SH2*scan*: Mapping SH2 domain-ligand binding selectivity for inhibitors and degraders

Luis M. Gonzalez Lira[1,†], Jennifer K. Wolfe-Demarco[1,†], Alexander M. Clifford[1], Tuan D. Le[1], Ghadeer M. Khasawneh[1], Medhanie Kidane[1], Michelle Nguyen[1], Julia A. Najera[1], Gabriel Pallares[1], Nicole B. Servant[1], Jean A. Bernatchez[1*]

[1]Eurofins DiscoverX, LLC, 11180 Roselle Street, Suite D, San Diego, CA 92121, United States of America
†Luis M. Gonzalez Lira and Jennifer K. Wolfe-Demarco contributed equally to this work.
[*]Author to whom correspondence should be addressed.

Abstract

Drug discovery targeting SH2 domains (key protein-protein interaction modules) has been hampered by a lack of assay systems evaluating SH2 domain-synthetic ligand binding selectivity to reduce potential off-target effects. Here, we developed SH2*scan*, a high-throughput competition binding assay platform used to quantify ligand-SH2 domain interactions, covering >80% of the target class. This platform can be used to design more selective compounds targeting SH2 domains.

Main

Protein-protein interactions (PPIs) remain a prime target for drug discovery programs, given their importance as key biochemical transducers in deregulated cell signaling pathways, as is the case in inflammatory disorders and cancer[1,2]. SH2 domains, of which there are 120 canonical members found in 110 proteins, are prototypical protein-protein interaction modules which generally recognize phosphotyrosine-containing motifs on their protein binding partners[3–5]. Despite divergent amino acid sequences among SH2 domains, the conserved structure among these PPI modules has been thought to present challenges in the development of highly selective inhibitors of targets in the class[6]. While large scale platforms evaluating SH2 domain-phosphopeptide ligand interactions have been developed[7,8], equivalent systems to study the selectivity of SH2 domain interactions with small molecule inhibitors or targeted protein degraders have yet to be developed. Such systems would be of high value to drug discovery programs targeting SH2 domains, given their potential to identify off-target liabilities for hit-to-lead programs in the early stages of drug development.

Here, we describe a high-throughput, plate-based binding assay platform called SH2*scan*, with 95 wildtype SH2 domain-containing constructs covering a total of 97 SH2 domains (>80% of the target class), as well as 7 mutant STAT SH2 domain-containing constructs (**Extended Data Table 1**). This platform utilizes a competition binding assay principle (**Fig. 1a and 1b**). Multiple binding reactions can be performed in parallel with a dose-response of a given competitor; these data can be plotted, and a dissociation constant (K_D) value can be calculated from the curve fitting for the given protein-ligand interaction.

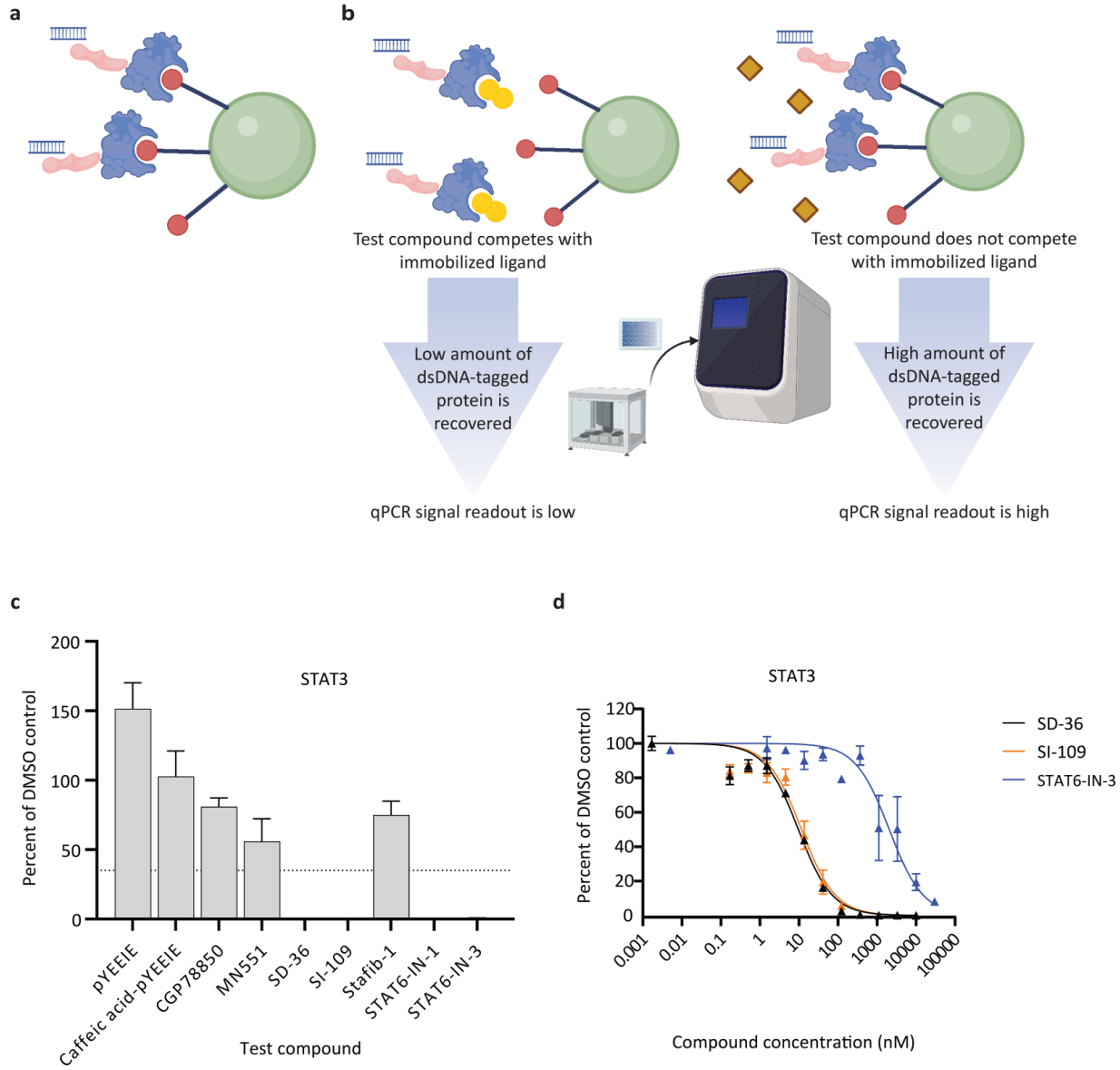


Fig.1|SH2*scan* assay principle for measuring compound dissociation constants. a, An SH2 domain-containing protein construct (blue) fused to the NFκB DNA binding domain (pink) is tagged with an exogenous double-stranded DNA (dsDNA) probe. This construct is incubated with a capture ligand (red) immobilized on magnetic beads (green). **b**, In the presence of a competitor compound (yellow), less tagged protein is captured on the beads. The protein remaining on the beads is eluted using a high concentration of sodium phenyl phosphate, a generic competitor of phosphopeptide binding to SH2 domains[9]. After elution, a lower qPCR signal is obtained at the end of the assay. In the presence of a non-competitor compound ligand (brown), more tagged protein is captured on the beads and a high qPCR signal is observed. **c**, Representative primary screening data for the STAT3 construct are shown. Compounds were tested at 10 µM and the hit cutoff for the compound screen was set at ≤35% of the average signal of the DMSO control wells for each construct (dashed line). Percent assay signal for each compound is expressed as the mean of at least two independent technical replicates from at least one independent experiment, ± standard deviation. **d**, K_D data for selected validated hits for the STAT3

construct are shown. Compounds were tested in dose-response, and the data points were fit to the Hill equation (see Methods). Data are presented as mean percent assay signal from at least four independent technical replicates collected over two independent experiment, ± standard deviation.

To examine the selectivity profiles of synthetic ligands across the SH2 domain target class, we chose 9 compounds from the literature to profile in our panel (**Extended Data Fig. 1**). These compounds were expected to have a range of primary targets based on literature sources and included inhibitors[10–15] and a protein degrader[16]. We profiled each of these compounds in a primary screen and then validated all hits in dose-response to extract K_D values for each protein-ligand interaction. Representative data for the primary screen are shown in **Fig. 1c**, while sample dose-response data are shown in **Fig. 1d**.

A phylogenetic tree was generated containing the 120 *Homo sapiens* SH2 domains, and for each compound, branch tips of the tree were assigned dots proportionally sized to the normalized percent of DMSO signal from the primary screening assay (**Fig. 2**). Hits from the complete primary screen (**Extended Data Tables 2 and 3**) were validated in dose-response to obtain K_D values (**Extended Data Tables 4 and 5**). Available literature values for these interactions are presented in **Extended Data Table 6**. Certain compound-SH2 domain target pairs which were previously described as having measured affinities in the literature[16] that did not come up as hits in the primary screen were tested in dose-response, and in those cases where we measured K_D values, these are indicated in **Extended Data Table 4**. Compounds which were above the hit cutoff in the primary screen that were validated as hits when tested in dose-response, and primary screening hits that did not confirm in the dose-response experiments, are also noted in this table.

We observed a wide range of affinities (picomolar to micromolar K_D values) and selectivity profiles for the compounds tested in SH2*scan*. This study has revealed that even small changes to the chemical structure of SH2 domain-targeting compounds can have profound effects on binding affinity and selectivity within the target class. While some compounds tested showed relatively narrow selectivity profiles (such as SD-36, SI-109 and MN551), others (such as caffeic acid-pYEEIE, STAT6-IN-1 and STAT6-IN-3) showed promiscuity in terms of their SH2 domain binding. These findings demonstrate the utility of compound screening across SH2 domains to reveal off-target binding events.

Many SH2 domain-containing proteins present in SH2*scan* have limited or non-existent biochemical and functional characterization in the scientific literature. This platform can be used to study these targets and potentially uncover new biological insights for them via phosphopeptide profiling. In sum, this work showcases SH2*scan* as an assay platform that can be leveraged at scale to both develop new therapeutics targeting SH2 domains as well as to uncover new protein-protein interaction networks.

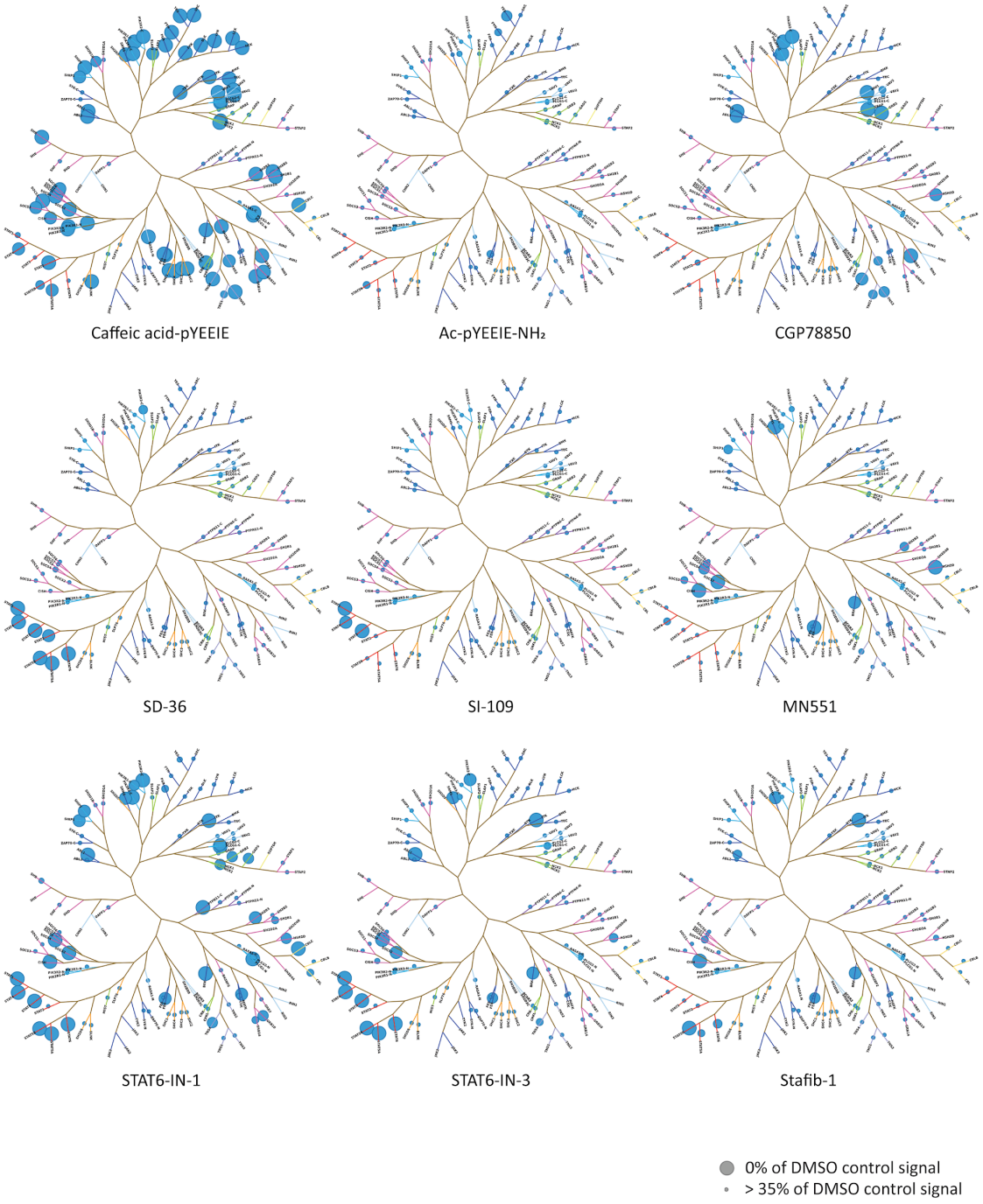


Fig. 2|Primary screening across SH2*scan* reveals unique binding selectivity signatures for synthetic ligands. Percent of DMSO control values were measured for each compound tested in this study and mapped as dots on a phylogenetic tree containing all 120 canonical human SH2 domains. Each dot plotted in a diagram represents a mean value collected from at least two independent technical replicates collected over at least one independent experiment, ± standard deviation. The screening results for the two SH2 domains of SYK and ZAP70 are mapped as equivalently sized dots, given that the SH2 domains for these targets are expressed in tandem constructs, and any competition observed cannot be unequivocally assigned to one or the other SH2 domain.

References

1. Fu, H., Mo, X. & Ivanov, A. A. Decoding the functional impact of the cancer genome through protein–protein interactions. *Nat. Rev. Cancer* **25**, 189–208 (2025).
2. Nada, H. *et al.* New insights into protein-protein interaction modulators in drug discovery and therapeutic advance. *Signal Transduct. Target. Ther.* **9**, 341 (2024).
3. Songyang, Z. *et al.* SH2 domains recognize specific phosphopeptide sequences. *Cell* **72**, 767–778 (1993).
4. Liu, B. A., Engelmann, B. W. & Nash, P. D. The language of SH2 domain interactions defines phosphotyrosine-mediated signal transduction. *FEBS Lett.* **586**, 2597–2605 (2012).
5. Diop, A. *et al.* SH2 domains: folding, binding and therapeutical approaches. *Int. J. Mol. Sci.* **23**, (2022).
6. Kraskouskaya, D., Duodu, E., Arpin, C. C. & Gunning, P. T. Progress towards the development of SH2 domain inhibitors. *Chem. Soc. Rev.* **42**, 3337–3370 (2013).
7. Jones, R. B., Gordus, A., Krall, J. A. & MacBeath, G. A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature* **439**, 168–174 (2006).
8. Machida, K. *et al.* High-throughput phosphotyrosine profiling using SH2 domains. *Mol. Cell* **26**, 899–915 (2007).
9. Surdo, P. L. *et al.* Structural and biochemical evaluation of the interaction of the phosphatidylinositol 3-kinase p85alpha Src homology 2 domains with phosphoinositides and inositol polyphosphates. *J. Biol. Chem.* **274**, 15678–15685 (1999).
10. Park, S.-H., Won, J. & Lee, K.-H. Design and characterization of non-phosphopeptide inhibitors for Src family SH2 domains. *Bioorg. Med. Chem. Lett.* **12**, 2711–2714 (2002).
11. Gilmer, T. *et al.* Peptide inhibitors of src SH3-SH2-phosphoprotein interactions. *J. Biol. Chem.* **269**, 31711–31719 (1994).
12. Furet, P. *et al.* Structure-based design and synthesis of high affinity tripeptide ligands of the Grb2-SH2 domain. *J. Med. Chem.* **41**, 3442–3449 (1998).
13. Ramachandran, S. *et al.* Structure-based design of a phosphotyrosine-masked covalent ligand targeting the E3 ligase SOCS2. *Nat. Commun.* **14**, 6345 (2023).
14. Mandal, P. K. *et al.* Targeting the Src Homology 2 (SH2) Domain of Signal Transducer and Activator of Transcription 6 (STAT6) with Cell-Permeable, Phosphatase-Stable Phosphopeptide Mimics Potently Inhibits Tyr641 Phosphorylation and Transcriptional Activity. *J. Med. Chem.* **58**, 8970–8984 (2015).
15. Elumalai, N., Berg, A., Natarajan, K., Scharow, A. & Berg, T. Nanomolar inhibitors of the transcription factor STAT5b with high selectivity over STAT5a. *Angew. Chem. Int. Ed* **54**, 4758–4763 (2015).
16. Bai, L. *et al.* A Potent and Selective Small-Molecule Degrader of STAT3 Achieves Complete Tumor Regression In Vivo. *Cancer Cell* **36**, 498-511.e17 (2019).

212 **Online Methods**
213
214 Small molecules and nucleic acids
215
216 Caffeic acid-pYEEIE, CGP78850, MN551, SD-36, SI-109, Stafib-1, STAT6-IN-1, and STAT6-IN-3
217 were purchased from MedChemExpress. The Ac-pYEEIE-NH_2 phosphopeptide was custom synthesized
218 and purchased from Biopeptide Co., Inc. The DNA probe containing the qPCR amplicon for tagging the
219 NFkB fusion domain was custom synthesized and purchased from Thermo Fisher Scientific. The
220 structures of all 9 competitor ligands used in this study are shown in **Extended Data Fig. 1**.
221
222 Protein constructs and protein expression
223
224 Wild-type and mutant SH2-domain containing protein constructs were engineered as N-terminal
225 fusions with the DNA-binding domain of NFkB (amino acids 35–36 fused to amino acids 41–359, with
226 UniProt entry P19838 being used as a reference)[17]. Expression of SH2 domain-containing proteins was
227 achieved through transient transfection of HEK293 cells. The resulting protein extracts were prepared
228 using M-PER extraction buffer (Pierce Biotechnology), with added 150 mM NaCl, 10 mM DTT,
229 cOmplete™ Protease Inhibitor Cocktail (Roche Diagnostics GmbH) and Phosphatase Inhibitor Cocktail Set
230 II (Merck KGaA) according to the manufacturer's guidelines. Phosphatase inhibitors were omitted for the
231 non-phosphorylated (np) STAT3 mutant extracts: np-STAT3(N647I), np-STAT3(S614R), np-STAT3(D661Y),
232 np-STAT3(Y640F), np-STAT3(N567K), and np-STAT3(E616del). These extracts were incubated at 30 °C for
233 45 min during the protein harvest protocol to allow for dephosphorylation of STAT3 phosphotyrosines
234 by endogenous phosphatases. All other extracts for the STAT3 mutant constructs were performed as per
235 the manufacturer's instructions. Construct details are provided in **Extended Data Table 1**.
236
237 Competition binding assays
238
239 Competition binding assays were adapted from protocols used for kinases, bromo domains and
240 KRAS[18–23]. For liganded beads, preparation was performed as follows: streptavidin-coated magnetic
241 beads (Thermo Fisher Scientific) were incubated with biotinylated phosphopeptide capture ligands
242 which were custom synthesized (Biopeptide Co., Inc) at 25 °C for 30 min. In order to remove the
243 unbound affinity ligand and to reduce the nonspecific binding of proteins in the cell lysate, excess biotin
244 (125 mM) was subsequently introduced to the ligand-coated beads. The beads were then washed with a
245 blocking buffer containing SeaBlock (Pierce Biotechnology), 1% BSA and 0.05% Tween 20.
246 The binding reactions were performed with protein extracts containing the DNA-tagged SH2
247 domain-containing construct, affinity ligand-coated beads and a given small molecule test compound in
248 a binding buffer (10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 0.01% Tween 20) in a deep well, natural
249 polypropylene 384-well plate, catalog number 784201 (Greiner Bio-One), in a final volume of 19.7 µL. No
250 purification of SH2 domain-containing constructs was conducted prior to adding the protein extracts to
251 the reaction mixture. The extracts were diluted 10,000-fold in the final reaction, resulting in a DNA-
252 tagged construct concentration of less than 0.1 nM. Binding assay reactions were incubated at 25 °C
253 with shaking for 1 h. After the incubation period, the affinity ligand-coated beads were thoroughly
254 washed with the buffer used in the binding step to remove any non-specifically bound protein from the
255 beads. The remaining bead-bound protein was then eluted from the beads via resuspension and
256 incubation (30 min, 25°C) in an elution buffer of 20 mM sodium phenyl phosphate dibasic hydrate
257 (Sigma-Aldrich) dissolved in the binding buffer. Quantitative PCR was then used to determine the
258 concentration of SH2-domain containing protein constructs in the eluates. Primary screens were
259 conducted using 10 µM of a given test compound. Hits were then validated in dose-response to extract

260 K_D values for each test compound and were determined using a series of 11-point, 3-fold serial dilutions.
261 The concentration of the capture ligand present on the magnetic beads was optimized in each assay to
262 ensure accurate determination of true thermodynamic K_D values for test compounds, as previously
263 described in detail for this type of assay[21]. The complete primary screening data set is shown in
264 **Extended Data Tables 2 and 3**, and the complete dose-response validation data set for this study is
265 presented in **Extended Data Tables 4 and 5**.
266
267 Curve fitting for competitive binding assays
268
269 Thermodynamic dissociation constants (K_Ds) were derived for each competitor-SH2 domain-
270 containing construct interaction by fitting the data to a standard dose-response curve using the Hill
271 equation:
272
273 $\text{response} = \text{background} + (\text{signal} - \text{background})/(1 + (K_D^{\text{Hill slope}}/\text{dose}^{\text{Hill slope}}))$
274
275 The Hill slope was set to -1 and a nonlinear least-square fit using the Levenberg-Marquardt
276 algorithm was performed to generate the final curves.
277
278 Generation of SH2 domain phylogenies
279
280 Phylogenetic trees were generated based on the sequence for each canonical *Homo sapiens* SH2
281 domain. Protein sequences for the SH2 domain regions were mined from NCBI reference sequences,
282 aligned using MUSCLE[24] in SeaView v.5.0.5[25] and were further refined using GBlocks[26] with parameters
283 allowing for the least amount of stringency for retention[27]. Phylogenetic analysis was conducted on the
284 Cyberinfrastructure for Phylogenetic Research (CIPRES) Science Gateway (www.phylo.org[28]) using the
285 RAxML software v.8.2.12[29] with the LG evolutionary and GTRGAMMA models[30]. Branch support was
286 estimated by bootstraps with 700 replications and the constructed phylogram was edited using ITOL
287 (itol.embl.de). SH2 protein nomenclature was determined using UNIPROT (uniprot.org). Accession
288 numbers for individual proteins are available in **Extended Data Table 1**.
289
290 Analysis of primary screening and dose-response data
291
292 Independent replicates for each primary screening and dose-response validation data set are noted in
293 **Extended Data Table 3** and **Extended Data Table 5**, respectively. For each independent target assessed
294 in both primary screening and dose-response formats, an assay signal was calculated from the vehicle
295 control (DMSO) and background from a positive control compound. Passing assay metrics were defined
296 as a signal to background ratio of ≥3 and a Z score of ≥0.3 calculated in Spotfire™ (Tibco). For primary
297 screening experiments, the percent of control (POC) for each replicate per compound was calculated by
298 dividing the assay signal of the test compound by that of the vehicle control. Positive hits were defined
299 as ≤35% POC. For dose-response experiments, K_D values were calculated for each individual replicate
300 curve.
301
302 Mapping of primary screening values onto phylogenetic trees
303
304 Phylogenetic trees were mapped in MS Paint and the resulting X,Y coordinates loaded into
305 Spotfire™(Tibco) along with the tree image described above. Primary screening data was mapped onto
306 the resulting tree template by subtracting the average POC for the test compound from 100 in order to
307 generate dots directly proportional to the strength of each hit. These values were then loaded into

Spotfire™(Tibco) along with the corresponding X,Y coordinates and used to generate a spot layer overlay onto the phylogenetic tree.

Data availability

Images in **Fig.1a** and **Fig.1b** were generated using Biorender.com and in **Fig.1c** and **Fig.1d** using GraphPad Prism. Chemical structures in **Extended Data Fig. 1** were generated using ChemDraw. Construct details, primary screening data and dose-response data for this study are presented in **Extended Data Tables 1-5**. Any other data that support the conclusions of this paper are available upon request.

References

17. Müller, C. W., Rey, F. A., Sodeoka, M., Verdine, G. L. & Harrison, S. C. Structure of the NF-kappa B p50 homodimer bound to DNA. *Nature* **373**, 311–317 (1995).
18. Fabian, M. A. *et al.* A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat. Biotechnol.* **23**, 329–336 (2005).
19. Karaman, M. W. *et al.* A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **26**, 127–132 (2008).
20. Davis, M. I. *et al.* Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **29**, 1046–1051 (2011).
21. Wodicka, L. M. *et al.* Activation state-dependent binding of small molecule kinase inhibitors: structural insights from biochemistry. *Chem. Biol.* **17**, 1241–1249 (2010).
22. Ciceri, P. *et al.* Dual kinase-bromodomain inhibitors for rationally designed polypharmacology. *Nat. Chem. Biol.* **10**, 305–312 (2014).
23. Kidane, M. *et al.* Suite of Biochemical and Cell-Based Assays for the Characterization of Kirsten Rat Sarcoma (KRAS) Inhibitors and Degraders. *ACS Pharmacol. Transl. Sci.* **7**, 3921–3934 (2024).
24. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004).
25. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224 (2010).
26. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552 (2000).
27. Talavera, G. & Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577 (2007).
28. Miller, M. A., Pfeiffer, W. & Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. in *2010 Gateway Computing Environments Workshop (GCE)* 1–8 (IEEE, 2010). doi:10.1109/GCE.2010.5676129.
29. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
30. Le, S. Q. & Gascuel, O. An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**, 1307–1320 (2008).

Acknowledgements

We thank members of Eurofins DiscoverX, LLC for reviewing the manuscript and providing feedback.

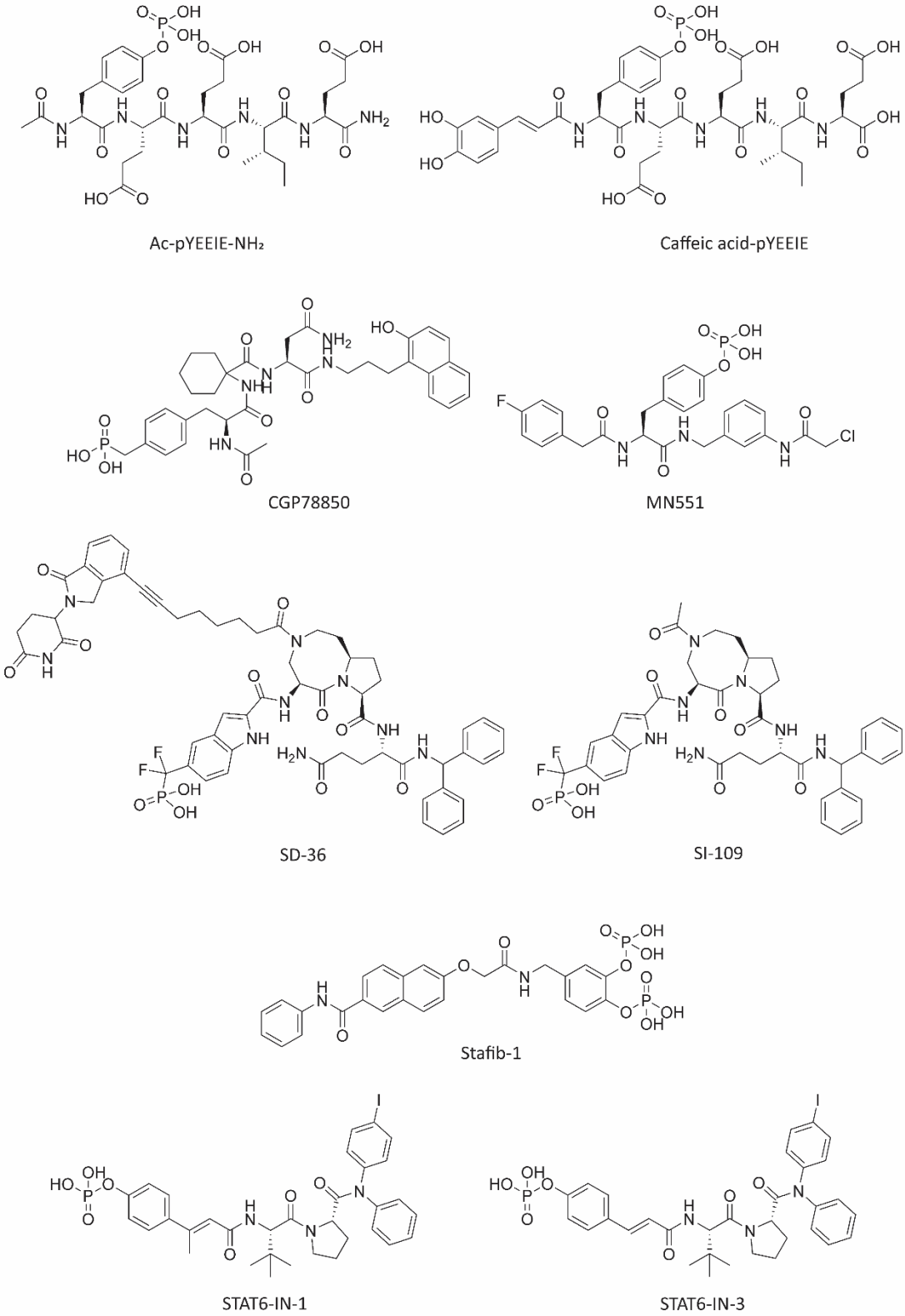
Author contributions

Design of the clones and capture ligand molecules was performed by L.M.G.L. and J.A.B. Construct cloning and protein lysate generation were performed by L.M.G.L., A.M.C., T.D.L., G.M.K., M.K., M.N., J.A.N. and J.A.B. Assay development and optimization was conducted by L.M.G.L., A.M.C., T.D.L., G.M.K., M.K., M.N., J.A.N. and J.A.B. Compound screening was performed by J.K.W.D. The phylogenetic tree used for compound selectivity graphics was generated by A.M.C. Mapping of screening data to the phylogenetic tree diagram was conducted by J.K.W.D. and G.P. Data analysis for compound screening data was performed by J.K.W.D., G.P., N.B.S. and J.A.B. Experimental conceptualization and project supervision were conducted by J.A.B. The paper was written by L.M.G.L., J.K.W.D., A.M.C., T.D.L., G.P. and J.A.B. All authors reviewed the manuscript.

Competing Interests

All authors were employees of Eurofins DiscoverX, LLC when this work was performed. J.A.B. is a co-inventor on a pending patent application filed by Eurofins DiscoverX, LLC for SH2 domain competition binding assays.

Extended Data Fig. 1|Chemical structures of the 9 synthetic SH2 domain ligands tested in SH2*scan*. Compounds tested for binding selectivity in this study are shown below.



408 **Extended Data Table 1|SH2*scan* protein construct details.** Shown are the gene names, construct
409 names, NCBI protein accession numbers, amino acid ranges and mutation status for all proteins in
410 SH2*scan*. SH2dom.C-term. indicates a construct containing the SH2 domain closer to the C-terminus of
411 the full-length protein when more than one SH2 domain is present in the protein. SH2dom.N-term.
412 indicates a construct containing the SH2 domain closer to the N-terminus of the full-length protein when
413 more than one SH2 domain is present in the protein. SH2dom1.dom2. signifies a construct containing
414 two tandem SH2 domains (these engage a doubly phosphorylated phosphopeptide capture ligand). Non-
415 phosphorylated constructs are harvested from protein lysates that are not treated with phosphatase
416 inhibitors, as described in the Methods section.
417

Gene Name	Construct Name	NCBI protein accession number	Amino acid range	Mutant
ABL1	ABL1	NP_009297.2	S140/Y251	No
ABL2	ABL2	NP_009298.1	V165/G273	No
APS (SH2B2)	APS (SH2B2)	NP_001346157.1	E409/Q518	No
BLK	BLK	NP_001706.2	V113/A224	No
BLNK	BLNK	NP_037446.1	G317/S456	No
BRK	BRK	NP_005966.1	S75/E174	No
CBL	CBL	NP_005179.2	P47/G351	No
CBLB	CBLB	NP_001308715.1	A67/F454	No
CBLC	CBLC	NP_036248.3	M1/G323	No
CISH	CISH	NP_037456.5	K76/L275	No
CRK	CRK	NP_058431.2	D6/G124	No
CRKL	CRKL	NP_005198.1	M1/Y105	No
CSK	CSK	NP_004374.1	M80/M173	No
DAPP1	DAPP1	NP_055210.2	Q31/R131	No
FER	FER	NP_005237.2	K453/K552	No
FES	FES	NP_001996.1	V451/K551	No
FGR	FGR	NP_005239.1	S138/A255	No
FRK	FRK	NP_002022.1	A106/V213	No
FYN	FYN	NP_002028.1	W149/K248	No
GADS	GADS	NP_004801.1	W58/H155	No
GRAP	GRAP	NP_006604.1	P59/E151	No
GRB10	GRB10	NP_001357938.1	I536/R640	No
GRB14	GRB14	NP_004481.2	I433/R537	No
GRB2	GRB2	NP_002077.1	P59/I151	No
GRB7	GRB7	NP_001025173.1	P415/L532	No
HCK	HCK	NP_002101.2	L139/K245	No
HSH2D	HSH2D	NP_116244.1	G24/D129	No
ITK	ITK	NP_005537.3	N232/C339	No
LCK	LCK	NP_005347.3	A119/T226	No
LNK (SH2B3)	LNK (SH2B3)	NP_005466.1	D356/V465	No
LYN	LYN	NP_002341.1	S115/P229	No

MATK	MATK	NP_002369.2	L118/G216	No
MIST	MIST	NP_443196.2	W291/T422	No
NCK1	NCK1	NP_001278928.1	P281/S377	No
NCK2	NCK2	NP_003572.2	E284/Q380	No
PIK3R1	PIK3R1(SH2dom.C-term.)	NP_852664.1	L616/R724	No
PIK3R1	PIK3R1(SH2dom.N-term.)	NP_852665.1	G21/Q135	No
PIK3R2	PIK3R2(SH2dom.C-term.)	NP_005018.2	D610/R717	No
PIK3R2	PIK3R2(SH2dom.N-term.)	NP_005018.2	A317/Q432	No
PIK3R3	PIK3R3(SH2dom.C-term.)	NP_001290357.1	N366/C477	No
PIK3R3	PIK3R3(SH2dom.N-term.)	NP_001290357.1	S74/E188	No
PLCG1	PLCG1(SH2dom.C-term.)	NP_002651.2	H663/E759	No
PLCG1	PLCG1(SH2dom.N-term.)	NP_002651.2	E533/A662	No
PLCG2	PLCG2(SH2dom.C-term.)	NP_002652.2	H641/E738	No
PLCG2	PLCG2(SH2dom.N-term.)	NP_002652.2	E516/P640	No
PTPN11	PTPN11(SH2dom.C-term.)	NP_001317366.1	T108/R220	No
PTPN11	PTPN11(SH2dom.N-term.)	NP_001317366.1	M1/P107	No
PTPN6	PTPN6(SH2dom.C-term.)	NP_536859.1	T106/R217	No
PTPN6	PTPN6(SH2dom.N-term.)	NP_536859.1	M1/D104	No
RASA1	RASA1(SH2dom.C-term.)	NP_002881.1	G340/Q444	No
RASA1	RASA1(SH2dom.N-term.)	NP_002881.1	T174/D280	No
SH2B (SH2B1)	SH2B (SH2B1)	NP_001139267.1	D519/S628	No
SH2D1A	SH2D1A	NP_002342.1	M1/K104	No
SH2D1B	SH2D1B	NP_444512.2	M1/R103	No
SH2D6	SH2D6	NP_001381392.1	S213/P335	No
SH3BP2	SH3BP2	NP_001139328.1	E503/R618	No
SHB	SHB	NP_003019.2	E397/L509	No
SHC1	SHC1	NP_001123512.1	A481/L584	No
SHC2	SHC2	NP_036567.2	E479/P582	No
SHC3	SHC3	NP_058544.3	E492/Q594	No
SHC4	SHC4	NP_976224.3	K518/K630	No
SHF	SHF	NP_001288097.2	L336/L441	No
SHIP1	SHIP1	NP_001017915.1	M1/P112	No
SHIP2	SHIP2	NP_001558.3	S20/V117	No
SLAP2	SLAP2	NP_115590.1	A88/L193	Np
SLP76	SLP76	NP_005556.1	N404/T527	No
SOCS1	SOCS1	NP_003736.1	S60/I211	No
SOCS2	SOCS2	NP_001257400.1	Q32/V198	No
SOCS3	SOCS3	NP_003946.3	M1/L225	No
SOCS4	SOCS4	NP_955453.1	L274/E437	No
SOCS6	SOCS6	NP_004223.2	V361/S499	No
SOCS7	SOCS7	NP_055413.2	S451/F576	No

SRC	SRC	NP_938033.1	Q147/T250	No
SRMS	SRMS	NP_543013.1	D117/M213	No
STAP1	STAP1	NP_036240.1	Y168/H293	No
STAP2	STAP2	NP_060190.2	V138/D247	No
STAT1	STAT1	NP_009330.1	L136/S710	No
STAT2	STAT2	NP_005410.1	G130/L706	No
STAT3	STAT3	NP_001356441.1	G127/I722	No
STAT3	STAT3(D661Y) -nonphosphorylated	NP_001356441.1	G127/I722	D661Y
STAT3	STAT3(E616del)-nonphosphorylated	NP_001356441.1	G127/I722	E616del
STAT3	STAT3(N567K)-nonphosphorylated	NP_001356441.1	G127/I722	N567K
STAT3	STAT3(N647I)-nonphosphorylated	NP_001356441.1	G127/I722	N647I
STAT3	STAT3(S614R)-nonphosphorylated	NP_001356441.1	G127/I722	S614R
STAT3	STAT3(Y640F) -nonphosphorylated	NP_001356441.1	G127/I722	Y640F
STAT4	STAT4	NP_001230764.1	S137/S702	No
STAT5A	STAT5A	NP_001275647.1	S128/D712	No
STAT5B	STAT5B	NP_036580.2	S128/D717	No
STAT5B	STAT5B(N642H)	NP_036580.2	S128/D717	N642H
STAT6	STAT6	NP_003144.3	M119/Q662	No
SYK	SYK(SH2dom1.dom2.)	NP_003168.2	S9/I262	No
TEC	TEC	NP_003206.2	T235/T355	No
TNS1	TNS1	NP_072174.3	D1458/D1574	No
TNS2	TNS2	NP_056134.2	D1145/D1259	No
TNS3	TNS3	NP_073585.8	D1167/D1284	No
TNS4	TNS4	NP_116254.4	D444/E558	No
TXK	TXK	NP_003319.2	N140/S251	No
VAV1	VAV1	NP_005419.2	H660/T772	No
VAV2	VAV2	NP_001127870.1	S663/S773	No
VAV3	VAV3	NP_006104.4	V661/S772	No
YES	YES	NP_005424.1	P149/T262	No
ZAP70	ZAP70(SH2dom1.dom2.)	NP_001070.2	D3/N256	No

Extended Data Table 2|Primary screening data for the 9 compounds tested in SH2*scan*. Data is expressed as the mean of the normalized percent of the DMSO vehicle control, ± standard deviation. The number of independent replicates and independent experiments for the shown data are presented in **Extended Data Table 3**. False positives are marked with an asterisk.

Target	Caffeic Acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH_2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
ABL1	0.0±1.4	3.1±1.7	35.7±4.3	106.1±10.2	87.4±24.0	131.9±19.8	19.9±26.7*	0±0*	0±46.8*
ABL2	0.7±0.4	6.2±0.1	94.6±1.6	110.7±17.0	125.7±12.9	122.1±5.7	106.0±17.7	98.1±4.2	108.3±3.5
APS (SH2B2)	2.2±0.3	128.1±7.6	94.2±3.2	121.8±20.8	140.1±6.1	130.4±1.6	64.6±31.3	45.3±5.2	66.4±17.9
BLK	0.1±0.0	42.1±4.6	115.5±14.2	47.8±12.5	70.6±5.9	105.3±14.7	91.8±35.5	49.1±3.4	67.6±11.0
BLNK	3.2±0.6	116.9±7.1	100.6±0.3	118.9±8.3	103.8±13.0	117.0±0.7	157.9±27.8	116.4±19.9	145.1±22.5
BRK	0.0±0.0	116.9±12.5	4.7±0.7*	101.4±15.6	76.9±3.5	124.3±16.8	5.4±60.2*	3.0±0.4*	4.1±47.5*
CBL	32.7±4.1	76.7±17.1	38.9±1.8	121.6±16.4	68.3±19.5	100.8±12.1	85.6±14.2	24.3±1.0	73.9±11.4
CBLB	32.7±22.6	122.1±4.2	86.8±3.2	116.1±4.8	60.9±5.8	128.3±5.4	78.9±15.9	36.0±0.8	69.0±7.6
CBLC	0.1±0.0	128.3±29.6	43.0±3.8	97.4±7.1	113.0±4.5	106.0±14.1	100.1±12.3	3.3±0.1	34.7±1.3
CISH	4.9±6.9	116.9±15.1	3.2±4.3	131.9±3.0	42.8±5.7	101.6±5.9	1.7±2.3	1.6±2.3	26.8±11.8
CRK	3.1±2.8	88.3±0.9	72.1±0.5	114.0±4.2	77.0±13.9	118.7±21.4	90.2±5.9	34.5±1.8	80.4±9.5
CRKL	42.1±6.0	36.4±18.5	202.2±6.2	131.3±16.6	119.4±31.5	125.3±32.5	120.1±104.6	42.4±0.9	102.1±19.9
CSK	103.0±4.8	119.6±4.5	105.0±7.9	106.4±17.0	97.3±14.8	121.2±1.9	111.7±5.5	26.3±0.8	65.2±5.7
DAPP1	39.2±3.2	136.1±2.8	81.2±19.3	134.5±1.4	125.9±18.7	148.7±2.8	100.2±22.0	41.1±13.9	58.9±43.2
FER	0.8±0.8	135.5±7.5	1.1±0.2*	37.4±14.9	74.4±18.6	107.2±24.0	10.8±42.1*	8.5±0.1	10.7±34.0
FES	2.2±0.4	52.0±0.5	82.1±4.3	24.1±2.4	46.3±7.5	159.5±1.3	115.8±98.4	66.3±6.4	82.0±13.1
FGR	0.0±0.0	43.3±1.3	49.8±22.3	49.9±0.9	82.0±28.0	127.6±6.4	34.3±37.0*	20.7±23.8	32.1±40.1
FRK	0.1±0.1	89.4±19.2	69.0±8.3	82.5±12.4	155.2±4.5	150.8±10.9	65.3±42.8	38.3±5.9	59.9±42.5
FYN	0.0±0.1	53.7±4.6	83.2±0.4	21.2±2.5	106.0±20.7	97.3±16.0	92.1±35.5	48.2±2.1	51.8±33.2
GADS	94.8±0.0	42.0±2.5	42.9±0.8	115.5±27.2	140.4±0.2	276.5±49.4	51.9±12.6	9.6±0.1	76.5±7.9
GRAP	109.9±0.3	0.0±0.0	48.7±17.4	112.9±3.8	89.2±4.6	140.7±11.0	76.2±27.0	52.0±15.9	73.5±23.6
GRB10	0.7±0.9	88.5±13.4	61.8±16.1	155.2±118.9	111.2±9.2	121.7±4.4	78.1±12.9	14.6±11.0*	58.8±17.6
GRB14	0.0±0.0	45.5±1.4	62.1±2.4	56.1±7.2	43.2±2.9	110.5±5.2	55.1±11.6	24.5±4.0	79.3±7.7
GRB2	73.4±9.2	0.0±0.0	103.7±0.2	77.1±23.9	92.9±1.6	104.9±5.4	79.5±8.8	14.7±3.1*	72.4±21.8
GRB7	0.2±0.1	0.0±0.0	71.3±3.5	123.8±0.6	66.9±4.6	134.7±14.0	57.9±22.2	47.7±4.2	81.2±4.6
HCK	0.0±0.3	60.0±12.1	92.7±11.3	82.2±11.9	56.4±3.2	91.9±27.0	77.0±19.3	48.8±5.2	77.4±4.3
HSH2D	80.3±18.6	8.1±6.6	3.0±0.2	127.4±18.6	72.7±14.5	154.5±38.0	27.1±53.9*	14.6±0.1*	19.6±29.7

Target	Caffeic Acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH_2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
ITK	3.1±0.3	113.6±8.8	51.4±1.1	122.5±7.6	94.0±11.1	124.2±3.0	0±61.0*	0.0±0.0	0±54.7*
LCK	0.0±0.0	46.0±3.7	98.7±0.5	44.2±3.3	97.3±3.9	126.8±18.2	97.4±7.6	48.3±1.3	72.0±20.4
LNK (SH2B3)	0.4±0.1	79.1±5.6	17.7±5.5	72.0±12.2	86.6±0.0	81.4±6.7	94.8±14.4	6.9±3.8	62.0±5.7
LYN	0.0±0.0	40.0±0.0	161.3±102.3	38.1±2.4	90.9±2.3	140.0±1.1	95.6±22.4	73.6±54.0	82.3±17.1
MATK	83.8±3.5	153.2±7.1	60.4±2.8	197.1±10.7	158.6±5.1	163.4±8.7	59.9±60.2	6.5±1.8	70.3±45.8
MIST	72.9±1.4	101.9±2.3	102.3±1.8	55.1±77.7	103.5±8.6	103.7±3.5	96.8±9.3	88.7±5.9	101.6±8.1
NCK1	0.8±0.1	104.4±4.8	67.4±5.3	90.1±14.4	106.4±10.8	124.1±9.7	53.7±39.3	10.6±2.6*	85.1±10.2
NCK2	2.8±3.1	111.3±23.7	81.5±3.0	74.9±14.5	80.0±38.4	121.1±35.8	91.4±35.1	66.2±8.2	101.7±68.4
PIK3R1(SH2dom.C-term.)	0.2±0.1	0.0±0.0	52.4±0.0	35.8±9.3	94.6±6.5	127.8±12.3	85.3±16.8	12.3±0.1	32.6±4.4
PIK3R1(SH2dom.N-term.)	0.7±0.1	109.2±5.7	99.8±1.6	87.7±20.5	123.7±16.3	133.8±18.4	88.8±24.3	44.8±5.0	85.6±12.9
PIK3R2(SH2dom.C-term.)	0.3±0.9	0.4±0.4*	30.3±1.1*	49.5±7.1	18.5±3.9*	120.4±15.9	84.2±34.7	0.2±0.1	6.1±1.6
PIK3R2(SH2dom.N-term.)	0.0±0.2	141.1±24.2	101.2±13.4	74.1±10.0	132.1±39.4	130.0±27.0	97.2±21.5	48.2±5.7	88.4±18.2
PIK3R3(SH2dom.C-term.)	0.8±0.1	0.0±0.0	27.3±3.8	27.7±0.9	57.8±6.8	137.4±20.4	86.5±0.0	1.0±0.1	26.9±10.2
PIK3R3(SH2dom.N-term.)	0.6±0.0	111.3±10.3	97.7±10.4	85.7±19.1	110.9±7.3	142.0±21.5	105.4±23.9	48.3±1.7	85.8±16.7
PLCG1(SH2dom.C-term.)	0.0±0.0	96.8±4.1	64.1±4.5	37.9±3.6	85.7±0.8	98.0±1.4	78.8±17.8	1.5±2.1	24.7±7.0
PLCG1(SH2dom.N-term.)	0.3±0.0	108.7±14.6	62.8±3.8	42.1±30.1	91.3±4.5	130.4±16.7	82.6±20.2	0.3±0.1	32.7±14.4
PLCG2(SH2dom.C-term.)	1.1±0.1	104.4±2.7	133.0±12.6	90.9±15.5	86.3±12.4	98.7±6.9	80.4±8.2	7.5±0.6	75.3±10.0
PLCG2(SH2dom.N-term.)	1.2±0.1	87.2±19.7	109.0±67.6	92.3±3.3	73.8±3.7	111.7±2.4	72.6±27.2	20.7±29.3	51.6±17.6
PTPN11(SH2dom.C-term.)	72.6±1.1	127.5±18.4	125.6±1.4	133.6±0.6	104.4±9.8	136.9±12.1	99.1±12.9	2.1±0.9	37.9±11.0
PTPN11(SH2dom.N-term.)	92.8±10.0	125.2±9.1	85.0±0.9	116.6±1.5	89.9±0.7	109.2±0.7	79.2±4.6	76.7±1.3	73.9±34.6
PTPN6(SH2dom.C-term.)	114.5±11.0	130.9±0.8	146.7±13.0	110.0±19.0	108.2±6.9	119.0±0.4	178.8±34.7	45.8±7.8	136.2±30.6
PTPN6(SH2dom.N-term.)	117.7±19.5	111.7±9.8	101.7±14.6	163.9±13.1	87.3±4.7	123.4±22.8	27.8±4.2	69.7±0.0	77.4±15.8
RASA1(SH2dom.C-term.)	78.0±6.4	131.2±42.7	107.1±2.8	119.3±0.8	86.6±9.6	118.3±10.2	97.0±11.7	76.6±7.8	123.4±14.3
RASA1(SH2dom.N-term.)	0.1±0.0	47.6±2.5	94.2±3.0	121.2±7.1	64.7±0.3	122.7±19.0	71.3±8.8	50.2±0.4	81.7±8.5
SH2B (SH2B1)	2.3±0.4	125.1±1.9	104.8±1.6	126.6±2.3	68.4±0.4	148.7±6.6	92.7±14.3	22.6±6.8*	78.1±4.8
SH2D1A	79.8±21.3	63.8±17.1	77.9±2.2	93.4±38.3	353.9±24.9	87.6±23.9	87.8±37.4	68.7±3.5	68.5±43.7
SH2D1B	12.9±2.3	77.3±6.6	63.2±6.6	47.1±1.3	99.6±1.3	223.4±5.4	62.9±9.5	49.6±1.2	46.0±22.7
SH2D6	36.4±4.5	89.5±8.4	93.6±4.8	100.6±11.5	99.6±3.4	96.9±1.7	88.5±5.5	67.5±5.3	97.5±15.9
SH3BP2	0.9±0.3	77.8±0.5	109.2±8.6	90.0±2.1	91.5±19.4	131.0±18.9	86.0±3.2	56.5±0.3	92.8±2.7
SHB	1.8±0.1	97.4±9.4	102.1±7.8	117.6±2.8	87.7±2.3	129.2±11.5	114.1±26.3	28.7±0.4	53.4±23.5
SHC1	0.0±0.0	49.8±6.2	112.5±5.9	105.0±6.6	62.0±8.8	112.5±12.4	67.3±22.1	48.2±4.5	62.5±13.9

Target	Caffeic acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH_2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
SHC2	0.3±0.1	100.1±8.5	82.9±1.7	138.0±4.7	94.9±2.3	131.3±12.5	83.4±24.4	47.3±3.3	85.2±11.3
SHC3	0.0±0.0	29.2±2.8	91.2±1.9	98.9±6.8	74.5±0.8	122.5±7.3	87.0±14.4	35.3±4.8	50.6±6.4
SHC4	0.0±0.0	108.5±7.4	93.0±4.4	119.5±2.1	100.3±4.5	135.5±12.2	103.0±14.0	77.1±7.5	90.2±12.0
SHF	50.6±4.3	77.1±2.6	102.9±9.4	60.2±17.5	100.0±5.2	78.9±3.0	79.7±9.4	43.6±7.2	89.5±4.6
SHIP1	9.0±2.9	77.7±20.9	19.1±0.9*	79.4±1.4	44.1±5.9	101.8±16.1	61.1±7.3	4.4±1.3	53.7±3.2
SHIP2	0.2±0.0	106.5±9.3	31.7±18.3*	62.5±0.8	67.3±2.6	121.0±4.5	56.2±25.7	6.9±3.9	36.5±29.6
SLAP2	5.3±3.8	91.4±15.3	81.6±2.1	202.8±149.8	95.7±9.9	121.1±25.2	90.4±6.1	55.8±2.8	92.1±16.6
SLP76	99.6±4.7	44.1±5.3	115.6±1.3	42.0±2.1	66.5±0.8	61.9±11.9	97.5±31.0	50.8±0.4	93.6±30.3
SOCS1	1.1±0.1	129.4±11.7	38.1±7.5	128.1±5.7	132.1±8.8	103.3±30.8	156.2±34.8	98.5±3.5	95.9±35.6
SOCS2	2.6±0.1	118.8±4.7	0.0±0.0	104.8±3.6	98.5±5.4	141.7±16.9	47.9±29.0	0.0±0.0	7.9±50.6
SOCS3	5.3±0.3	115.6±4.4	24.4±16.1*	149.7±16.3	100.4±1.2	153.3±21.3	72.7±29.0	39.3±22.1	52.6±50.6
SOCS4	2.5±0.1	100.6±1.6	9.6±3.9	32.2±11.1	30.0±17.2	122.6±6.9	25.5±12.4	0.0±0.0	0.0±0.0
SOCS6	8.2±0.1	63.3±5.3	65.2±0.1	95.6±6.2	57.2±3.4	91.9±7.4	96.9±12.0	44.1±2.1*	90.2±12.1
SOCS7	0.5±0.4	126.4±4.2	21.9±5.0	110.9±12.0	270.7±32.0	124.3±32.9	81.0±17.8	34.6±0.1	50.9±16.4
SRC	0.0±0.0	99.2±4.1	70.2±6.8	68.6±4.6	81.4±1.5	116.7±8.5	90.6±5.2	53.3±8.1	67.0±21.0
SRMS	0.1±0.0	65.3±5.9	1.3±0.6*	62.7±1.1	102.2±1.9	123.4±4.6	8.9±56.2*	7.6±0.5*	8.5±53.2*
STAP1	77.9±1.0	113.1±2.6	84.3±2.8	101.0±11.6	79.9±5.8	103.5±11.2	92.0±6.8	65.6±5.1	76.9±9.2
STAP2	58.7±7.4	86.0±7.4	143.3±67.6	80.1±5.7	36.1±5.4	129.5±10.5	78.7±36.1	58.4±51.4	56.9±15.1
STAT1	27.1±21.3	71.2±2.1	49.8±4.1	69.3±10.9	9.1±0.0	0.0±1.1	48.2±5.4	0.0±0.0	0.7±0.6
STAT2	0±0.0*	27.3±0.0*	216.9±22.6	53.6±1.3	0.1±1.3	100.4±0.7	152.6±67.8	35.1±16.6	60.3±12.5
STAT3	102.5±22.8	80.9±6.4	67.8±17.7	151.4±18.8	0.1±0.1	0.4±0.1	66.9±10.0	0.0±0.0	0.7±0.2
STAT3(D661Y)-nonphosphorylated	108.2±9.3	97.4±6.3	171.4±3.1	91.4±3.8	0.0±0.1	0.6±0.2	157.3±49.3	0.0±0.0	4.4±3.7
STAT3(E616del)-nonphosphorylated	85.5±0.3	71.8±6.4	203.2±18.0	86.7±0.4	0.0±0.0	0.1±0.0	135.3±32.6	0.0±0.0	0.4±0.4
STAT3(N567K)-nonphosphorylated	91.0±2.3	78.3±2.0	86.5±6.2	94.9±0.8	0.1±0.0	0.2±0.0	97.5±16.8	0.0±0.0	0.2±0.3
STAT3(N647I)-nonphosphorylated	85.3±2.1	80.6±2.9	81.2±11.7	97.9±8.6	0.0±0.0	0.5±0.1	78.6±15.4	0.0±0.0	2.1±3.7
STAT3(S614R)-nonphosphorylated	81.4±6.7	82.4±3.2	95.7±8.1	96.3±8.9	1.6±0.0	0.1±0.0	87.4±4.6	0.0±0.0	0.1±0.1
STAT3(Y640F)-nonphosphorylated	130.6±28.4	117.6±3.3	49.0±5.7	156.0±2.6	0.2±0.6	5.8±3.7	75.3±26.0	0.0±0.0	2.4±6.1
STAT4	8.8±0.4	91.4±14.0	22±0.8*	106.6±3.0	1.1±0.2	7.3±0.1	57.6±15.3	0.0±0.0	0.2±0.3
STAT5A	6.2±0.4	84.9±21.0	39.0±0.8	94.4±4.9	3.4±0.4	100.8±9.7	22.0±26.6	0.0±0.0	0.0±0.0
STAT5B	19.3±2.8	89.6±6.1	49.1±7.6	21.8±4.7	5.7±0.1	109.2±6.0	0.0±0.0	0.0±0.0	0.0±0.0
STAT5B(N642H)	9.4±1.0	107.1±9.8	76.9±22.3	115.5±14.7	3.2±0.8	122.8±9.5	12.5±13.5	0.0±8.8	1.3±7.7
STAT6	102.6±16.9	103.6±11.0	37.2±13.5	162.5±10.0	4.2±0.1	44.5±2.9	18.2±25.5*	0.0±9.7	0.0±0.0

Target	Caffeic acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
SYK(SH2dom1.dom2.)	72.6±6.4	85.1±0.1	61.9±4.0	126.1±1.2	104.6±11.7	121.8±4.5	66.1±28.1	61.2±11.0	65.2±43.9
TEC	0.1±0.0	126.5±0.2	120.4±2.4	119.3±6.6	121.1±3.0	135.0±2.6	159.2±17.1	75.1±3.7	135.7±22.9
TNS1	25.5±1.1	16.4±0.3	133.6±23.5	161.1±11.7	112.9±1.0	152.5±3.2	131.7±25.9	87.6±10.0	88.4±24.5
TNS2	4.8±1.3	16.0±0.0	74.1±2.1	104.0±5.1	65.4±3.3	126.0±5.7	86.0±8.8	55.9±0.4	86.8±7.1
TNS3	1.8±0.6	9.8±0.1	48.5±7.3	83.2±4.0	57.6±6.0	127.9±17.2	77.8±9.0	57.0±10.1	55.6±12.1
TNS4	3.1±0.7	5.0±1.1	25.9±0.1*	61.9±0.0	58.1±8.9	109.6±29.7	32.7±33.4*	14.0±3.4	21.6±25.3*
TXK	2.7±0.3	74.4±2.5	89.1±44.1	109.9±5.6	100.7±9.7	118.1±15.2	143.5±46.3	87.1±5.9	85.4±12.9
VAV1	1.6±0.6	1.8±0.4	108.9±7.6	87.2±8.1	116.5±3.0	112.0±1.7	106.1±16.9	55.3±2.6	67.6±32.1
VAV2	0.5±0.1	50.4±7.8	77.8±4.8	49.5±2.8	116.7±1.5	119.4±0.5	95.8±29.1	79.9±1.6	87.5±11.3
VAV3	0.4±0.1	21.7±3.7	113.3±17.1	75.8±17.0	82.1±3.8	130.2±0.9	101.7±15.4	40.3±1.3	76.8±21.6
YES	0.0±0.0	66.7±8.3	46.5±2.1	14.2±1.5	73.3±11.9	112.7±26.4	63.1±24.3	29.5±0.0	37.7±27.8
ZAP70(SH2dom1.dom2.)	32.3±1.6	91.6±5.3	33.5±1.9*	134.0±4.4	95.6±1.3	129.0±1.2	58.0±19.5	51.4±0.0	60.7±32.9

463 **Extended Data Table 3| Replicate numbers for primary screening data for the 9 compounds tested in**
464 **SH2*scan*.** Independent replicates numbers (N) for each primary screening value presented in **Extended**
465 **Data Table 2** are shown below and were generated from at least one independent experiment.
466
467

Target	Caffeic Acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH_2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
ABL1	4	4	4	4	4	4	8	4	8
ABL2	2	2	2	2	2	2	4	2	4
APS (SH2B2)	2	2	2	2	2	2	4	2	4
BLK	4	4	2	4	4	4	6	2	6
BLNK	2	2	2	2	2	2	4	2	4
BRK	2	2	2	2	2	2	4	2	4
CBL	4	4	2	4	4	4	6	2	6
CBLB	2	2	2	2	2	2	4	2	4
CBLC	2	2	2	2	2	2	4	2	4
CISH	2	2	2	2	2	2	4	2	4
CRK	2	2	2	2	2	2	4	2	4
CRKL	4	4	2	4	4	4	6	2	6
CSK	2	2	2	2	2	2	4	2	4
DAPP1	2	2	2	2	2	2	4	2	4
FER	4	4	2	4	4	4	6	2	6
FES	2	2	4	2	2	2	4	4	4
FGR	2	2	2	2	2	2	6	2	6
FRK	2	2	2	2	2	2	4	2	4
FYN	4	4	2	4	4	4	6	2	6
GADS	2	2	2	2	2	2	4	2	4
GRAP	2	2	2	2	2	2	4	2	4
GRB10	2	2	2	2	2	2	4	2	4
GRB14	2	2	2	2	2	2	4	2	4
GRB2	2	2	2	2	2	2	4	2	4
GRB7	2	2	2	2	2	2	4	2	4
HCK	4	4	2	4	4	4	6	2	6
HSH2D	4	4	2	4	4	4	6	2	6
ITK	2	2	2	2	2	2	4	2	4
LCK	2	2	2	2	2	2	4	2	4
LNK (SH2B3)	2	2	2	2	2	2	4	2	4
LYN	2	2	2	2	2	2	4	2	4
MATK	2	2	2	2	2	2	4	2	4
MIST	2	2	2	2	2	2	4	2	4
NCK1	2	2	2	2	2	2	4	2	4
NCK2	4	4	2	4	4	4	6	2	6
PIK3R1(SH2dom.C-term.)	2	2	2	2	2	2	4	2	4

Target	Caffeic Acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH_2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
PIK3R1(SH2dom.N-term.)	2	2	2	2	2	2	4	2	4
PIK3R2(SH2dom.C-term.)	4	4	2	4	4	4	6	2	6
PIK3R2(SH2dom.N-term.)	4	4	2	4	4	4	6	2	6
PIK3R3(SH2dom.C-term.)	2	2	2	2	2	2	4	2	4
PIK3R3(SH2dom.N-term.)	2	2	2	2	2	2	4	2	4
PLCG1(SH2dom.C-term.)	2	2	2	2	2	2	4	2	4
PLCG1(SH2dom.N-term.)	2	2	2	2	2	2	4	2	4
PLCG2(SH2dom.C-term.)	2	2	2	2	2	2	4	2	4
PLCG2(SH2dom.N-term.)	2	2	2	2	2	2	4	2	4
PTPN11(SH2dom.C-term.)	2	2	2	2	2	2	4	2	4
PTPN11(SH2dom.N-term.)	2	2	2	2	2	2	4	2	4
PTPN6(SH2dom.C-term.)	2	2	2	2	2	2	4	2	4
PTPN6(SH2dom.N-term.)	2	2	2	2	2	2	4	2	4
RASA1(SH2dom.C-term.)	2	2	2	2	2	2	4	2	4
RASA1(SH2dom.N-term.)	2	2	2	2	2	2	4	2	4
SH2B (SH2B1)	2	2	2	2	2	2	4	2	4
SH2D1A	4	4	2	4	4	4	6	2	6
SH2D1B	2	2	2	2	2	2	4	2	4
SH2D6	2	2	2	2	2	2	4	2	4
SH3BP2	2	2	2	2	2	2	4	2	4
SHB	2	2	2	2	2	2	4	2	4
SHC1	2	2	2	2	2	2	4	2	4
SHC2	2	2	2	2	2	2	4	2	4
SHC3	2	2	2	2	2	2	4	2	4
SHC4	2	2	2	2	2	2	4	2	4
SHF	2	2	2	2	2	2	4	2	4
SHIP1	2	2	2	2	2	2	4	2	4
SHIP2	2	2	2	2	2	2	6	2	6
SLAP2	2	2	2	2	2	2	4	2	4
SLP76	2	2	2	2	2	2	4	2	4
SOCS1	2	2	2	2	2	2	4	2	4
SOCS2	2	2	2	2	2	2	4	2	4
SOCS3	2	2	4	2	2	2	6	4	6
SOCS4	2	2	2	2	2	2	4	2	4
SOCS6	2	2	2	2	2	2	4	2	4
SOCS7	2	2	2	2	2	2	4	2	4
SRC	2	2	2	2	2	2	4	2	4
SRMS	2	2	2	2	2	2	4	2	4
STAP1	2	2	2	2	2	2	4	2	4
STAP2	2	2	2	2	2	2	4	2	4
STAT1	2	2	2	2	2	2	4	2	4

Target	Caffeic acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH_2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
STAT2	2	2	2	2	2	2	4	2	4
STAT3	3	3	4	3	3	3	4	4	4
STAT3(D661Y)-nonphosphorylated	2	2	2	2	2	2	4	2	4
STAT3(E616del)-nonphosphorylated	2	2	2	2	2	2	4	2	4
STAT3(N567K)-nonphosphorylated	2	2	2	2	2	2	4	2	4
STAT3(N647I)-nonphosphorylated	2	2	2	2	2	2	4	2	4
STAT3(S614R)-nonphosphorylated	2	2	2	2	2	2	4	2	4
STAT3(Y640F)-nonphosphorylated	2	2	2	2	2	2	4	2	4
STAT4	2	2	2	2	2	2	4	2	4
STAT5A	2	2	2	2	2	2	4	2	4
STAT5B	2	2	2	2	2	2	4	2	4
STAT5B(N642H)	2	2	6	2	2	2	8	6	8
STAT6	2	2	4	2	2	2	4	4	4
SYK(SH2dom1.dom2.)	2	2	2	2	2	2	4	2	4
TEC	2	2	2	2	2	2	4	2	4
TNS1	2	2	2	2	2	2	4	2	4
TNS2	2	2	2	2	2	2	4	2	4
TNS3	2	2	2	2	2	2	4	2	4
TNS4	2	2	2	2	2	2	4	2	4
TXK	2	2	2	2	2	2	4	2	4
VAV1	2	2	2	2	2	2	4	2	4
VAV2	2	2	2	2	2	2	4	2	4
VAV3	2	2	2	2	2	2	4	2	4
YES	2	2	2	2	2	2	4	2	4
ZAP70(SH2dom1.dom2.)	2	2	2	2	2	2	4	2	4

Extended Data Table 4|Dissociation constant values for the 9 compounds tested in SH2*scan*. Mean K_D values collected in this study are listed below in nM, ± standard deviation. The number of independent replicates and independent experiments for the shown data are presented in **Extended Data Table 5**. False positives are colored in orange and are denoted with "no hit". False negatives are colored in green. The false positive rate for the study was calculated as 3.8% (35 false positives/918 total tested conditions x 100). The false negative rate for the study was calculated as 0.9% (9 false negatives/918 total tested conditions x 100).

Target	Caffeic acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH_2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
ABL1	105±17	560±360					No Hit	No Hit	No Hit
ABL2	1400±360	1100±120							
APS (SH2B2)	2300±450								
BLK	1.5±0.69								
BLNK	240±24								
BRK	24±3.9						No Hit	No Hit	No Hit
CBL	11000±2200							8100±1300	
CBLB	5700±1400								
CBLC	37±23							1800±490	2800±520
CISH	430±980		46±1.4				3900±8500	3600±640	4200±1100
CRK	2300±670							2004±5.9	
CRKL									
CSK								3050±490	
DAPP1									
FER	62±86						1900±0	1400±71	8000±8700
FES	24±4.9			4200±37					
FGR	2.7±5.3						No Hit	2300±550	No Hit
FRK	2.2±0.47								
FYN	0.53±0.13			3300±900					
GADS								2400±0	
GRAP		3.3±0.46							
GRB10	480±79							No Hit	
GRB14	1200±920							3500±420	
GRB2		3.4±2.7						6040±1400	
GRB7	6100±320	5.2±0.59							
HCK	6.0±0.97								
HSH2D		1600±350	8700±130				No Hit	No Hit	7600±1600
ITK	570±630						No Hit	10300±4900	
LCK	3.7±0.62								
LNK (SH2B3)	500±410		3900±710					2300±140	
LYN	1.0±0.34								

Target	Caffeic acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH_2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
MATK								3000±490	
MIST									
NCK1	97±71							No Hit	
NCK2	210±59								
PIK3R1 (SH2dom.C-term.)	140±100	150±28						3400±1100	3000±730
PIK3R1 (SH2dom.N-term.)	140±17								
PIK3R2 (SH2dom.C-term.)	No Hit	No Hit	No Hit		No Hit			1000±78	490±80
PIK3R2(SH2dom.N-term.)	180±52								
PIK3R3 (SH2dom.C-term.)	230±180	48±20	4400±1600	14900±28000				780±42	1020±260
PIK3R3 (SH2dom.N-term.)	150±20								
PLCG1 (SH2dom.C-term.)	14±11							2050±210	3800±280
PLCG1 (SH2dom.N-term.)	180±200							890±720	3700±330
PLCG2 (SH2dom.C-term.)	660±700							3800±500	
PLCG2 (SH2dom.N-term.)	1070±320							550±14	
PTPN11 (SH2dom.C-term.)								1060±350	
PTPN11 (SH2dom.N-term.)									
PTPN6 (SH2dom.C-term.)									
PTPN6 (SH2dom.N-term.)							13000±6800		
RASA1 (SH2dom.C-term.)									
RASA1 (SH2dom.N-term.)	230±410								
SH2B (SH2B1)	910±310								
SH2D1A									
SH2D1B	65±14								
SH2D6									
SH3BP2	1020±600								
SHB	1200±530							No Hit	
SHC1	320±280								
SHC2	580±330								
SHC3	130±120	1800±240						3600±570	
SHC4	39±6.1								
SHF									
SHIP1	5800±1600		No Hit					4400±420	
SHIP2	230±120		No Hit					2030±420	
SLAP2	310±100								
SLP76									
SOCS1	30±6.5								

Target	Caffeic acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH_2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
SOCS2	1300±880		400±35					1500±870	1010±13
SOCS3	1800±2000		No Hit						No Hit
SOCS4	1500±810		64±28		5900±920		3900±870	13±0.71	10±5.3
SOCS6	1900±240								
SOCS7	280±140		8100±2500					No Hit	
SRC	4.8±1.3								
SRMS	1.2±0.73		No Hit				No Hit	No Hit	No Hit
STAP1									
STAP2									
STAT1	No Hit				180±61	180±35	37000±8600	330±73	670±330
STAT2	3.6±0.18	No Hit			450±510	1300±1300	9300±4300		510±83
STAT3					21±24	15±9.0		180±0.01	370±162
STAT3(D661Y)-nonphosphorylated					24±12	21±6.8		1000±800	1900±2300
STAT3(E616del)-nonphosphorylated					28±11	8.7±3.4		140±35	280±120
STAT3(N567K)-nonphosphorylated					30±11	12±1.8		190±107	500±200
STAT3(N647I)-nonphosphorylated					63±23	25±8.3		800±120	1500±750
STAT3(S614R)-nonphosphorylated					9.7±4.1	5.0±0.56		88±31	140±71
STAT3(Y640F)-nonphosphorylated					80±27	78±44		102±44	95±40
STAT4	4400±370		No Hit		120±110	240±106		220±67	520±270
STAT5A	4020±1500				1600±670	7030±2200	28000±14000	9.1±4.3	46±19
STAT5B	4200±2800				2030±820	18000±4000	1800±560	16±3.2	69±42
STAT5B(N642H)	2400±2300				2400±740	19000±8500	3500±1700	5.6±2.1	36±27
STAT6					1070±740	1400±0.15	7300±2300	0.68±0.50	8.6±4.5
SYK (SH2dom1.dom2.)									
TEC	39±7.8								
TNS1	3800±610	408±160							
TNS2	4300±2500	1200±110							
TNS3	1000±210	620±51							
TNS4	2600±740	508±120	No Hit				No Hit	1800±420	No Hit
TXK	950±250								
VAV1	840±530	260±4.7							
VAV2	58±40								
VAV3	70±12	860±170							
YES	1.0±0.47			3800±780				9200±380	
ZAP70 (SH2dom1.dom2.)	6900±3.8								

496
497

Extended Data Table 5 | Replicate numbers for dissociation constant values for the 9 compounds tested in SH2*scan*. Independent replicate numbers (N) for each dose-response value presented in **Extended Data Table 4** are shown below and were generated from at least one independent experiment. False positives are colored in orange and false negatives are colored in green.

Target	Caffeic acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH_2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
ABL1	4	4					2	2	2
ABL2	4	4							
APS (SH2B2)	4								
BLK	4								
BLNK	2								
BRK	4						2	2	2
CBL	4							2	
CBLB	2								
CBLC	10							2	4
CISH	38		2				20	2	4
CRK	12							2	
CRKL									
CSK								2	
DAPP1									
FER	40						2	2	4
FES	4			2					
FGR	40						2	4	2
FRK	4								
FYN	4			2					
GADS								2	
GRAP		4							
GRB10	4							2	
GRB14	24							2	
GRB2		8						2	
GRB7	3	4							
HCK	4								
HSH2D		2	10				2	2	2
ITK	30						2	4	
LCK	4								
LNK (SH2B3)	12		2					2	
LYN	4								

Target	Caffeic acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH_2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
MATK								2	
MIST									
NCK1	3							2	
NCK2	3								
PIK3R1 (SH2dom.C-term.)	29	3						2	4
PIK3R1 (SH2dom.N-term.)	8								
PIK3R2 (SH2dom.C-term.)	2	2	2		2			2	4
PIK3R2(SH2dom.N-term.)	4								
PIK3R3 (SH2dom.C-term.)	34	2	2	3				2	4
PIK3R3 (SH2dom.N-term.)	4								
PLCG1 (SH2dom.C-term.)	15							2	4
PLCG1 (SH2dom.N-term.)	21							2	4
PLCG2 (SH2dom.C-term.)	13							4	
PLCG2 (SH2dom.N-term.)	10							2	
PTPN11 (SH2dom.C-term.)								2	
PTPN11 (SH2dom.N-term.)									
PTPN6 (SH2dom.C-term.)									
PTPN6 (SH2dom.N-term.)							4		
RASA1 (SH2dom.C-term.)									
RASA1 (SH2dom.N-term.)	3								
SH2B (SH2B1)	3								
SH2D1A									
SH2D1B	2								
SH2D6									
SH3BP2	3								
SHB	10							2	
SHC1	3								
SHC2	3								
SHC3	10	4						2	
SHC4	3								
SHF									
SHIP1	4		2					2	
SHIP2	12		2					3	
SLAP2	5								
SLP76									
SOCS1	5								

Target	Caffeic acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE-NH_2	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
SOCS2	12		2					4	4
SOCS3	18		2						2
SOCS4	10		2		2		4	2	4
SOCS6	4								
SOCS7	12		2					2	
SRC	4								
SRMS	31		2				2	2	2
STAP1									
STAP2									
STAT1	2				87	13	6	9	11
STAT2	2	10			99	21	6		4
STAT3					122	13		9	11
STAT3(D661Y)-nonphosphorylated					55	13		9	10
STAT3(E616del)-nonphosphorylated					8	8		9	11
STAT3(N567K)-nonphosphorylated					8	6		9	11
STAT3(N647I)-nonphosphorylated					9	6		9	10
STAT3(S614R)-nonphosphorylated					9	6		9	11
STAT3(Y640F)-nonphosphorylated					48	10		7	11
STAT4	2		2		74	13		9	11
STAT5A	2				69	13	8	9	11
STAT5B	3				79	12	8	9	11
STAT5B(N642H)	2				56	12	8	9	11
STAT6					156	13	6	7	10
SYK (SH2dom1.dom2.)									
TEC	4								
TNS1	4	10							
TNS2	3	5							
TNS3	3	4							
TNS4	12	4	2				2	2	4
TXK	4								
VAV1	3	4							
VAV2	3								
VAV3	4	4							
YES	31			2				4	
ZAP70 (SH2dom1.dom2.)	2								

Extended Data Table 6|Literature SH2 domain binding measurements for compounds relevant to this study. Literature values for the SH2 domain binding of compounds tested or related to those tested in this study are shown below in nM. Units reported for these measurements are listed in brackets following the value in the table.

Target	Caffeic acid-pYEEIE	CGP78850	MN551	Ac-pYEEIE	SD-36	SI-109	Stafib-1	STAT6-IN-1	STAT6-IN-3
FYN	84 (IC_{50})	>50000 (IC_{50})		1100 (IC_{50})					
GRB2		40 (IC_{50})							
LCK	42 (IC_{50})			1300 (IC_{50})					
PIK3R1 (SH2dom.N-term.)		>50000 (IC_{50})							
PTPN11 (SH2dom.C-term.)		>50000 (IC_{50})							
PTPN11 (SH2dom.N-term.)		>50000 (IC_{50})							
SOCS2			2180 (K_D)						
SRC	64 (IC_{50})			2000 (IC_{50})					
STAT1					1086 (K_D)	1270 (K_D)			
STAT2					4883 (K_D)	6630 (K_D)			
STAT3					44.4 (K_D)	51.2 (K_D)			
STAT4					804 (K_D)	1980 (K_D)	3290 (K_I)		
STAT5A					6180 (K_D)		2420 (K_I)		
STAT5B					11700 (K_D)		44 (K_I)		
STAT6					11325 (K_D)	3370 (K_D)	3470 (K_I)	28 (IC_{50})	44 (IC_{50})
References	Park *et al.*, Bioorganic & Medicinal Chemistry Letters, 2002	Gay *et al.*, International Journal of Cancer, 1999	Ramachandran *et al.*, Nature Communications, 2023	Park *et al.*, Bioorganic & Medicinal Chemistry Letters, 2002	Bai *et al.*, Cancer Cell, 2019	Bai *et al.*, Cancer Cell, 2019	Elumalai *et al.*, Angewandte Chemie International Edition, 2015	Mandal *et al.*, Journal of Medicinal Chemistry, 2015	Mandal *et al.*, Journal of Medicinal Chemistry, 2015