRNAs and MCP protein fusions. a, In-silico folding predictions for circRNA architectures pre- and post-ribozyme excision. RNAfold was used for predictions and the colors indicate positional entropy values. Parent architectures do not contain the MS2 hairpin and are the designs published by Litke and Jaffrey[19] (addgene plasmids #124360 (left) and #124362 (right)). **b**, MS2-circRNA designs that contain a 16 nucleotide barcode and 19 nucleotide MS2 hairpin. MS2-circRNA 1b contains an added priming site specifically for reverse transcription. **c**, In-gel DFHBI-1T fluorescence of total RNA extracted from cells expressing the MS2-circRNAs shown in **b**. 2-3 barcode sequences of varying GC content for each MS2-circRNA were tested (n=1). Gels 1, 4, and 5 were non-denaturing 6% acrylamide gels and gels 2 and 3 were denaturing 10% acrylamide gels. **d**, Structure (PDB ID: 2BU1) of the MCP dimer bound to an MS2 hairpin (top). Linear representations of monomeric and tandem-dimer (td) MCP protein fusions (bottom). In a tandem dimer, the N-terminus of one MCP protomer is linked to the C-terminus of a second MCP protomer by a short linker. **e**, Co-enrichment of co-expressed MS2-circRNA 2 with Lck fused to different MCP variants. Lck was enriched with dasatinib-conjugated agarose beads in the presence of DMSO or a dasatinib competitor. Co-enriched MS2-circRNA was quantified with in-gel DFHBI-1T fluorescence (n=2 parallel transfections and enrichments. Representative blot shown. All MCP variants contain the high affinity V29I mutation. td- indicates tandem dimer designs. VE/AG variants contain V75E and A81G mutations, which prevent MCP oligomerization. delFG variants possess a FG loop (V67-A81) deletion, which also prevents MCP oligomerization. The linkers are positioned between MCP protomers. The td-delFG, linker Ala variant was used in all subsequent experiments and is referred to as ‘tdMCP’ throughout the manuscript.


a
MS2-circRNA Barcode
library
tdMCP-protein fusion
Bxb1 attB
Sequence
PacBio Long Read Sequencing
Variant:Barcode Sequence Map
Variant Barcode
P160A ACGTGCCA
P160A CGAGGGCA
P160A TAACTTGT
P160T TTGGCATG
P160F AGTCATCG
P160F CGATACCG
Genomic Integration
+ Bxb1 Recombinase
Genomic Locus
Tet (Pol II)
Bxb1attP
U6 (Pol III)
Variant A
Variant Library
Variant B
Barcode Extraction
Illumina Short Read Sequencing
Count
Barcode
Count Barcodes
b
circRNA-Protein Landing Pad
Tet
U6
BFP
P2A
iCasp9
CMV
Reverse Tet Transactivator
PolyA
PolyA
Bxb1 attP
+Bxb1 Recombinase
Bxb1 Integrated Cassette
Amp. resistance
MS2-circRNA Barcode
Protein Variant
mCherry P2A Puromycin
Term.
Bxb1 attB
IRES
PolyA
Final Configuration
Tet
Protein Variant
mCherry P2A Puromycin
Amp. resistance
RNA Barcode
U6
BFP
P2A
iCasp9
IRES
PolyA
Term.
PolyA
c
Pearson's R²
Number of cells/barcode sequence
Enrichment
Dilution & Amplification
d
R² = 0.98
Replicate 2 Counts
Replicate 1 Counts
e
MS2
MS2
Barcode
RT Priming Site
MS2-circRNA 1b
dummy MS2-circRNA
f
Myc Enriched, 45min
Flag
Myc
Frequency
Count
g
off target counts/total counts
off target/total counts
No dummy circRNA
+dummy MS2-circRNA
Input
Incubation Time (min)

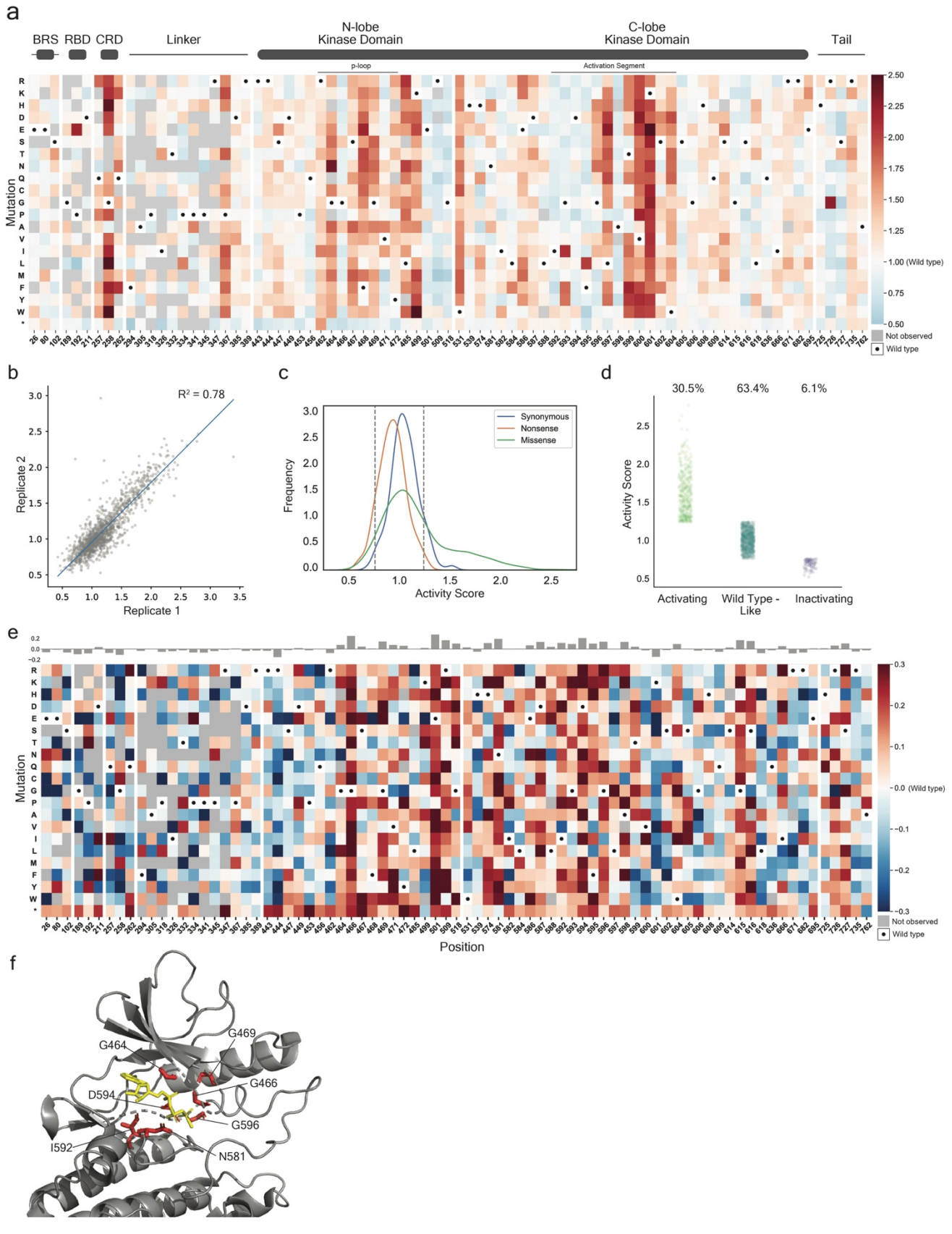
Extended Data Fig. 2 | Engineering of a multiplexed MS2-circRNA barcoding platform. a, Subassembly, which is used to link protein variants to MS2-circRNA barcode sequences, is achieved by sequencing the cloned plasmid library using long-read, high-throughput sequencing. In parallel, the plasmid library is integrated into the RPLP line to generate a cell line stably expressing tdMCP:MS2-circRNA complexes for use in profiling experiments. **b,** Schematic of the RPLP cell line. Recombination of Bxb1 attB-containing plasmids into the RPLP stably integrates protein variants and MS2-circRNA barcodes under genomic transcriptional control. Selection for successful recombinants is performed using AP1903 and/or puromycin. **c,** Parallel Flag immunoprecipitations were performed using serial dilutions of lysate generated from RPLP cells stably expressing Flag-tdMCP-BRaf WT and a library of ~20,000 MS2-circRNA 1b barcode sequences. Pearson's R^2 correlation coefficients of barcode sequence counts are shown for replicate enrichments or RT-PCR amplifications. **d,** Sequenced barcode counts derived from two independent expressions and Flag immunoprecipitations of Flag-tdMCP-BTK co-expressed with a library of MS2-circRNA 2 barcode sequences. Pearson's $R^2 = 0.98$. **e,** Schematics of MS2-circRNA 1b, the architecture used to encode variants in the tdMCP- and untagged BRaf variant libraries, and dummy MS2-circRNA, which is included in the lysis buffer to minimize barcode sequence missassociation. **f,** Relative frequencies of barcode sequence counts matched to Myc-tdMCP-BRaf and Flag-tdMCP-BRaf libraries after 45 min Myc immunoprecipitation. **g,** Quantification of the data shown in **f**. The y-axis is the fraction of barcode sequence counts encoding the Flag-tdMCP-BRaf (n=2 expression, lysis, enrichment, and sequencing replicates).


a
Wild Type
G466V
R509H
HEK293
CRaf
tdMCP-BRaf (Flag)
pMek
GAPDH
b
tdMCP-BRaf
BRaf
HEK293T
pS729
BRaf
Vinculin
BRaf
d
Untagged
tdMCP-fusion
Density
Unique Barcode Sequences
c
N-terminus and BRaf-specific domain
Ras-binding domain
Cystine-rich domain
linker
Kinase domain
tail
#Positions in library (coverage of domain)
4 (2.7%)
4 (4.9%)
3 (6.0%)
12 (7.3%)
53 (19.2%)
5 (11.6%)
e
DMSO
250nM SJF0628
wild type
G466V
wild type
G466V
Flag
Myc
g
Classification: Increased Abundance Wild Type-Like Decreased Abundance
Abundance Score
Arg Lys His Asp Glu Ser Thr Asn Gln Cys Gly Pro Ala Val Ile Leu Met Phe Tyr Trp Ter
Mutation
f
R² = 0.75
Replicate 2
Replicate 1
h
p-loop
kinase domain
linker
activation loop
cysteine-rich domain
tail
i
Fold
Number of Positions (% of 80)
Classified
β-Stranded 10 (12.50%)
Helical 9 (11.25%)
Bend 17 (21.25%)
Turn 6 (7.500%)
Unclassified
Resolved in structure but unclassified 21 (26.25%)
Unresolved in structure 17 (21.25%)
j
I592
I582

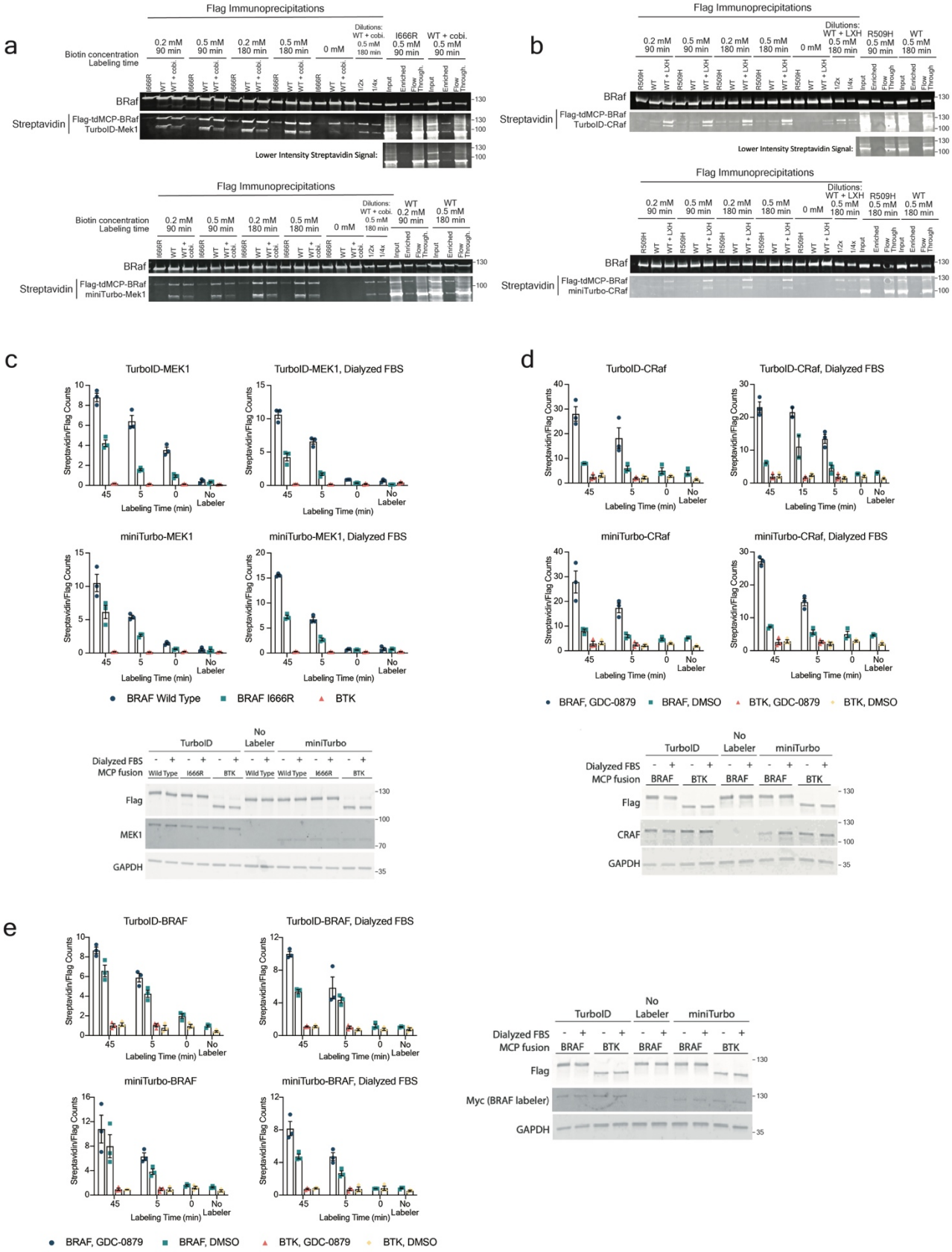
Extended Data Fig. 3 | BRaf variant library construction and abundance. a, Western blot analysis of RPLP cells co-expressing tdMCP-BRaf variants and CRaf (n=1). **b,** Representative western blot of HEK293T LLP cells stably expressing Flag-tdMCP-BRaf or untagged BRaf (n=2). Top two rows are the same samples as the bottom two rows run on separate blots. **c,** Domain architecture of BRaf and the number of positions in each domain included in the tdMCP- and untagged BRaf variant libraries. Positions were chosen by compiling all missense and nonsense BRaf variants appearing in the COSMIC database and The Cancer Genome Atlas and choosing positions where there are records of at least 3 total variants and at least 2 unique variants. **d,** Distributions of unique barcode sequences assigned to each variant in the MCP-tagged and untagged BRaf variant libraries. **e,** Western blots for the data quantifications shown in Fig. 3d (n=1). **f,** Scatterplot showing abundance score correlations between the two independent biological replicates shown in Fig. 3e (Pearson's $R^2 = 0.75$). **g,** Abundance scores for all variants in the tdMCP-BRaf variant library, grouped by substitution. **h,** Varied positions in the untagged and tdMCP-BRaf variant libraries shown as orange spheres on autoinhibited BRaf (PDB ID: 6NYB). **i,** DSSP (Define Secondary Structure of Proteins) structural predictions for BRaf library positions resolved in 6NYB. DSSP predictions for structures 7MFE and 5VR3 were referenced for positions not resolved in 6NYB. **j,** Alignment of BRaf's ATP-binding site in the active (black, PDB 4MNE) and autoinhibited (gray, PDB 6NYB) conformations, with the two positions (I582 and I592) with the lowest average abundance scores shown as sticks.


a
WCL
Enriched
Enrichment Time (hrs)
EGF (50ng/mL, 10min)
pErk
total Erk
dim
bright
+EGF, rep 1
4hr enrichment
1/2 1/4
Separate gel, Flag
b
Assay 1 Activity Score
Assay 2 Activity Score
R² = 0.74
c
Barcode Activity Score
WT
V600E
K601E
L597V
G469A
N581S
G466E
G596D
Variant
Replicate 1
Replicate 2
d
Frequency
Activity Score
Synonymous Wild Type
Nonsense
Missense
e
32.8%
64.9%
2.3%
Activity Score
Activating
Wild Type - Like
Inactivating
f
Activity Score
Position
Classification
Activating
Wild Type-Like
Inactivating
g
BRaf protomer 1
BRaf protomer 2
G518
Q456
A598
R509
14-3-3
Q636
R726
S727
h
Q262
G258
Q257
i
P367

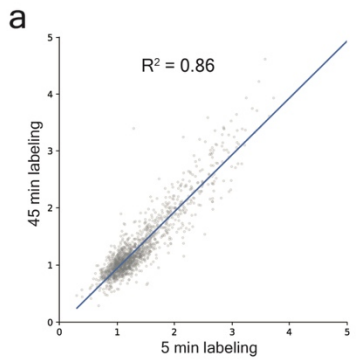
Extended Data Fig. 4 | BRaf variant activity with overexpressed CRaf. a, Western blot analysis of HEK293T cells expressing the pErk reporter and treated with EGF **b,** Correlation (Pearson's R^2 = 0.74) between BRaf variant activity scores with overexpressed CRaf derived from the mean of replicates 1 and 2 (y-axis, represented in Fig. 4e) and the mean of replicates 3 and 4 (x-axis). Each replicate involved independent cell culturing, lysis, immunoprecipitations, and quantification. Replicates 1 and 2 were performed on a separate day than replicates 3 and 4. **c,** LABEL-seq activity scores with overexpressed CRaf for previously characterized BRaf variants. Center line = median, box limits extend from 25-75th percentiles, and whiskers extend from min to max. Each box corresponds to the barcode scores from one replicate. P and n values are listed in Supplementary Table 4. **d,** LABEL-seq activity scores for BRaf variants with overexpressed CRaf. Gray dashed lines indicate the activity scores we defined as activating or inactivating (>2 standard deviations above or below the mean of the synonymous WT distribution). **e,** Classification of BRaf variants as activating, Wild Type-like, or inactivating when overexpressed with CRaf. **f,** Activity scores of BRaf variants when overexpressed with CRaf. **g,** The 10 positions (blue spheres) with the lowest average activity scores projected onto an active BRaf complex (PDB ID: 6Q0K). Positions 26, 189, and 443 were not resolved in this structure. **h,** The autoinhibited BRaf complex (PDB ID: 6NYB) with BRaf's CRD (blue), non-CRD BRaf positions (black), and 14-3-3 (green) highlighted. G258 is located within the CRD, which plays a role in autoinhibition, Ras binding, and membrane localization[40]. In autoinhibited BRaf, the phi-psi angles at G258 are only energetically accessible for glycine, suggesting that substitutions at this position alter the backbone conformation and disrupt autoinhibitory contacts. **i,** Autoinhibited BRaf (PDB ID: 6NYB) with BRaf P367 shown as red sticks, non-P367 BRaf positions shown in blue and 14-3-3 in green. P367 is located within the CRD-kinase domain linker, which interacts with 14-3-3 in the autoinhibited complex. Based on the known binding specificity of 14-3-3 proteins[42], P367 substitutions likely lead to BRaf activation by disrupting autoinhibitory 14-3-3 interactions.

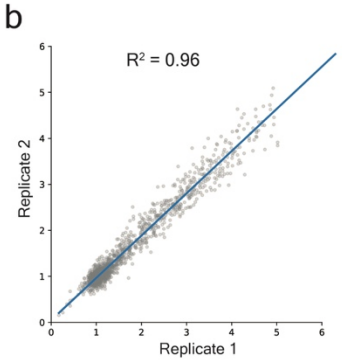

a
BRS
RBD
CRD
Linker
N-lobe
Kinase Domain
C-lobe
Kinase Domain
Tail
p-loop
Activation Segment
Mutation
Not observed
Wild type
1.00 (Wild type)
b
R² = 0.78
Replicate 2
Replicate 1
c
Synonymous
Nonsense
Missense
Frequency
Activity Score
d
30.5%
63.4%
6.1%
Activity Score
Activating
Wild Type - Like
Inactivating
e
Mutation
Position
0.0 (Wild type)
Not observed
Wild type
f
G464
G469
G466
D594
G596
I592
N581

Extended Data Fig. 5 | BRaf protein variant activity measured without overexpressed CRaf. **a,** Sequence-activity map for the untagged BRaf variant library without CRaf overexpression. Activity scores are the average of two replicates, each involving independent pErk reporter transfections, lysis, parallel Flag and pErk immunoprecipitations, and quantification by high-throughput sequencing. Black dots indicate the WT amino acid and gray tiles indicate missing data. Position-averaged activity scores are shown in the bar graph above the heatmap. **b,** Scatterplot showing activity score (BRaf variants without CRaf overexpression) correlations between the two independent replicates represented in **a** (Pearson's $R^2 = 0.78$). **c,** LABEL-seq activity scores for BRaf variants without overexpressed CRaf. Gray dashed lines indicate the activity scores we defined as activating and inactivating (>2 standard deviations above or below the mean of the synonymous WT distribution). **d,** Classification of BRaf variants as activating, Wild Type-like, or inactivating in the absence of CRaf overexpression. **e,** Sequence-CRaf-dependent activity map for BRaf variants. Black dots indicate the WT amino acid and gray tiles indicate missing data. Position-averaged scores are shown in the bar graph above the heatmap. **f,** Highly conserved positions (red sticks) across the human kinome projected onto a structure of BRaf bound to ATP-γ-S (PDB ID: 6NYB).


a
Flag Immunoprecipitations
Biotin concentration
Labeling time
0.2 mM 90 min
0.5 mM 90 min
0.2 mM 180 min
0.5 mM 180 min
0 mM
Dilutions: WT + cobi. 0.5 mM 180 min
I666R 0.5 mM 90 min
WT + cobi. 0.5 mM 90 min
WT 0.2 mM 90 min
WT 0.5 mM 180 min
I666R
WT
WT + cobi.
1/2x
1/4x
Input
Enriched
Flow Through
BRaf
Streptavidin
Flag-tdMCP-BRaf
TurboID-Mek1
miniTurbo-Mek1
Lower Intensity Streptavidin Signal:
b
Flag Immunoprecipitations
R509H
WT + LXH
R509H 0.5 mM 90 min
Dilutions: WT + LXH 0.5 mM 180 min
1/2x
1/4x
BRaf
Streptavidin
Flag-tdMCP-BRaf
TurboID-CRaf
miniTurbo-CRaf
Lower Intensity Streptavidin Signal:
c
TurboID-MEK1
TurboID-MEK1, Dialyzed FBS
miniTurbo-MEK1
miniTurbo-MEK1, Dialyzed FBS
Streptavidin/Flag Counts
Labeling Time (min)
No Labeler
BRAF Wild Type
BRAF I666R
BTK
TurboID
miniTurbo
Dialyzed FBS
MCP fusion
Wild Type
I666R
BTK
Flag
MEK1
GAPDH
d
TurboID-CRaf
TurboID-CRaf, Dialyzed FBS
miniTurbo-CRaf
miniTurbo-CRaf, Dialyzed FBS
BRAF, GDC-0879
BRAF, DMSO
BTK, GDC-0879
BTK, DMSO
BRAF
CRAF
e
TurboID-BRAF
TurboID-BRAF, Dialyzed FBS
miniTurbo-BRAF
miniTurbo-BRAF, Dialyzed FBS
Myc (BRAF labeler)

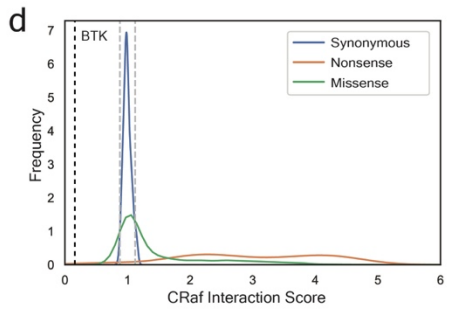
Extended Data Fig. 6 | Optimization of proximity labeling conditions. a, Biotinylation levels of Flag-tdMCP-BRaf variants in HEK293 cells co-expressing TurboID-Mek1 (top) or miniTurbo-Mek1 (bottom), treated with 10 μM cobimetinib or DMSO for 4 h, and labeled with biotin. **b,** Biotinylation levels of Flag-tdMCP-BRaf variants in HEK293 cells co-expressing TurboID-CRaf (top) or miniTurbo-CRaf (bottom), treated with 10 μM LXH254 or DMSO for 2 h, and labeled with biotin. (**c**)**,** Biotinylation levels (represented as the ratio of streptavidin and Flag barcode sequence counts) of Flag-tdMCP-BRaf variants and the background control Flag-tdMCP-BTK in HEK293T cells co-expressing TurboID-Mek1 and MS2-circRNAs. Cells were cultured in regular or dialyzed FBS and treated with biotin for 0, 5, or 45 min. (**d,e**), Biotinylation levels (represented as the ratio of streptavidin and Flag barcode sequence frequencies) of Flag-tdMCP-BRaf variants and the background control Flag-tdMCP-BTK in DMSO- or GDC-0879-treated HEK293T cells co-expressing TurboID-CRaf (**d**) or TurboID-BRaf (**e**). Cells were cultured in regular or dialyzed FBS and treated with biotin for 0, 5, or 45 min. For **c**-**e**, after biotin treatment, cells were lysed and pooled, parallel streptavidin and Flag enrichments were performed, co-enriched barcodes were sequenced by high-throughput sequencing, and the ratio of streptavidin and Flag counts for each co-enriched MS2-circRNA barcode sequence was calculated (n=3 barcodes/treatment, error bars represent mean +/- SEM). Whole-cell lysate samples were reserved for western blots, which are shown below the bar graphs (corresponds to the data shown in Fig. 5c,e,f). TurboID-Mek1 (dialyzed FBS), TurboID-CRaf (dialyzed FBS), and TurboID-BRaf (dialyzed FBS) data are the same as in Fig. 5c,e,f.

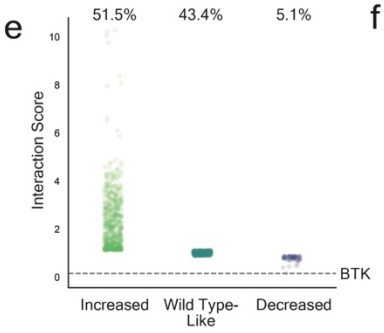


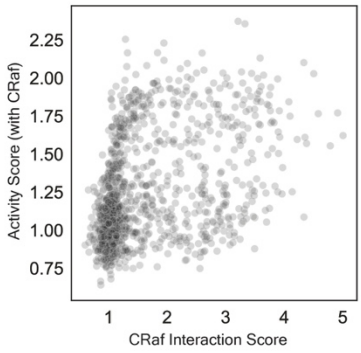


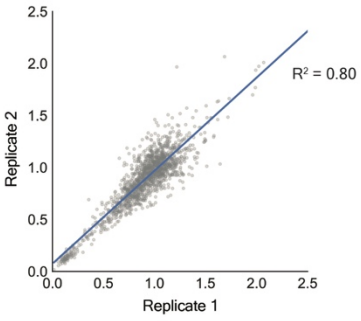
c

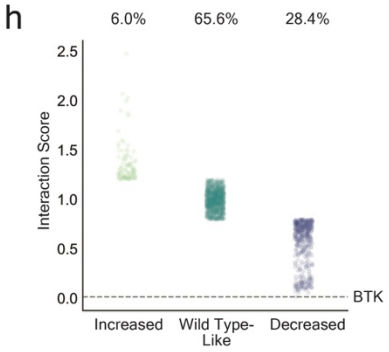
	Replicate 1	Replicate 2
WT	1.00 +/- 0.05	1.00 +/- 0.04
BTK	0.17 +/- 0.02	0.15 +/- 0.01
D594N	3.05 +/-0.22	2.92 +/- 0.32
D594G	4.24 +/- 0.31	4.20 +/- 0.45
G466V	3.81 +/- 0.25	3.42 +/- 0.24
G466E	3.57 +/- 0.19	3.28 +/- 0.18

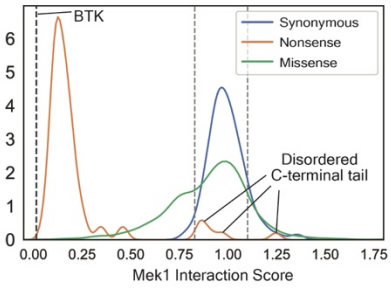






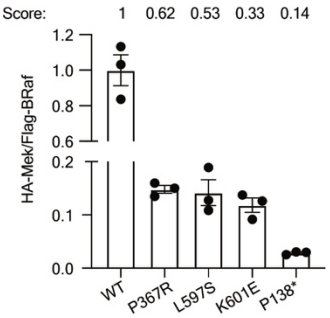


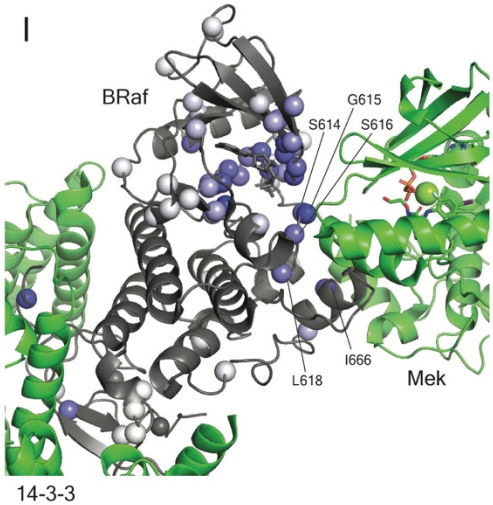


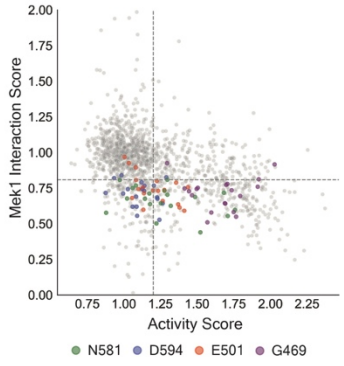


j

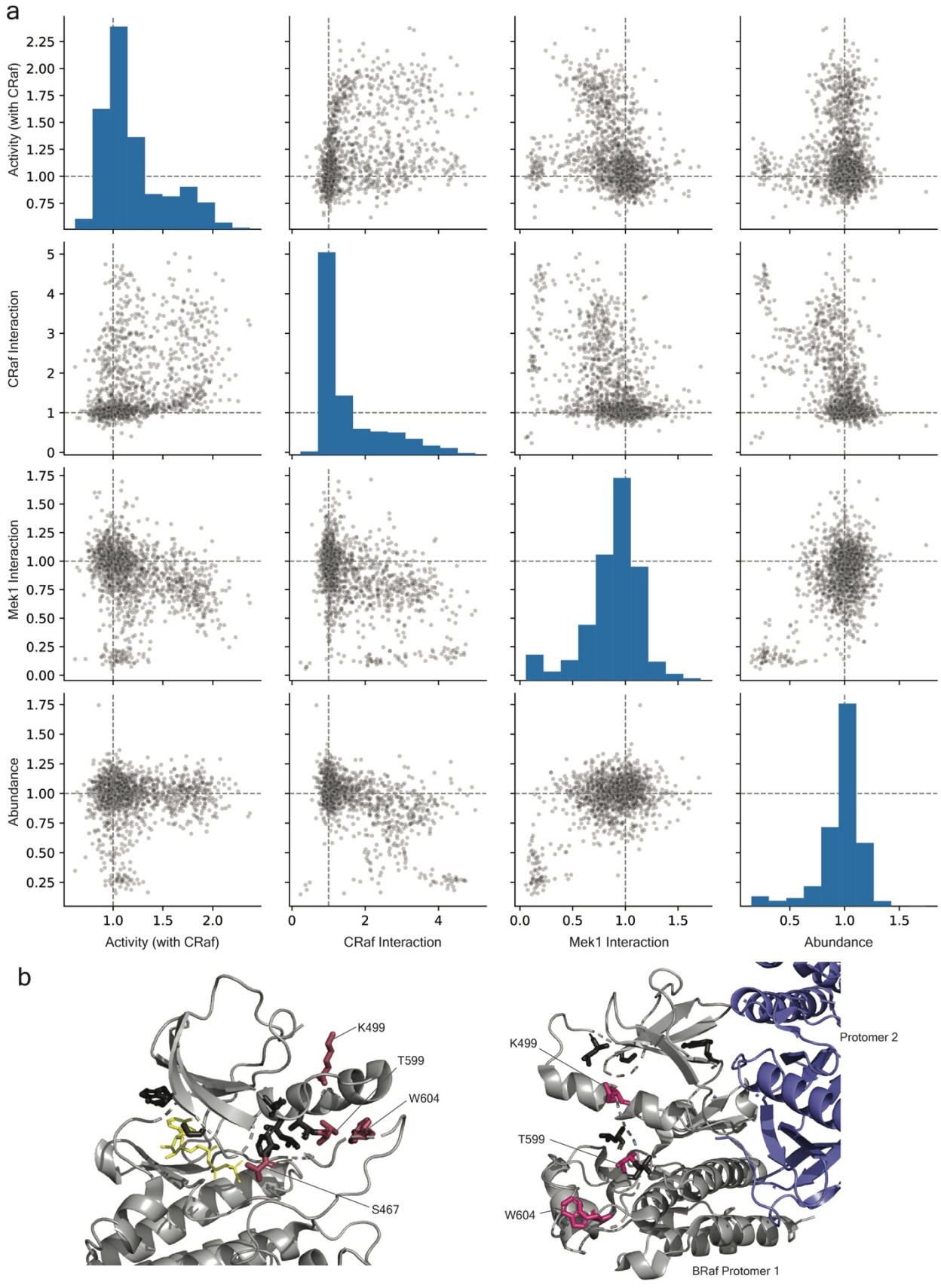
	Replicate 1	Replicate 2
WT	1.03 +/- 0.05	1.02 +/- 0.05
I666R	0.24 +/- 0.02	0.21 +/- 0.02
BTK	0.01 +/- 0.00	0.02 +/- 0.00



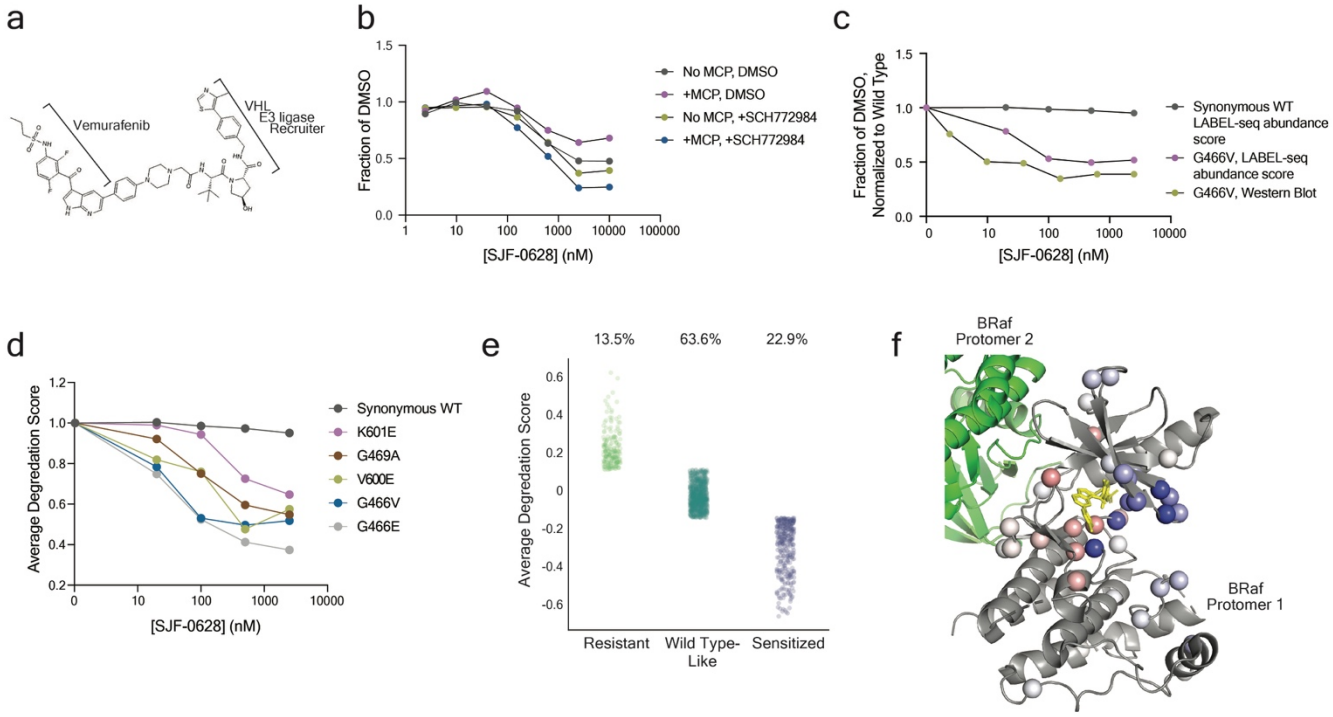


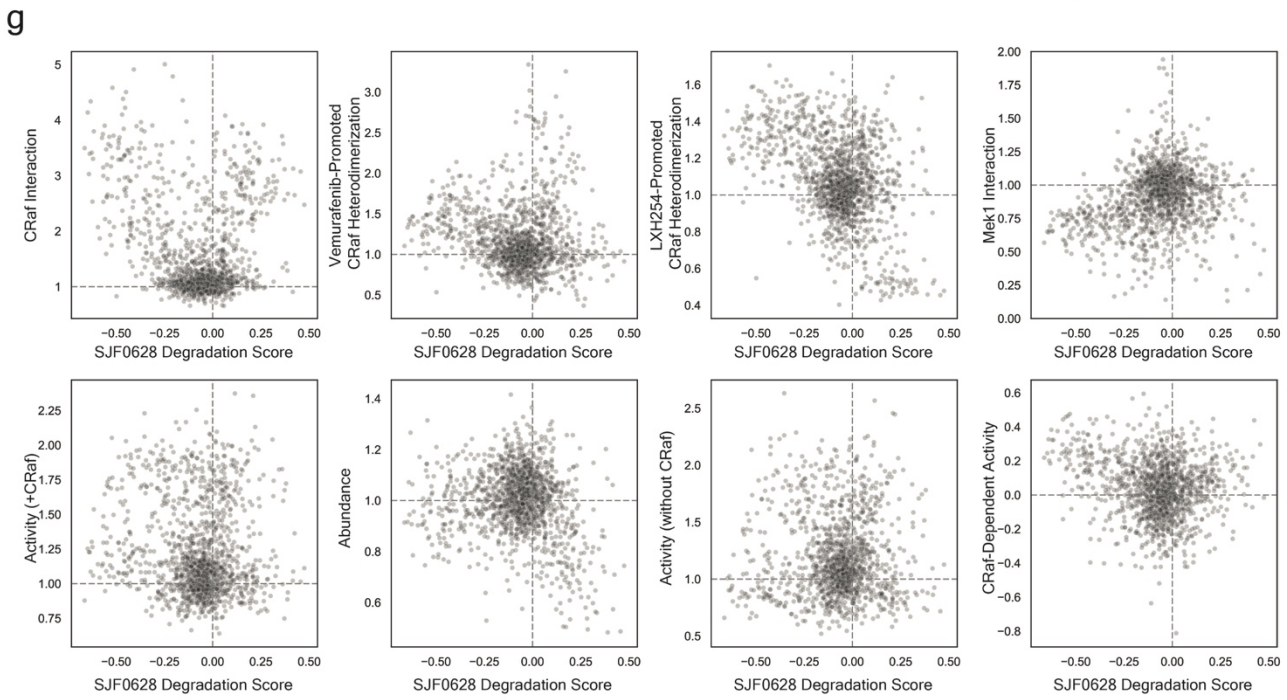


Extended Data Fig. 7 | BRaf variant interactions with CRaf and Mek1. a, CRaf interaction score correlation between 5 and 45 min biotin labelings . Scores are the mean of two replicates corresponding to parallel biotin labelings, enrichments, and quantifications. **b,** CRaf interaction score correlation between the replicates represented in Fig. 5h. **c,** Interaction scores (Fig. 5h) for BRaf variants previously shown to exhibit elevated heterodimerization with CRaf. Mean +/- SEM of all barcode scores. **d,** CRaf interaction score distributions. The gray dashed lines indicate the score values we defined as increased or decreased CRaf interaction (>2 standard deviations above or below the mean of the synonymous WT distribution). Black dashed line = BTK score. **e,** Classification of BRaf variants as increased, Wild Type-like, or decreased interaction with CRaf. **f,** Scatterplot comparing activity scores with CRaf to CRaf interaction scores. **g,** Mek1 interaction score correlation between the two independent replicates shown in Fig. 5j. **h,** Classification of BRaf variants as increased, Wild Type-like, or decreased interaction with Mek1. **i,** Mek1 interaction score distributions. The gray dashed lines indicate score values we defined as increased and decreased interaction (>2 standard deviations above or below the mean of the synonymous WT distribution). Black dashed line = BTK score. **j,** Mek1 interaction scores for Flag-tdMCP BRaf WT, I666R (decreased interaction control) and Flag-tdMCP-BTK. Values shown are mean +/- SEM of all barcode scores. **k,** HA-Mek1 and Flag-BRaf variant co-immunoprecipitations. Mek1 interaction scores are shown above the plot (n=3, value shown = mean, error bars = SEM). **l,** Number of variants at each position classified as decreased Mek1 interaction (Fig. 5j) projected onto autoinhibited BRaf (PDB ID: 6NYB). White = 0, darkest blue = 19. **m,** Mek1 interaction and activity (with CRaf) scores. BRaf variants at the highlighted positions showed similarly decreased Mek1 interaction scores but varied activity scores, indicating that reduced Mek1 interactions cannot solely be attributed to the formation of lower affinity pMek[5]. Scores to the right of the vertical dashed line are classified as activating and scores below the horizontal dashed line are classified as decreased Mek1 interaction.

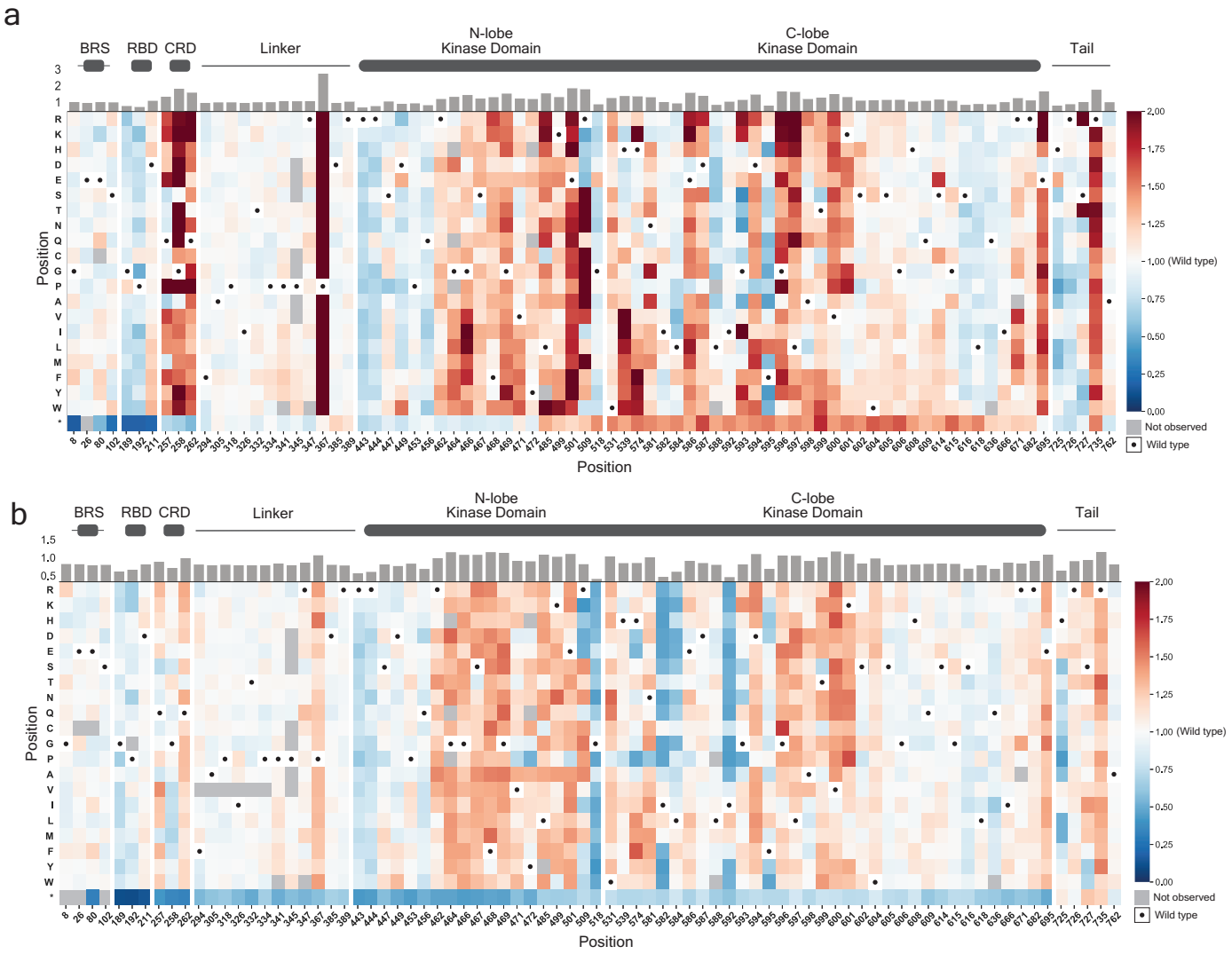

a
Activity (with CRaf)
CRaf Interaction
Mek1 Interaction
Abundance
b
K499
T599
W604
S467
Protomer 2
BRaf Protomer 1

Extended Data Fig. 8 | Pairwise comparisons of protein variant scores. a, All pairwise comparisons of activity scores (with CRaf), CRaf interaction scores, Mek interaction scores, and abundance scores. The histograms are the relative frequency of each score in the library. **b,** Cluster 1 (black sticks) and cluster 2 (maroon sticks) positions projected onto autoinhibited BRaf (left, PDB ID: 6NYB) and the active BRaf dimer (right, PDB ID: 4MNE).


a
Vemurafenib
VHL
E3 ligase
Recruiter
b
Fraction of DMSO
[SJF-0628] (nM)
No MCP, DMSO
+MCP, DMSO
No MCP, +SCH772984
+MCP, +SCH772984
c
Fraction of DMSO, Normalized to Wild Type
[SJF-0628] (nM)
Synonymous WT LABEL-seq abundance score
G466V, LABEL-seq abundance score
G466V, Western Blot
d
Average Degredation Score
[SJF-0628] (nM)
Synonymous WT
K601E
G469A
V600E
G466V
G466E
e
13.5%
63.6%
22.9%
Average Degredation Score
Resistant
Wild Type-Like
Sensitized
f
BRaf Protomer 2
BRaf Protomer 1


g
CRaf Interaction
Vemurafenib-Promoted CRaf Heterodimerization
LXH254-Promoted CRaf Heterodimerization
Mek1 Interaction
Activity (+CRaf)
Abundance
Activity (without CRaf)
CRaf-Dependent Activity
SJF0628 Degradation Score

Extended Data Fig. 9 | SJF0628-mediated degradation of BRaf protein variants. a, Structure of SJF0628. **b,** Expression levels of BRaf WT without tdMCP (No MCP) and tdMCP-BRaf WT expressed in HEK293T cells treated with various concentrations of SJF0628 alone or the Erk inhibitor SCH772984 and various concentrations of SJF0628. Expression measured using western blotting. Mean of n=2 replicates shown. **c,** SJF0628-promoted degradation of tdMCP-BRaf synonymous WT and G466V variants measured using western blotting (n=1) and the LABEL-seq abundance assay. **d,** Average LABEL-seq-derived degradation scores at the concentrations of SJF0628 shown for five BRaf variants previously demonstrated to be sensitive to SJF028 (n=2). **e,** Classification of BRaf variants as resistant, Wild Type-like, or sensitized to degradation by 2.5 μM SJF0628. Variants with degradation scores >2 standard deviations above or below the mean of the synonymous WT distribution were classified as resistant or sensitized, respectively. **f,** Position-averaged SJF0628 degradation scores (represented as spheres) at 2.5 μM SJF0628 projected onto a structure of BRaf's kinase domain bound to Vemurafenib (PDB ID: 3OG7). The average degradation score at a position is represented by the shade of red or blue: darkest red = 0.18, darkest blue = - 0.51. **g,** Scatterplots showing relationship between SJF0628 degradation score and other LABEL-seq biochemical measurements. Data shown for missense variants only. Gray dashed lines correspond to the WT score.



Extended Data Fig. 10 | a-b, Sequence-CRaf interaction maps for variants in the tdMCP-BRaf library co-expressed with TurboID-CRaf, treated with Vemurafenib (**a**) or LXH254 (**b**), and labeled with biotin. The interaction scores shown are the average of two replicates, each involving an independent cell culturing, biotin labeling, parallel streptavidin and Flag enrichments, and quantification with high-throughput sequencing. Black dots indicate the WT amino acid and gray tiles indicate missing data. Position-averaged CRaf interaction scores are shown in the bar graph above the heatmaps.

Supplementary Methods

Generating single protein variant barcode libraries

To generate a library of barcodes, attB plasmid containing genes for tdMCP-protein fusions and a MS2circRNA were digested overnight at 37°C to remove existing barcode sequence (2 µg plasmid, 5 µL 10x CutSmart buffer, 2 µL BsrGI-HF, nuclease-free water to 50uL). The next day, fragments were dephosphorylated by adding 5 µL antarctic phosphatase buffer and 2 µL enzyme and incubating for 1 h at 37°C. Dephosphorylated fragments were separated by gel electrophoresis and the band corresponding to digested vector was extracted using a Zymo gel extraction kit and eluted into 20 µL nuclease-free water.

To generate a double-stranded barcode fragment (dsBC), 50pmol of the single stranded barcode oligo (JJS29), 50pmol of the second strand synthesis primer (ssSynth primer, JJS30), 5 µL NEB2.1 10x buffer, and nuclease-free water to a total volume of 48 µL were combined. Oligos were annealed by heating to 94°C for 5 min, ramping 0.1°C/s to 71°C (Tm) and holding for 5 min, then ramping 0.1°C/s to 37°C. ssSynth primer was extended by adding 1 µL 10mM dNTP's and 1 µL Klenow polymerase then incubating at 37°C for 1hr. The resulting dsBC was cleaned using a Zymo DNA cleanup column and eluted into 12 µL nucleasefree water.

dsBC was cloned into the digested vector by combining 200ng vector (~0.035pmol, 8uL), 10ng dsBC (~0.175pmol, 2uL) and 10 µL 2x NEBuilder® HiFi DNA Assembly Master Mix and incubating at 50°C for 1hr. A separate vector-only reaction was run in parallel using 200ng vector, 10 µL NEBuilder Master Mix, and nuclease-free water to 20 µL (without barcode) to

estimate vector-only background transformants. After incubation, reactions were cleaned using a Zymo DNA cleanup column and eluted into 10 µL nucleasefree water.

Cleaned reactions were transformed into NEB® 5-alpha Competent *E. coli* following the manufacturer's instructions. Following outgrowth, transformed cells were added to 50 mL LB broth supplemented with 100 µg/mL carbenicillin. The library was bottlenecked to the desired number of barcodes by making serial dilutions of the culture in LB+carb. 200 µL of each dilution was plated on LB+carb plates to enumerate transformants. Liquid cultures were grown at 37°C overnight to saturation and DNA was extracted using QIAGEN® Plasmid Plus Midi Kit.

Cloning tdMCP-BRaf variant library

A site saturation mutagenesis library of full length BRaf consisting of variants at 80 chosen residues was purchased from Twist Bioscience. The double stranded DNA pool contained the entire 3xFlag-tdMCP-BRaf open reading frame with flanking 5' (TGAGCGGACTCTAGCGTTTAAAC) and 3' (CGCCGGTCTTAACTGATCACTGA) sequences complementary to the vector. The gene fragment library was cloned into a linearized attB vector using Gibson assembly as follows. The fully cloned plasmid is available on addgene (#226163).

To generate a linearized vector the following was combined and incubated at 37°C overnight: 2 µg attB library vector, 1.5 µL AflII, 1.5 µL NheI-HF, 5 µL CutSmart 10x buffer, nuclease-free water to 50 µL. The next day, 2.5 µL recombinant shrimp alkaline phosphatase was added and incubated at 37°C for 2 hr.

Digested DNA was run on a 1% agarose gel, cleaned using a Zymo gel extraction kit, and eluted into 30 µL nuclease-free water.

The gene fragment library was cloned into the linearized vector using a 1:2 molar ratio of vector:insert by combining the following and incubating at 50°C for 1hr: 80ng linearized vector, 95ng gene fragment library, 40 µL NEBuilder® HiFi DNA Assembly Master Mix, and nuclease-free water to 80uL. To estimate background transformants, a vector-only reaction was run in parallel using 80ng linearized vector, 40 µL NEBuilder Master Mix, and water to 80 µL (no gene fragment library). After incubation, reactions were cleaned using a Zymo DNA cleanup column and eluted into 20 µL nuclease-free water.

Cleaned reactions were transformed into NEB® 5-alpha Competent E. coli (High Efficiency) per the manufacturer's instructions. Briefly, 4 replicate transformations were performed on the library, each using 40 µL cells and 4 µL cleaned gibson reaction and single replicate was performed in parallel to estimate vector-only transformants. Cells were incubated with DNA on ice for 30 min, heat shocked at 42°C for 30 s, and put back on ice for 5 min. Subsequently, 760 µL SOC was added to each replicate and cells were incubated at 37°C for 1 h at 220 rpm. All library replicates were then pooled and added to 40 mL LB supplemented with 100 µg/mL carbenicillin. From this culture, 200uL, 20uL, and 2 µL was plated onto carbenicillin-selective agar plates to estimate transformants. 200 µL vector-only was also plated onto agar to estimate background. The 40 mL liquid culture was grown overnight to saturation at 37°C and 220 rpm. The following day DNA was harvested using a QIAGEN Plasmid Plus Midi Kit.

Variant library was barcoded and bottlenecked to desired number of transformants as described in 'Generating single protein variant barcode libraries'. All oligonucleotides used to generate libraries and amplicons are reported in Supplementary Table 5.

Cloning untagged BRaf variant library

To clone a BRaf variant library without the 3xFlag-tdMCP fusion (fully constructed plasmid available on addgene, #226164), the attB 3xFlag-tdMCP-BRaf variant library, cloned into the attB vector, was digested by combining the following and incubating overnight at 37°C: 2 µg plasmid + 5 µL CutSmart buffer (10x) + 2 µL AflII + 2 µL BlpI + nuclease-free water to 50 µL. The next day, 1 µL recombinant shrimp alkaline phosphatase was added to the reaction and incubated at 37°C for 1hr. The digest was run on a 1% agarose gel, the correct fragment extracted, then eluted into 30 µL nuclease-free water.

To generate a double stranded bridging fragment the following was combined: 1 µg (36.1pmol) JJS6 + 870ng (36pmol) JJS7 + 5 µL NEBuffer 2 (10x) + nuclease-free water to 48uL. Oligos were annealed by denaturing at 94°C for 5 min, ramping 0.1°C/s to 65.5°C (Tm), holding at 65.5°C for 5 min, then ramping 0.1°C/s to 37°C and holding. 1 µL dNTP's (10mM each) and 1 µL Klenow polymerase was added to the annealed oligos. Oligos were extended by incubating at 37°C for 1hr, inactivating the enzyme at 75°C for 20 min, re-annealing the DNA by ramping 0.1°C/s to 42°C, then holding at 4°C. DNA was cleaned and concentrated using a Zymo DNA cleanup column, then eluted into 30 µL nuclease-free water.

Digested variant library was re-circularized using gibson assembly at a 1:5 molar ratio of vector:bridge as recommended by the manufacturer. Briefly, the following was combined and incubated at 50°C for 1hr: 200ng (5 µL) digested vector, 20ng (2 µL) dsBridge, 13 µL nuclease-free water, 20 µL 2x HiFi DNA Assembly master mix. After incubation, reactions were cleaned using a Zymo DNA cleanup column and eluted into 20 µL nuclease-free water.

Cleaned gibson reactions were transformed into NEB® 5-alpha Competent E. coli (High Efficiency), transformants enumerated, and DNA harvested as described in 'Cloning MCP-tagged gene library into attB library vector'.

Variant library was barcoded and bottlenecked to desired number of transformants as described in 'Generating single protein variant barcode libraries'.

Generating stable cell lines expressing single protein variants

Stable cell lines expressing protein and RNA barcodes were generated in the HEK293 RPLP. Stable lines not expressing circRNA barcodes were generated in the HEK293T lenti landing pad (LLP-iCasp9-Blast[24]). The following protocol applies to both cell lines.

Cells were maintained in high glucose DMEM supplemented with 10% FBS and 2 µg/mL doxycycline (DOX, 0.01% DMSO final concentration). Genes of interest were cloned into Bxb1 attB-containing plasmids using standard Gibson assembly cloning techniques. To stably recombine the attB plasmids under control of the landing pad promoters, cells were plated the day before transfection in 6-well dishes to 40% confluence using DOX-free media. The

following day, cells were transfected with a mixture of 1.7 µg attB plasmid + 0.2 µg pCAG-NLS-Bxb1 (Addgene#51271) + 0.2 µg pMAX-GFP + 7.8 µL Fugene 6 (for LLP lines) or Fugene HD (for RPLP lines) + 100 µL serum-free DMEM. 48-hours after transfection, cells were expanded to 10cm dishes in media containing 2 µg/mL DOX. 24-48 hr after DOX induction, media was supplemented with 10nM AP1903 to select for successful recombinants. 12-24 hr after addition of AP1903, media was exchanged for AP1903-free media (+2 µg/mL DOX), the resulting stables were expanded in +DOX media, and stocks frozen until required.

Generating variant library cell lines

Barcoded libraries were stably recombined into the HEK293 RPLP cell line and maintained in high glucose DMEM supplemented with 10% FBS and 2 µg/mL doxycycline (0.01% DMSO final concentration). The day before transfection, RPLP cells were plated into two 15 cm dishes using DOX-free media (+10% FBS) to reach ~85% confluent during transfection. On the day of transfection the media was replaced with doxfree DMEM (+10% FBS) supplemented with 100nM SCH772984. Transfection mixes were prepared by combining per 15cm dish: 16.5 µg attB-plasmid library + 3 µg pCAG-NLS-Bxb1 + 2 µg pMAX-GFP + 2150 µL serum-free DMEM + 80 µL FugeneHD, then incubated at room temperature for 15 min and added dropwise to cells. 48 hr after transfection cells were expanded and maintained in DMEM (+10% FBS) supplemented with 2 µg/mL doxycycline and 100nM SCH772984 moving forward. 4 days after transfection cells were further expanded and a sample analyzed by flow cytometry to quantify recombination efficiency via the mCherry reporter. 6 days after transfection, selection was initiated by supplementing media with 10nM AP1903. After 6 hr, media was replaced with fresh +AP1903 media. 24 hr after initiating selection, cells were further expanded in media

supplemented with 1 μg/mL puromycin (+2 μg/mL DOX, +100nM SCH772984) until ready to freeze aliquots.

Preparation of RNA for Illumina sequencing

To prepare barcodes for sequencing, RNA was extracted from the enrichment beads using IBI scientific TRI-Isolate RNA Pure Kit and eluted into 50 μL nuclease-free water. RNA was further concentrated using Zymo RNA Clean & Concentrator, eluted into 10 μL nuclease-free water, and stored at -80°C. Following the manufacturer's instructions, concentrated RNA was reverse transcribed using SuperScript™ IV Reverse Transcriptase and custom primer JJS1. Briefly, primer/template mixes corresponding to a single library enrichment were generated by combining 7 μL concentrated RNA + 0.625 μL JJS1 (2uM) + 0.625 μL dNTP's (10mM ea.). The primer was annealed by heating mixes to 75°C for 5 min then set on ice. 0.625 μL DTT (0.1M) + 0.625 μL Ribolock RNase inhibitor + 0.625 μL SSIV enzyme + 2.5 μL SSIV buffer were added to each reaction. cDNA was generated by heating reaction at 45°C for 30 min, 80°C for 10 min, then holding at 4°C.

cDNA was either stored at -20 or immediately used in PCR round #1 to generate double stranded DNA and add amplicon adapters, amplifying only cDNA derived from circRNA. Specifically, each RT reaction was split into two PCR reactions, each including 6.25 μL RT reaction + 2.5 μL JJS8 (10uM) + 2.5 μL JJS9-JJS15 (10uM) + 13.75 μL nuclease-free water + 1.5 μL DMSO + 25 μL Q5® Hot Start High-Fidelity 2X Master Mix. Cycling included 98°C for 30 s, five cycles of (98°C for 10 s > 64°C for 20 s > 72°C for 30 s), 72°C for 2 min, and 4°C hold. Replicate reactions were pooled then cleaned with Sera-Mag Select beads using a 2.5:1 ratio of

beads:reaction and eluted using 20 µL nuclease-free water. Illumina adapters were added in a second round of PCR. This reaction included 5-20 µL of cleaned reaction #1 + 2.5 µL JJS16-JJS23 (10uM) +2.5 µL JJS24 (10uM) + 0.5 µL 10x Sybr gold + 25 µL Q5 MM + 1.0 µL DMSO + nuclease-free water to 51.5uL. PCR cycling included one round at 98°C for 30s, then cycling at (98°C for 10 s > 63°C for 20 s > 72°C for 30 s). Reaction progress was monitored on a BioRad CFX Connect and stopped when RFU reached ~3,000 (1015 cycles). Amplicons were pooled based on final RFU values, separated by gel electrophoresis and the band at ~220bp excised using Freeze 'N Squeeze™ DNA Gel Extraction Spin Columns. The DNA was further concentrated using Zymo DNA cleanup columns and quantified using Qubit™ dsDNA HS Assay Kit. Amplicons were sequenced on a NextSeq™ 550 or 2000 using custom primers JJS25-28, 16 cycle pairedend reads and 8 cycle dual indexing.

Preparation of RNA for Genewiz Amplicon-EZ sequencing

Enriched RNA was extracted from TRIzol using IBI scientific TRI-Isolate RNA Pure Kit, eluted into 25 µL nuclease-free water and stored at -80°C. SuperScript™ IV Reverse Transcriptase was used to generate cDNA from circRNA barcodes following the manufacturer's instructions for gene-specific priming. Total reaction volume was 5 µL including 1 µL of extracted RNA and JJS1 as the primer.

To generate amplicons, 5 rounds of PCR were run using Q5® Hot Start High-Fidelity 2X Master Mix following the manufacturer's instructions with a 63°C annealing temperature and 30s extension. A 25 µL reaction scale was used including 12.5 µL Q5® Master Mix, 1.25 µL each of JJS2 & JJS3 (10 µM stocks), 2.5 µL cDNA reaction, 7.5µL nuclease-free water, and 0.75 µL

DMSO. PCR reactions were cleaned with SeraMag Select beads using a 2.5:1 ratio of beads:reaction and eluted using 25 µL nuclease-free water.

A second round of PCR was run on the eluted amplicons using 12.5 µL Q5® Master Mix, 1.25 µL each JJS4 & JJS5 (10 µM stocks), 5 µL eluted amplicon, 5 µL nuclease-free water, 1x SYBR gold, and 2% DMSO. Reaction progress was monitored using a Bio-Rad CFX Connect Real-Time PCR Detection System and halted in the linear phase of amplification. Reactions were equally pooled based on their final RFU values and amplicons of the correct size were selected using Sera-Mag Select beads with a 0.65:1 ratio of beads:reaction. Amplicons were eluted into nuclease free water and submitted to Genewiz using their Amplicon-EZ service for sequencing.