The pathogenic T42A mutation in SHP2 rewires interaction specificity and enhances signaling

Anne E. van Vlimmeren[1,2,†], Rashmi Voleti[1,†], Cassandra A. Chartier[1], Ziyuan Jiang[1], Deepti Karandur[3], Neel H. Shah[1,*]

[1] Department of Chemistry, Columbia University, New York, NY 10027

[2] Department of Biological Sciences, Columbia University, New York, NY 10027

[3] Department of Biochemistry, Vanderbilt University, Nashville, TN 37232

[†] These authors contributed equally.

* corresponding author: neel.shah@columbia.edu

ORCiDs

Anne E. van Vlimmeren: 0000-0003-0379-4945

Rashmi Voleti: 0000-0002-3705-7460

Cassandra A. Chartier: 0009-0004-2474-9937

Ziyuan Jiang: 0009-0003-6035-4126

Deepti Karandur: 0000-0002-6949-6337

Neel H. Shah: 0000-0002-1186-0626

Abstract

Most disease-associated mutations in the tyrosine phosphatase SHP2 increase its basal catalytic activity by destabilizing its auto-inhibited state. By contrast, some mutations do not increase the basal activity of SHP2 and likely exert their pathogenicity through alternative mechanisms. We lack a molecular understanding of how these mutations impact SHP2 structure, activity, and signaling. Here, we characterize five such mutations that occur within the ligand binding pockets of the regulatory phosphotyrosine-recognition domains of SHP2, using high-throughput biochemical screens, biophysical and biochemical measurements, molecular dynamics simulations, and cellular assays. While many of these mutations impact binding affinity, the T42A mutation is unique in that it also alters ligand binding specificity. As a result, the T42A mutant has biased sensitivity toward a subset of activating phosphoproteins. Our study highlights an example of a nuanced mechanism of action for a disease-associated mutation, characterized by a change in protein-protein interaction specificity that alters enzyme activation.

Keywords

Tyrosine phosphatase, SH2 domain, sequence specificity, Noonan syndrome, SHP2 mutations, SHP2 activation

Introduction

SHP2 (Src homology-2 domain-containing protein tyrosine phosphatase-2) is a ubiquitously expressed protein tyrosine phosphatase, encoded by the *PTPN11* gene. It has critical roles in many biological processes, including cell proliferation, development, immune-regulation, metabolism, and differentiation.[1–3] Mutations in *PTPN11* are associated with a variety of diseases, most notably the congenital disorder Noonan syndrome. Germline *PTPN11* mutations underlie approximately 50% of Noonan Syndrome cases, which are characterized by a wide range of developmental defects.[4–6] In addition, somatic mutations in *PTPN11* have been found in roughly 35% of patients with juvenile myelomonocytic leukemia (JMML), a rare pediatric cancer.[7,8] *PTPN11* mutations also occur in other types of leukemia, such as acute myeloid leukemia, acute lymphoid leukemia and chronic myelomonocytic leukemia, as well as solid tumors, albeit at lower incidence.[9]

SHP2 has three globular domains: a protein tyrosine phosphatase (PTP) domain, which catalyzes the dephosphorylation of tyrosine-phosphorylated proteins, and two Src Homology 2 (SH2) domains, which are phosphotyrosine (pTyr)-recognition domains (**Figure 1A**)**.** The SH2 domains regulate SHP2 activity by dictating localization and through allosteric control of catalytic activity. Interactions between the N-SH2 domain and PTP domain limit substrate access by blocking the catalytic site, leading to an auto-inhibited state with low basal catalytic activity. Conformational changes of the N-SH2 domain, caused by its binding to tyrosine-phosphorylated proteins, disrupt the N-SH2/PTP interaction to activate SHP2 in a ligand-dependent manner (**Figure 1B,C**).[10] Thus, the N-SH2 domain couples the localization of SHP2 to its activation by specific upstream signals.

Over 100 disease-associated mutations in *PTPN11* have been reported, yielding amino acid substitutions at more than 30 different residues spanning all three globular domains.[4] Most mutations in SHP2 are at the N-SH2/PTP auto-inhibitory interface and shift the conformational equilibrium of SHP2 towards the active state (**Figure 1D**).[10–13] These mutations cause SHP2 to populate a catalytically active state irrespective of localization or activating stimuli. By contrast, some pathogenic mutations are found in the pTyr-binding pockets of the N- and C-SH2 domains and are mechanistically distinct, as they have the potential to change the nature of SH2-phosphoprotein interactions. T42A, a Noonan Syndrome mutation in the pTyr-binding pocket of the N-SH2 domain, has been reported to enhance binding affinity for various SHP2 interactors (**Figure 1E**).[14] Thus, this mutation is thought to make SHP2 more readily activated by upstream phosphoproteins, while still requiring binding and localization to those phosphoproteins for functional signaling. Beyond its known effect on ligand binding affinity, the precise effect of the T42A mutation on specific cell signaling processes remains elusive. Two nearby mutations in the N-SH2 domain, L43F and T52S (**Figure 1E**), are associated with non-syndromic heart defects and JMML, respectively, but very little is known about their effects on ligand binding or cell signaling.[15,16] The C-SH2 domain mutant R138Q has been identified in melanoma, whereas the E139D mutation has been associated with both JMML and Noonan Syndrome (**Figure 1F**).[9,17] Insights into the molecular mechanisms underlying these pathogenic pTyr-binding pocket mutations could further our understanding of how they dysregulate cellular signaling, and in turn, tumorigenesis or development.

In this study, we extensively characterize the binding properties of five disease-associated SH2-domain mutations in SHP2. Through a series of biophysical measurements and high-throughput peptide-binding screens, we demonstrate that the T42A mutation in the N-SH2 domain is unique among these mutations in that it selectively enhances binding to specific phosphopeptide sequences. Through molecular dynamics simulations and further biochemical experiments, we identify structural changes caused by the T42A mutation that likely explain its altered ligand binding specificity. We show that this change in specificity within the N-SH2 domain results in sequence-dependent changes to the activation of full-length SHP2 by phosphopeptide ligands. Finally, we demonstrate that these findings are robust in a cellular context, by showing that $SHP2^{T42A}$ binds tighter than $SHP2^{WT}$ to several full-length phosphoprotein interactors and enhances downstream signaling. Our results suggest that the pathogenicity of $SHP2^{T42A}$ could be due to biased sensitization to specific upstream signaling partners, caused by rewiring of its interaction specificity.

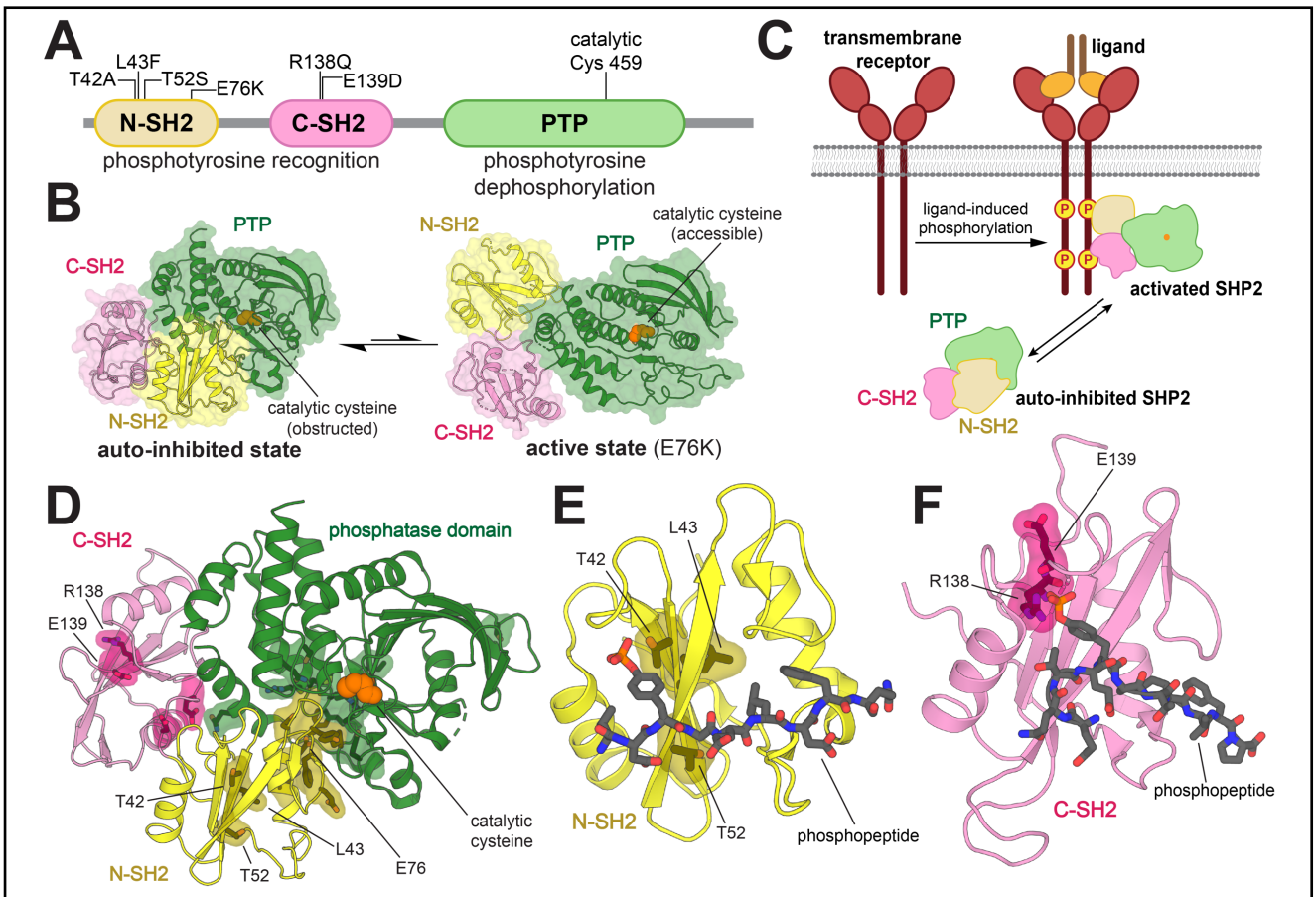

Figure 1. Structure and regulation of SHP2. (**A**) Domain architecture diagram of SHP2. Relevant mutations and the catalytic cysteine (C459) are indicated. (**B**) SHP2 is kept in its auto-inhibited state by interactions between the N-SH2 and PTP domain (PDB: 4DGP). In its active state, the catalytic cysteine is accessible. The structure of SHP2^{E76K} (PDB: 6CRF) is used to represent the active state. (**C**) The SH2 domains of SHP2 bind to tyrosine-phosphorylated upstream proteins, such as transmembrane receptors, inducing a conformational change that activates SHP2. (**D**) Disease-associated mutations cluster largely, but not exclusively, on the interdomain interface between the N-SH2 and the PTP domain (PDB: 4DGP). (**E**) Mutations in or near the N-SH2 binding pocket (PDB: 6ROY). (**F**) Mutations in or near the C-SH2 binding pocket (PDB: 6R5G).

Results

<u>Mutations in the SH2 domains of SHP2 impact both binding affinity and sequence specificity</u>

Mutations in the SH2 domains of SHP2, proximal to the ligand binding region, have the potential to change both the affinity and the specificity of the SH2 domain, thereby affecting SH2 domain functions such as recruitment, localization, and allosteric regulation of SHP2 activity. We focused on three mutations in the N-SH2 domain (T42A, L43F, and T52S) that are both disease-relevant and close to the pTyr-binding pocket (**Figure 1E**). These mutations do not cause a significant increase in basal phosphatase activity, in contrast to the well-studied JMML mutation E76K, which lies at the auto-inhibitory interface and dramatically enhances phosphatase activity (**Figure S1 and Table S1**).[17–19] Additionally, we studied R138Q and E139D, two disease-associated mutations in the pTyr-binding pocket of the C-SH2 domain (**Figure 1F**). E139D causes a 15-fold increase in basal phosphatase activity (**Figure S1 and Table S1**), as has been reported previously.[17,20] The R138Q mutation is expected to disrupt phosphopeptide binding, as Arg 138 is part of the conserved FLVR motif in SH2 domains, and directly coordinates the phosphoryl group of phosphotyrosine ligand.[21,22] This mutation had no impact on the basal catalytic activity of SHP2 (**Figure S1 and Table S1**).

Using a fluorescence polarization assay, we measured the binding affinity of a fluorescent phosphopeptide derived from a known SHP2 binding site (pTyr 1179) on insulin receptor substrate 1 (IRS-1) against all the four N-SH2 domain variants (**Figure S2A and Table S2**).[23] We found that N-SH2^{T42A} binds 5-fold tighter to this peptide compared to N-SH2^{WT}, consistent with previous literature demonstrating enhanced binding for N-SH2^{T42A}.[14] Next, in competition binding experiments, we tested 8 unlabeled phosphopeptides derived from known SHP2 binders, and one unlabeled phosphopeptide

(Imhof-9) based on an unnatural ligand discovered in a previously reported peptide screen (**Figure 2, Figure S2B, and Table S2**).[23–29] For N-SH2^{L43F} and N-SH2^{T52S}, we observed a 2- to 3-fold increase in binding affinity for some peptides when compared to N-SH2^{WT} (**Figure 2C,D**). By contrast, we observed a broad range of effects on binding affinity for N-SH2^{T42A}. N-SH2^{T42A} displayed a 28-fold increase in affinity for the PD-1 pTyr 223 phosphopeptide, while a 20-fold increase was observed for Gab2 pTyr 614 (**Figure 2B**). The increase in affinity for other peptides was more moderate, ranging from 4- to 6-fold. This suggests that the T42A mutation selectively enhances the affinity of the N-SH2 domain for specific peptides.

For C-SH2 domain mutants R138Q and E139D, we first measured binding against two fluorescent phosphopeptides: one derived from a known binding site (pTyr 248) on PD-1, as well as the designed ligand Imhof-9 (**Figure S2C and Table S2**).[29,30] As expected, C-SH2^{R138Q} binding to phosphopeptides was severely attenuated (**Figure S2C**), and this mutant was excluded from further binding analyses. C-SH2^{E139D} binding was comparable to C-SH2^{WT} binding against the two fluorescent phosphopeptides (**Figure S2C**) and against the 9 unlabeled peptides used for N-SH2 binding assays (**Figure S2D and Table S2**). Notably, most of these peptides (aside from PD-1 pTyr 248) are derived from known N-SH2 binding sites and have a relatively weak affinity for the SHP2 C-SH2 domain, compared to the N-SH2 domain. An effect of the E139D mutation on ligand binding affinity or specificity may be more apparent with native C-SH2 ligands. Collectively, these N-SH2 and C-SH2 binding experiments with a small panel of peptides suggest that N-SH2^{T42A} is unique among the SH2 mutants in its impact on both phosphopeptide binding affinity and specificity.

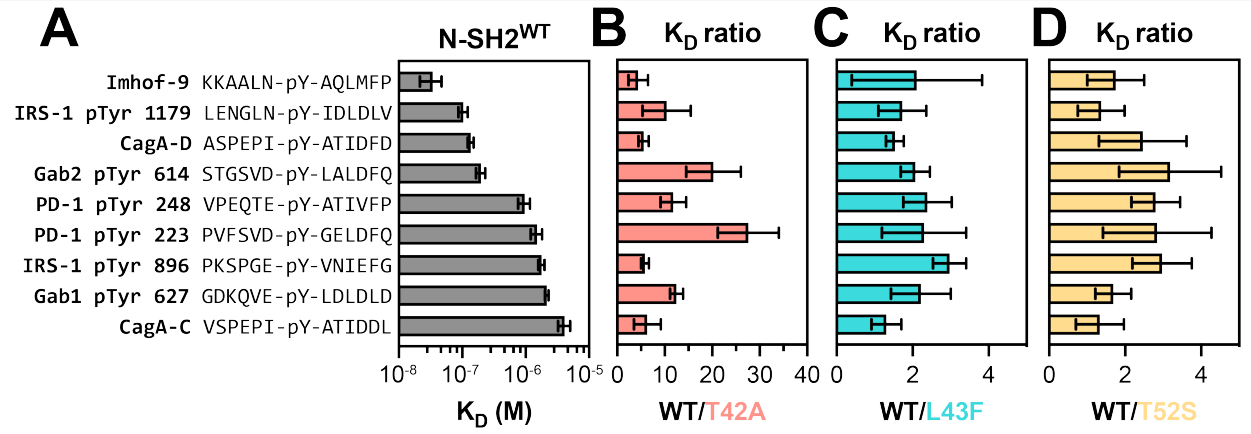

Figure 2. K_D measurements reveal sequence-specific enhancement of binding affinity in SHP2^{T42A}. (**A**) Measured binding affinities of N-SH2^{WT} against peptides derived from various known SHP2 interactors. (**B**) Fold-change in K_D for N-SH2^{T42A} compared to N-SH2^{WT}, for each of the peptides shown in panel (A). (**C**) Same as (B), but for N-SH2^{L43F}. (**D**) Same as (B), but for N-SH2^{T52S}. Source data can be found in Table S2.

Human phosphopeptide profiling reveals the scope of specificity differences in SHP2 SH2 domain mutants

Next, we characterized the sequence preferences of these SH2 mutants on a larger scale. We employed a bacterial peptide display screen recently developed in our lab to profile SH2 domain sequence specificity.[31] In this method, a genetically-encoded library of peptides is expressed on the surface of bacterial cells, and tyrosine residues on these peptides are enzymatically phosphorylated. Cells are then mixed with SH2-coated magnetic beads to enrich for cells displaying phosphopeptides with high affinity for the SH2 domain. The peptide-coding DNA sequences in the enriched sample and an unenriched input library sample are deep-sequenced. For each peptide, an enrichment score is calculated as the frequency of each peptide in the SH2-enriched sample divided by the frequency of that peptide in the unenriched input sample.[31] For this study, we used two largely non-overlapping peptide libraries, both encoding known human phosphosites. The pTyr-Var Library contains 3,065 sequences corresponding to wild-type tyrosine phosphosites, with an additional 6,833 sequences encoding disease-associated point mutations, natural polymorphisms, or control mutations.[31] The Human pTyr Library consists of 1,916 sequences derived from known phosphorylation sites in the human proteome, along with another 617 control mutants.[32]

We profiled N-SH2^{WT}, N-SH2^{T42A}, N-SH2^{L43F}, N-SH2^{T52S}, C-SH2^{WT}, and C-SH2^{E139D}, against both libraries described above. The libraries were screened separately, but under identical conditions, and the spread of peptide enrichment scores was similar across both libraries. Thus, the results of both screens were combined for the analyses described below. We omitted sequences from our analysis that contained more than one tyrosine residue, yielding 9,281 relevant sequences across both libraries. For most phosphopeptides the screens showed a strong correlation between enrichment scores for the wild-type SH2 domain and the corresponding SH2 mutants. However, some phosphopeptides had larger enrichment scores for the mutant N-SH2 domains when compared to N-SH2^{WT} (**Figure 3A-C and Table S3**). This effect was strongest for N-SH2^{T42A}, both in magnitude and in number of phosphopeptides that were disproportionately enriched in the N-SH2^{T42A} screens. In the C-SH2 domain screens, C-SH2^{E139D} showed slightly weakened binding to some peptides when compared to C-SH2^{WT} (**Figure 3D and Table S3**), in contrast to our binding affinity measurements (**Figure S2D**). This result is in alignment with previous work showing a change in binding preferences for C-SH2^{E139D}, and it reinforces the importance of screening a large number of peptides for an unbiased assessment of specificity.[14]

To validate the stark difference between N-SH2^{WT} and N-SH2^{T42A} in our screens, we synthesized and purified three representative sequences to use in competition binding assays (**Figure 3E**). One of these peptides, DNMT3B pTyr 815, was derived from DNA methyltransferase 3β, a protein that is not known to interact with SHP2. The second peptide, PGFRB pTyr 763, was derived from the platelet-derived growth factor receptor β, which is known to interact with SHP2 through this phosphosite.[33] The third peptide, MILR1 pTyr 338, was derived from a known SHP2 binding site on the mast cell immunoreceptor Allegrin-1.[34] Competition fluorescence polarization assays with these peptides revealed large differences in binding affinity between N-SH2^{WT} and N-SH2^{T42A}, as predicted by the screens (**Figure 3F and Table S2**). N-SH2^{T42A} bound 17-fold tighter to DNMT3B pTyr 815 than N-SH2^{WT} and 24-fold tighter to PGFRB pTyr 763. The difference in binding affinity between N-SH2^{WT} and N-SH2^{T42A} was largest for MILR1 pTyr 338, for which the mutation caused a 90-fold enhancement.

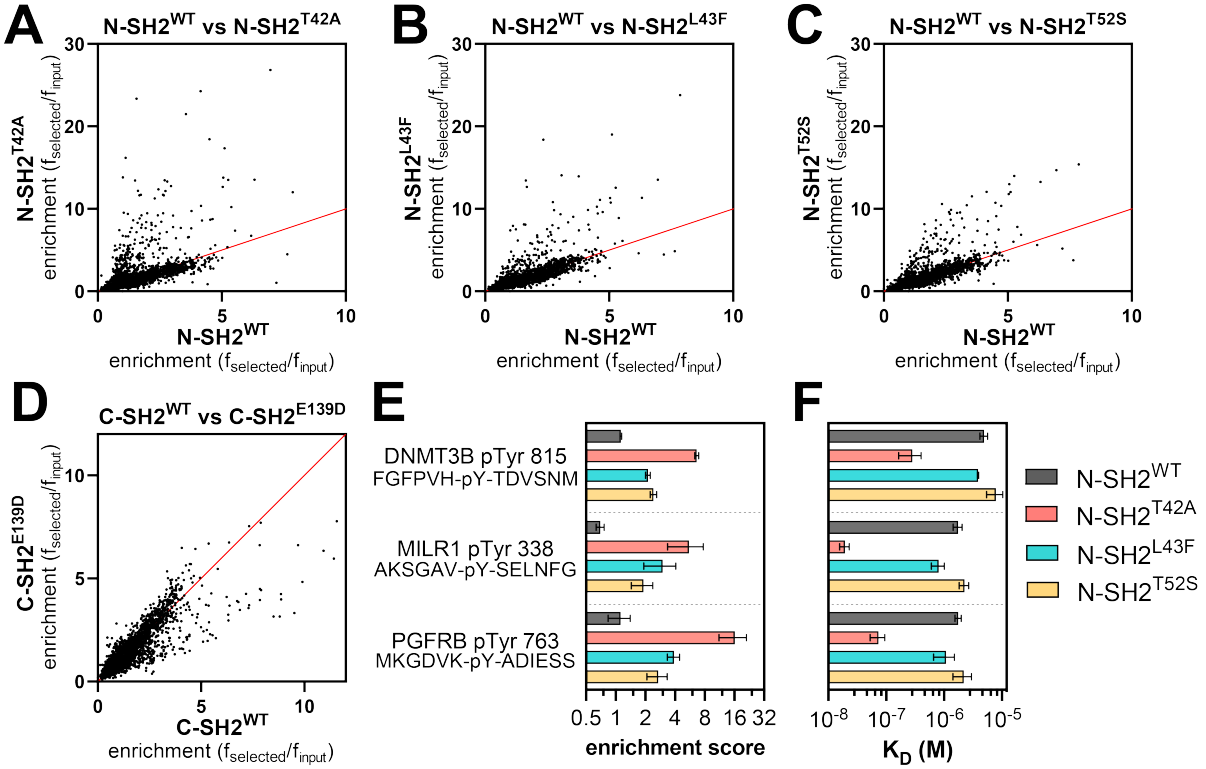

Figure 3. Peptide library screens identify sequences with enhanced SHP2^{T42A} binding. (**A-C**) Comparison of enrichment scores for N-SH2^{WT} and each N-SH2 mutant. (**D**) Comparison of enrichment scores for C-SH2^{WT} and C-SH2^{E139D}. For panels (A)-(D), the red line denotes where the wild-type and mutant SH2 have the same enrichment values. For each SH2 domain, data are the average of two pTyr-Var Library screens and three Human pTyr Library screens. Source data for panels (A)-(D) can be found in Table S3. (**E**) Enrichment scores from peptide display screens for three representative peptides that showed enhanced binding to N-SH2^{T42A} relative to N-SH2^{WT}. (**F**) Binding affinity measurements for the three peptides shown in panel (E).

SHP2 N-SH2^{WT} and N-SH2^{T42A} display distinct position-specific sequence preferences

To fully describe the differences in specificity between SHP2 N-SH2^{WT} and the N-SH2 mutants, we examined the sequence features of the peptides enriched in the screens with each of these domains. Our peptide libraries collectively contain 392 sequences lacking tyrosine residues, which serve as negative controls in our screens. Less than 2% of the negative control peptides had enrichment scores above of 3.2, and so we used this value as a stringent cutoff to identify true binders in each screen, as done previously.[31] We identified 168 enriched sequences for N-SH2^{WT} and approximately 250 enriched sequences for each of the N-SH2 mutants, indicative of overall tighter binding by the mutants (**Figure 4A, Figure S3A, and Table S3**). Consistent with its unique change in binding specificity, the enriched peptide set for N-SH2^{T42A} had less overlap with that of N-SH2^{WT} when compared with N-SH2^{L43F} or N-SH2^{T52S}. Probability sequence logos, derived by comparing the amino acid composition of these enriched peptide sets to the full library, showed that N-SH2^{T42A} had the most distinctive sequence preferences of all four N-SH2 variants (**Figure S3B-E**).[35]

Due to the small number of enriched sequences in the N-SH2^{WT} screens, the corresponding sequence logo has low signal-to-noise ratio. Even still, the logo highlights several hallmarks of the SHP2 N-SH2 domain, such as a preference for a -2 Ile, Leu or Val, -1 His, +3 hydrophobic residue, and +5 Phe or His.[29,36] As expected, the N-SH2 mutants share many of these features with N-SH2^{WT} (**Figure S3B-E**). However, we observed distinct changes in specificity at positions closest to the pTyr residue. N-SH2^{T42A} prefers the smaller Val over Ile and Leu on the -2 position (**Figure 4B and Figure S3B,C**). At the -1 position, although His is strongly favored for all four SH2 domains, N-SH2^{T42A} had broadened tolerance of other amino acids, including Pro and polar residues Gln, Ser, Thr, and Arg (**Figure 4C and Figure S3B,C**). At the +1 position, N-SH2^{WT} favors large hydrophobic residues (Leu, Ile, or Phe), as well as His and Asn. By contrast, Ala is the dominant preference for N-SH2^{T42A} and is also enriched for N-SH2^{L43F} and N-SH2^{T52S}, but to a lesser extent (**Figure 4D and Figure S3B-E**). At the +2 position, we found that N-SH2^{T42A} uniquely has an enhanced preference for hydrophilic residues. One notable difference, which will be discussed in subsequent sections, is a switch for +2 Glu from a disfavored residue for N-SH2^{WT} to a favored residue for N-SH2^{T42A} (**Figure 4E and Figure S3B,C**). Finally, at the +3 residue, each N-SH2 variant shows a slightly different preference for specific hydrophobic residues, with N-SH2^{T42A} strongly preferring Leu (**Figure S3B-E**).

The sequence logos represent the position-specific amino acid preferences of each N-SH2 variant, without taking into account the surrounding context of a specific sequence. The pTyr-Var Library used in the screens encodes wild-type and point-mutant sequences derived from human phosphorylation sites, providing an internal control for sequence-specific mutational effects. Upon closer inspection of individual hits for N-SH2^{WT} and N-SH2^{T42A}, we identified several sets of sequences that corroborate the overall preferences described above. These include an enhanced preference in N-SH2^{T42A} for Pro or Thr at the -1 position over Leu or Ile and a strong preference for a +2 Glu residue (**Figure 4F**).

To more comprehensively analyze sequence preferences in a physiologically-relevant sequence context, we generated a saturation mutagenesis library based on the sequence surrounding PD-1 pTyr 223, and we screened this library against the N-SH2^{WT} and N-SH2^{T42A} proteins using our bacterial peptide display platform. The immunoreceptor tyrosine-based inhibitory motif (ITIM) surrounding PD-1 pTyr 223 was chosen because this was a sequence for which we observed a large change in binding affinity between N-SH2^{WT} and N-SH2^{T42A} (**Figure 2B**). Due to the relatively weak binding of N-SH2^{WT} to the wild-type PD-1 pTyr 223 peptide, the differentiation of neutral mutations and loss-of-function mutations was poor in our screen (**Figure 4G, Figure S3F, and Table S4**). However, we could confidently detect gain-of-function mutations for N-SH2^{WT}. For N-SH2^{T42A}, the overall tighter binding affinity allowed for reliable measurement of both gain- and loss-of-function point mutations on this peptide (**Figure 4G, Figure S3F, and Table S4**).

Our results show that the two domains have modestly correlated binding preferences with respect to this scanning mutagenesis library (**Figure S3F**). Interestingly, the -1 Asp and +1 Gly residues in the ITIM sequence are suboptimal for both N-SH2^{WT} and N-SH2^{T42A}, as most substitutions at these positions enhance binding. However, differences were observed in which mutations were tolerated by each SH2 domain at these positions (**Figure 4G**). For example, the substitution of the +1 Gly to Ala or Thr is favored by N-SH2^{WT}, consistent with previous studies,[36] but large hydrophobic residues are also favorable for N-SH2^{WT} at this position. By contrast, N-SH2^{T42A} strongly disfavors a

+1 Trp and Phe. This recapitulates our analysis of sequences enriched in the human phosphopeptide library screens, where we observed a N-SH2^{WT} preference for larger residues (Leu, Ile, Phe), whereas N-SH2^{T42A} had a strong preference for the smaller alanine (**Figure 4D**). Also consistent with our analysis of the human phosphopeptide screens, most substitutions at the -2 Val or +2 Glu in the ITIM are heavily disfavored by N-SH2^{T42A} (**Figure 4G**). Taken together, our experiments with the phosphopeptide library screens and the scanning mutagenesis library highlight consistent differences in the sequence preferences of N-SH2^{WT} and N-SH2^{T42A}, suggestive of distinct modes of phosphopeptide engagement by these two domains.

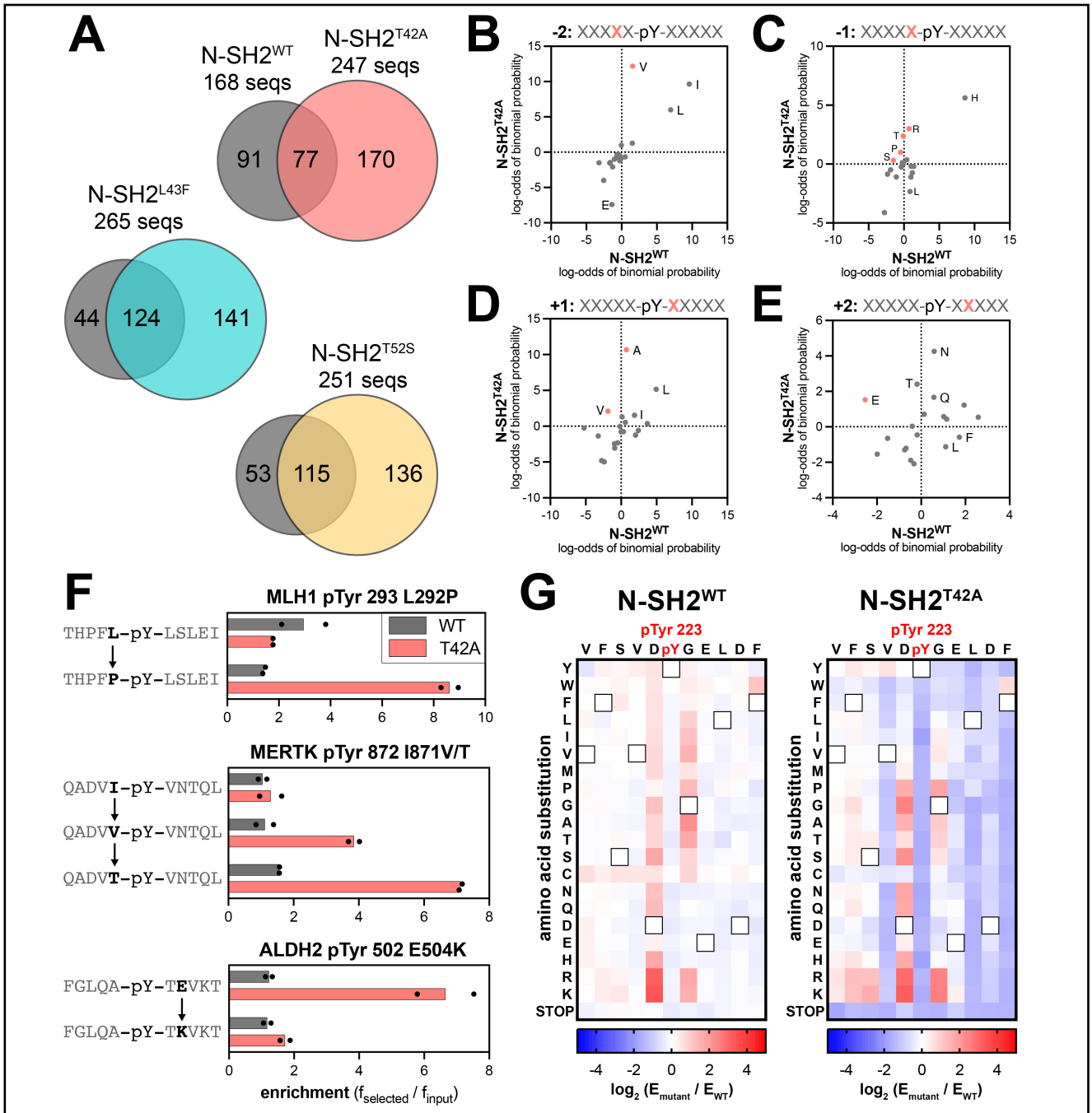

Figure 4. Analysis of enriched sequence features reveals specificity differences at several positions. (**A**) Overlap in sequences enriched by each N-SH2 domain above an enrichment score cutoff of 3.2. (**B**) Log-transformed probabilities of amino acid enrichment at the -2 position relative to the pTyr residue, derived from peptide display screens with N-SH2^{WT} and N-SH2^{T42A}. (**C**) Same as panel (B), except at the -1 position. (**D**) Same as panel (B), except at the +1 position. (**E**) Same as panel (B), except at the +2 position. (**F**) Enrichment scores for representative sets of peptides from the pTyr-Var Library screens that highlight different sequence preferences for N-SH2^{WT} and N-SH2^{T42A}. (**G**) Heatmaps of N-SH2^{WT} and N-SH2^{T42A} for PD-1 pTyr 223 (ITIM) scanning mutagenesis screen (average of 5 replicates). The wild-type residue on each position is indicated by a black square. Color scheme: blue indicates weaker binding, white indicates same as wild-type, red indicates tighter binding. Source data for the heatmaps in panel (G) can be found in Table S4.

The T42A mutation enhances binding by remodeling the N-SH2 phosphotyrosine binding pocket

Several structural explanations for tighter phosphopeptide binding by N-SH2^{T42A} have been postulated previously, but no studies have addressed the molecular basis for a change in specificity.[14,17,37–39] Crystal structures of N-SH2^{WT} bound to different peptides show that the hydroxyl group of the Thr 42 side chain hydrogen bonds to a non-bridging oxygen atom on the phosphotyrosine moiety of the ligand (**Figure 5A**).[30,40–43] The loss of this hydrogen bond in the T42A mutant is thought to be counterbalanced by enhanced hydrophobic interactions between the pTyr phenyl ring and Ala 42, but this cannot explain differences in the recognition of the surrounding peptide sequence, which is over 10 Å away. Many SH2 domains have a hydrophobic residue (Val, Ala, or Leu) at the position corresponding to Thr 42 (**Figure S4A**),[44] but the impact of this residue on sequence specificity has not been systematically explored. High-affinity SH2 domains have been engineered using bulky hydrophobic substitutions at this and other pTyr-proximal positions.[45,46] Structural analyses of these "superbinder" mutants suggest that those mutations can enhance hydrophobic interactions and also pre-organize the hydrogen bonding network around the phosphoryl group to further enhance binding.[37,45,46] Inspired by the structural rearrangements observed in superbinder mutants, we used molecular dynamics (MD) simulations to examine how the T42A mutation impacts SHP2 N-SH2 peptide engagement.

We carried out simulations of SHP2 N-SH2^{WT} and N-SH2^{T42A} in the apo state and bound to five different phosphopeptide ligands (PD-1 pTyr 223, MILR1 pTyr 338, Gab2 pTyr 614, IRS-1 pTyr 896, and Imhof-9). Each system was simulated three times, for 1 μs per trajectory. We first calculated the per-residue root mean squared fluctuation (RMSF) in each system (**Figure S5A**). Simulations of the peptide-bound state showed rigidification of the BC loop (residues 32-40) relative to apo-state simulations. This loop is responsible for coordinating the phosphoryl moiety through a series of hydrogen bonds. Peptide binding also reduced fluctuations around the EF and BG loops (residues 64-70 and 86-96, respectively). These regions recognize the strongly preferred +5 Phe found in all five peptides. Differences in RMSF values between N-SH2^{WT} and N-SH2^{T42A} were largely negligible in both the apo and peptide-bound states, with one noteworthy exception: the BC loop in two-thirds of the N-SH2^{WT} simulations with the PD-1 and MILR1 peptides showed significantly more fluctuation than in the N-SH2^{T42A} simulations (**Figure S5A**). These two peptides showed the largest fold-changes in binding affinity upon T42A mutation (**Figures 2B and 3F**). Our simulations suggest that N-SH2^{WT} cannot engage the phosphotyrosine residue in these peptides as stably as N-SH2^{T42A}, which may contribute to the large enhancement in binding affinity for the mutant.

Closer inspection of BC-loop interactions in the N-SH2^{WT} and N-SH2^{T42A} simulations revealed substantial reorganization of the hydrogen bond network around the phosphoryl group, which alters the positioning of phosphotyrosine within the N-SH2 binding pocket. In every N-SH2^{WT} simulation, Thr 42 makes a persistent hydrogen bond with a non-bridging oxygen on the phosphoryl group, and this interaction constrains the orientation of the phosphotyrosine residue (**Figure 5B and Figure S5B,C**). This bond breaks intermittently in N-SH2^{WT} simulations with the PD-1, MILR1, and Gab2 peptides, but not the tighter binding IRS-1 and Imhof-9 peptides (**Figure S5C**). In general, phosphotyrosine engagement in the N-SH2^{WT} simulations resembles that seen in crystal structures (**Figure 5A,B and Figure S5B**).

In almost every N-SH2^{T42A} simulation, where the phosphoryl group is not tethered to Thr 42, the phosphotyrosine residue relaxes into a new orientation characterized by a distinct hydrogen bond network (**Figure 5C and Figure S5B**). The most notable change is how Arg 32 of the conserved FLVR motif[44] interacts with the phosphoryl group: in the N-SH2^{WT} simulations, the bidentate guanidium-phosphoryl interaction involves one non-bridging oxygen and the less electron-rich bridging oxygen, but in the N-SH2^{T42A} simulations, Arg 32 coordinates two non-bridging oxygens, in a presumably stronger interaction (**Figures 5B,C and Figure S5B,D**).[47] Additionally, some interactions that exist in both the N-SH2^{WT} and N-SH2^{T42A} simulations become more persistent in the N-SH2^{T42A} simulation, such as a hydrogen bond between backbone amide nitrogen of Ser 36 and a non-bridging oxygen on the phosphoryl group (**Figure S5B,E**). Collectively, our analyses show that the T42A mutation remodels the phosphotyrosine binding pocket, which enhances phosphopeptide binding affinity. These observations also hint at subtle peptide-specific differences in complex stability between N-SH2^{WT} and N-SH2^{T42A}.

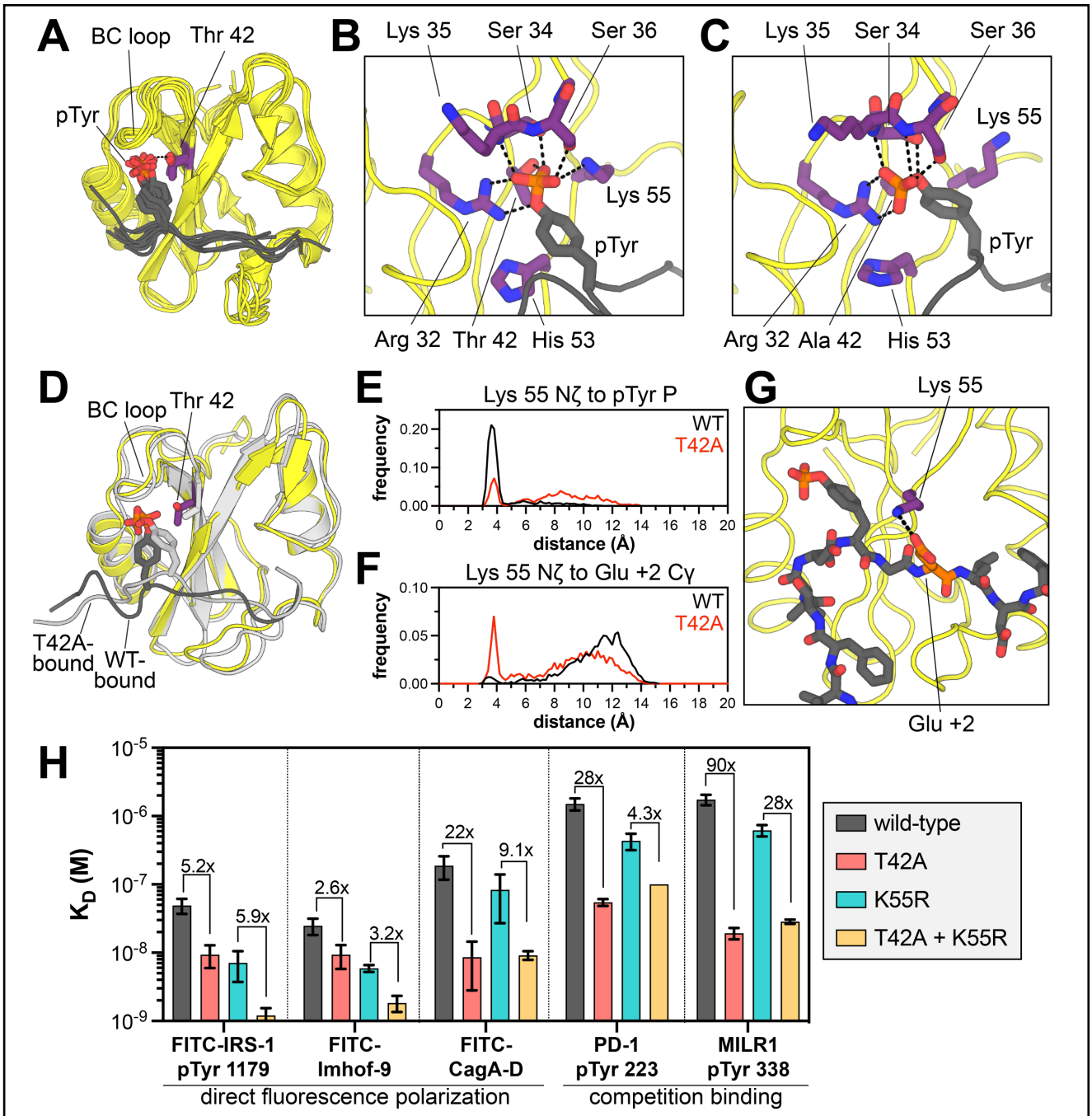

Figure 5. Structural impact of the T42A mutation on phosphotyrosine and proximal sequence recognition. (**A**) Hydrogen bonding of Thr 42 in SHP2 N-SH2^{WT} to the phosphoryl group of phosphopeptide ligands in several crystal structures (PDB: 6ROY, 1AYA, 1AYB, 3TL0, 5DF6, 5X94, and 5X7B). (**B**) Structure of N-SH2^{WT} bound to the PD-1 pTyr 223 (ITIM) peptide at the end of a 1 μs MD simulation. (**C**) Structure of N-SH2^{T42A} bound to the PD-1 pTyr 223 (ITIM) peptide at the end of a 1 μs MD simulation. (**D**) Overlay of the states shown in panels B and C. The N-SH2^{WT} state is in yellow with a dark-gray ligand. The N-SH2^{T42A} state is in light gray, with a light gray ligand. (**E**) Distribution of distances between the Lys 55 Nζ atom and the phosphotyrosine phosphorus atom in simulations of the PD-1 pTyr 223 peptide bound to N-SH2^{WT} (black) or N-SH2^{T42A} (red). (**F**) Distribution of distances between the Lys 55 Nζ atom and the +2 Glu Cδ atom in simulations of the PD-1 pTyr 223 peptide bound to N-SH2^{WT} (black) or N-SH2^{T42A} (red). (**G**) An ion pair between Lys 55 and the +2 Glu residue (Glu 225) in the PD-1 pTyr 223 (ITIM) peptide, frequently observed in N-SH2^{T42A} simulations. (**H**) Peptide-specific effects of the T42A mutation in the presence and absence of the K55R mutation.

T42A-dependent changes in phosphotyrosine engagement drive changes in sequence recognition

A major consequence of the reshuffling of hydrogen bonds between N-SH2^{WT} and N-SH2^{T42A} is that the phosphotyrosine residue is positioned slightly deeper into the ligand binding pocket of the mutant SH2 domain. The phenyl ring moves distinctly closer to residue 42 in the mutant simulations, presumably engaging in stabilizing hydrophobic interactions, as suggested for some superbinder mutants (**Figure 5D and Figure S5F**).[45,46] On the other side of the phenyl ring, the phosphotyrosine residue moves further away from His 53, which lines the peptide binding pocket and plays a role in recognizing residues beyond the phosphotyrosine (**Figure 5B,C and Figure S5G**). Overall, the peptide main chain from the -2 to +2 residues appears closer to the body of the SH2 domain (**Figure 5D**). This likely explains why N-SH2^{T42A} prefers smaller residues at the -2 position (Val over Leu/Ile) and +1 position (Ala over Leu/Ile) (**Figure 4**). Consistent with this, the Cα to Cα distance between the peptide +1 residue and Ile 54, which lines the peptide binding pocket, is frequently closer in N-SH2^{T42A} simulations than N-SH2^{WT} simulations (**Figure S5H**).

One of the most dramatic differences between the N-SH2^{WT} and N-SH2^{T42A} simulations in the peptide-bound state is the positioning and movement of Lys 55. In crystal structures and in our N-SH2^{WT} simulations, the Lys 55 ammonium group interacts with the phosphotyrosine phosphoryl group or engages the phenyl ring in a cation-π interaction (**Figure 5B,E and Figure S6A**).[30,40–43] In the N-SH2^{T42A} simulations, the phosphoryl group rotates away from Lys 55 and more tightly engages the BC loop and Arg 32, thereby liberating the Lys 55 side chain (**Figure 5C,E and Figure S6A**). This shift alters the electrostatic surface potential of N-SH2^{T42A} in the peptide binding region when compared to N-SH2^{WT} (**Figure S6B**). In some simulations, the Lys 55 side chain ion pairs with the Asp 40 side chain (**Figure S6C**). For the N-SH2^{T42A} simulations with the PD-1 pTyr 223 peptide, we observed significant sampling of a distinctive state, where the Lys 55 side chain formed an ion pair with the +2 Glu residue (PD-1 Glu 225) (**Figure 5F,G**). This interaction was not observed in the N-SH2^{WT} simulations, and indeed, our peptide display screens showed enhanced preference for a +2 Glu by N-SH2^{T42A} over N-SH2^{WT} (**Figure 4E-G**). Other peptides had Asp and Glu residues at nearby positions, but stable ion pairs between these residues and Lys 55 were not observed in our simulations.

Only three human SH2 domains have an Ala and Lys at the positions that are homologous to residues 42 and 55 in SHP2 N-SH2: the SH2 domains of Vav1, Vav2, and Vav3 (**Figure S4B,C**).[44] Experimental structures of the Vav2 SH2 domain bound to phosphopeptides show that the Lys 55-analogous lysine residue can form electrostatic interactions with acidic residues at various positions on the peptide ligands, including the +2 residue (**Figure S6D**).[48,49] Furthermore, Vav-family SH2 domains are known to prefer a +2 Glu on their ligands,[50] further corroborating a role for SHP2 Lys 55 in substrate selectivity in an T42A-context.

Based on our simulations and the Vav SH2 structures, we hypothesized that Lys 55 plays an important role in the specificity switch caused by the T42A mutation. Thus, we conducted a double-mutant cycle analysis in which we examined the effect of the T42A mutation in the presence and absence of a K55R mutation, measuring binding affinities to five phosphopeptides (**Figure 5H**). Many SH2 domains have an Arg at this position (**Figure S4B**), and the Arg residue forms a cation-π interaction with the pTyr phenyl ring, interacts with the phosphoryl group, or engages a conserved acidic residue at the end of the BC-loop, all of which would likely be tighter than the interactions observed for Lys 55. Indeed, the K55R mutation, on its own, enhanced binding to all five phosphopeptides, with an effect ranging from 2-fold to 7-fold (**Figure 5H**, *gray vs. cyan*). As discussed earlier, the T42A mutation, on its own, enhanced binding anywhere from 4-fold to 90-fold (**Figure 5H**, *gray vs. red*). In the presence of the K55R mutation, however, peptides that showed a large T42A effect now had a diminished impact of the T42A mutation. For example, for the MILR1 peptide, there was a drop in the T42A effect from 90-fold with Lys 55 to 28-fold with Arg 55. By contrast, for peptides where T42A alone had a small effect, this small effect was unchanged by the K55R mutation (**Figure 5H**). These results strongly suggest that Thr 42 and Lys 55 are energetically coupled, and that the peptide-specific effects of the T42A mutation are partly dependent on Lys 55.

We also examined the impact of a K55M mutation. The K55M mutation weakens binding to all peptides by a factor of 6- to 12-fold, indicating an important role for the epsilon ammonium group of Lys 55 in peptide binding (**Figure S6E**). Similar to the K55R mutation, we observed a peptide-

dependent impact of the K55M mutation on the effect of the T42A mutation. The large T42A effect seen for some peptides was substantially diminished in the K55M background. We interpret these double-mutant cycle experiments as follows: The T42A mutation liberates Lys 55 to engage some peptides in new electrostatic interactions, and this is one cause of the T42A-driven change in specificity. With the K55R mutation, this T42A-dependent role for Lys 55 is lost, because the arginine remains tethered to the phosphotyrosine or to Asp 40 on the BC loop. With the K55M mutation, the T42A-dependent electrostatic role for Lys 55 is lost due to removal of the epsilon amine. Overall, the simulations and experiments in these past two sections provide a structural explanation for how the T42A mutation remodels the ligand binding pocket of the N-SH2 domain, resulting in a change in peptide selectivity.

T42A-dependent changes in N-SH2 specificity drive changes in SHP2 activation

Given that N-SH2 engagement is thought to be the main driver of SHP2 activation,[10,37] we hypothesized that the binding specificity changes caused by the T42A mutation would sensitize SHP2 to some activating peptides but not others. To assess enzyme activation, we measured the catalytic activity full-length $SHP2^{WT}$ against the fluorogenic substrate DiFMUP, in the presence of the phosphopeptides used in our binding affinity measurements (**Figure 6A**). $SHP2^{WT}$ activity was enhanced with increasing phosphopeptide concentration, demonstrating ligand-dependent activation (**Figure 6B**). The concentration of phosphopeptide required for half-maximal activation (EC_{50}) was different for each phosphopeptide and correlated well with the binding affinity of the phosphopeptide for the N-SH2 domain (**Figure 6C and Table S5**), substantiating the importance of N-SH2 domain engagement for activation.

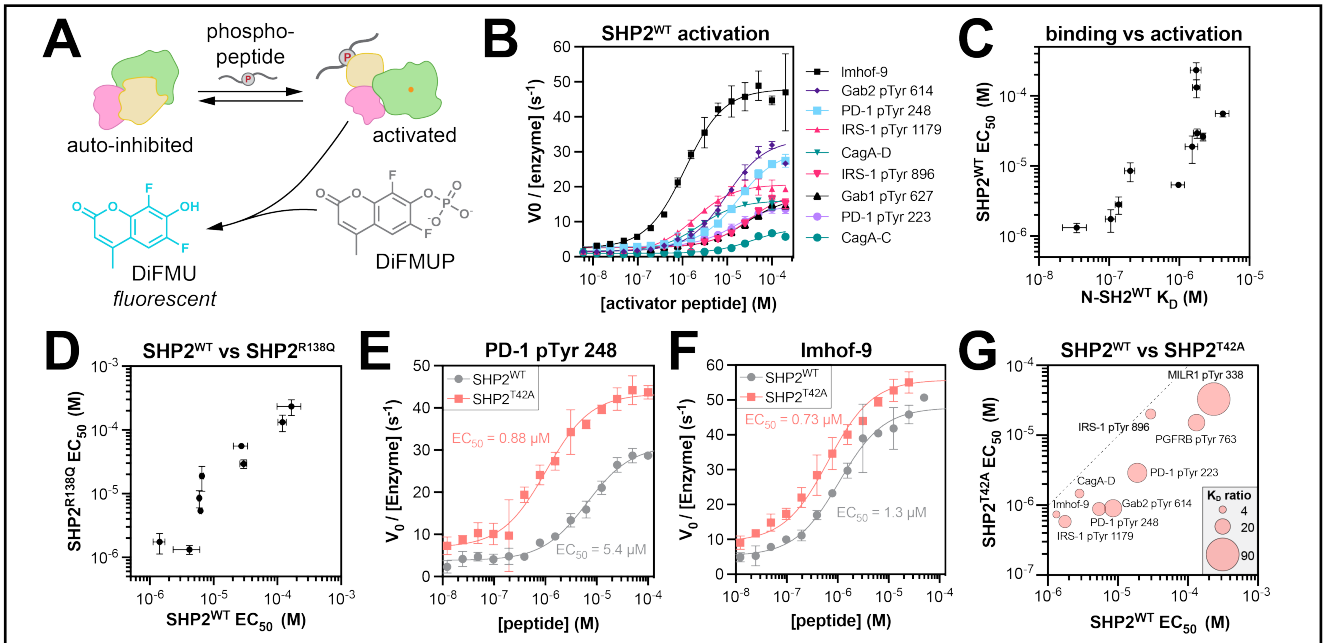

Figure 6. T42A-dependent changes in the activation of full-length SHP2. (**A**) SHP2 activation is measured by incubation with phosphopeptide ligands, followed by monitoring dephosphorylation of the small-molecule substrate DiFMUP to generate fluorescent DiFMU. (**B**) Representative activation curves for $SHP2^{WT}$. (**C**) Correlation between the EC_{50} of $SHP2^{WT}$ activation by phosphopeptides and the K_D of those phosphopeptides for the $N\text{-}SH2^{WT}$ domain. (**D**) Correlation between activation EC_{50} values for $SHP2^{WT}$ and $SHP2^{R138Q}$. (**E**) Comparison of $SHP2^{WT}$ and $SHP2^{T42A}$ activation curves for the PD-1 pTyr 248 peptide. (**F**) Comparison of $SHP2^{WT}$ and $SHP2^{T42A}$ activation curves for the Imhof-9 peptide. (**G**) Bubble plot juxtaposing the EC_{50} values for activation of $SHP2^{WT}$ and $SHP2^{T42A}$ by nine peptides, alongside the fold-change in K_D for binding of those peptides to $N\text{-}SH2^{WT}$ vs $N\text{-}SH2^{T42A}$. The dotted line indicates where EC_{50} for $SHP2^{WT}$ equals EC_{50} for $SHP2^{T42A}$. Peptides with a large fold-change in binding affinity (larger bubble) have a large fold-change in EC_{50} values for $SHP2^{T42A}$ over $SHP2^{WT}$ (distance from dotted line). All EC_{50} values can be found in Table S5.

We also measured EC_{50} values for activation of full-length $SHP2^{R138Q}$, which has negligible C-SH2 binding to phosphopeptides (**Figure S2C**). The EC_{50} values for $SHP2^{WT}$ and $SHP2^{R138Q}$ activation were strongly correlated, further supporting the notion that phosphopeptide binding to the N-SH2 domain, not the C-SH2 domain, is a major driver of SHP2 activation in our experiments (**Figure 6D and Table S5**). We note that some SHP2-binding proteins are bis-phosphorylated, unlike the mono-phosphorylated peptides tested in this work, and in that context, the C-SH2 domain can play a significant role in activating SHP2 by localizing the N-SH2 domain to a binding site for which the N-SH2 otherwise has a weak affinity.[51]

Next, we measured activation of $SHP2^{T42A}$ using the same panel of phosphopeptides. For some peptides, the T42A mutation dramatically shifted this EC_{50} value to lower concentrations, whereas other activation curves were only marginally affected by the mutation (**Figure 6E,F and Table S5**). The peptides that showed a large enhancement in binding affinity to N-$SH2^{T42A}$ over N-$SH2^{WT}$ (**Figure 6G**, *large bubbles*) also showed the largest enhancement in activation (**Figure 6G**, *distance from dotted line*). These results demonstrate that the T42A mutation can sensitize SHP2 to specific activating ligands over others by altering N-SH2 binding affinity and specificity.

Interestingly, we observed that the sequence of the peptide and the mutation status of SHP2 not only impacted the EC_{50}, but also the amplitude of activation (**Figure 6B,E,F**). This variation in amplitude has also been observed by others.[17,30,43] Although the molecular basis for this amplitude effect is currently unknown, a growing body of evidence suggests that SHP2 can adopt multiple distinct active states.[43,52] We speculate that the phosphopeptide sequence can tune SHP2 activation by stabilizing subtly different active-state conformations, which may be further impacted by different SHP2 mutations.[52]

The T42A mutation in SHP2 impacts its cellular interactions and signaling

All of the previous experiments were done using purified proteins and short phosphopeptide ligands. We next sought to determine if the effects of the T42A mutation could be recapitulated in a cellular environment with full-length proteins. First, we assessed the impact of the T42A mutation on phosphoprotein binding through co-immunoprecipitation experiments. We expressed myc-tagged $SHP2^{WT}$ or $SHP2^{T42A}$ in human embryonic kidney (HEK) 293 cells, along with a constitutively active variant of the tyrosine kinase c-Src and a SHP2-interacting protein of interest (**Figure 7A**). We chose Gab1, Gab2, and PD-1 as our proteins of interest, as these proteins play important roles in SHP2-relevant signaling pathways.[25,30,53–56] For all three proteins, we found that $SHP2^{T42A}$ co-immunoprecipitated more with the phosphorylated interacting protein than $SHP2^{WT}$, confirming that the T42A mutation enhances binding to full-length phosphoproteins in cells (**Figure 7B-D**).

Next, we assessed whether the enhancement in SHP2 binding caused by the T42A mutation impacts downstream cell signaling. SHP2 is a positive regulator of Ras GTPases, and it can promote activation of the Ras/MAPK pathway downstream of receptor tyrosine kinases.[57,58] Its functions in this context can be mediated by the adaptor proteins Gab1 and Gab2 (**Figure 7E**).[59] Thus, we transfected $SHP2^{WT}$ or $SHP2^{T42A}$ into HEK 293 cells, along with either Gab1 or Gab2. We then stimulated the cells with epidermal growth factor (EGF) and analyzed Erk phosphorylation as a marker of MAPK pathway activation downstream of the EGF receptor (**Figure 7E**). We observed a stimulation time-dependent increase then decrease in phospho-Erk levels, but the duration of the response was longer for $SHP2^{T42A}$ than for $SHP2^{WT}$ samples (**Figure 7F,G**). Our results are consistent with a previous phospho-proteomics study that examined $SHP2^{WT}$ and $SHP2^{T42A}$ interactors in HeLa cells, which found that the T42A mutation enhances interactions with a variety of growth factor signaling proteins, including those involved in MAPK signaling.[60] Taken together, our cellular experiments demonstrate enhanced binding and signaling in the context of the T42A mutation.

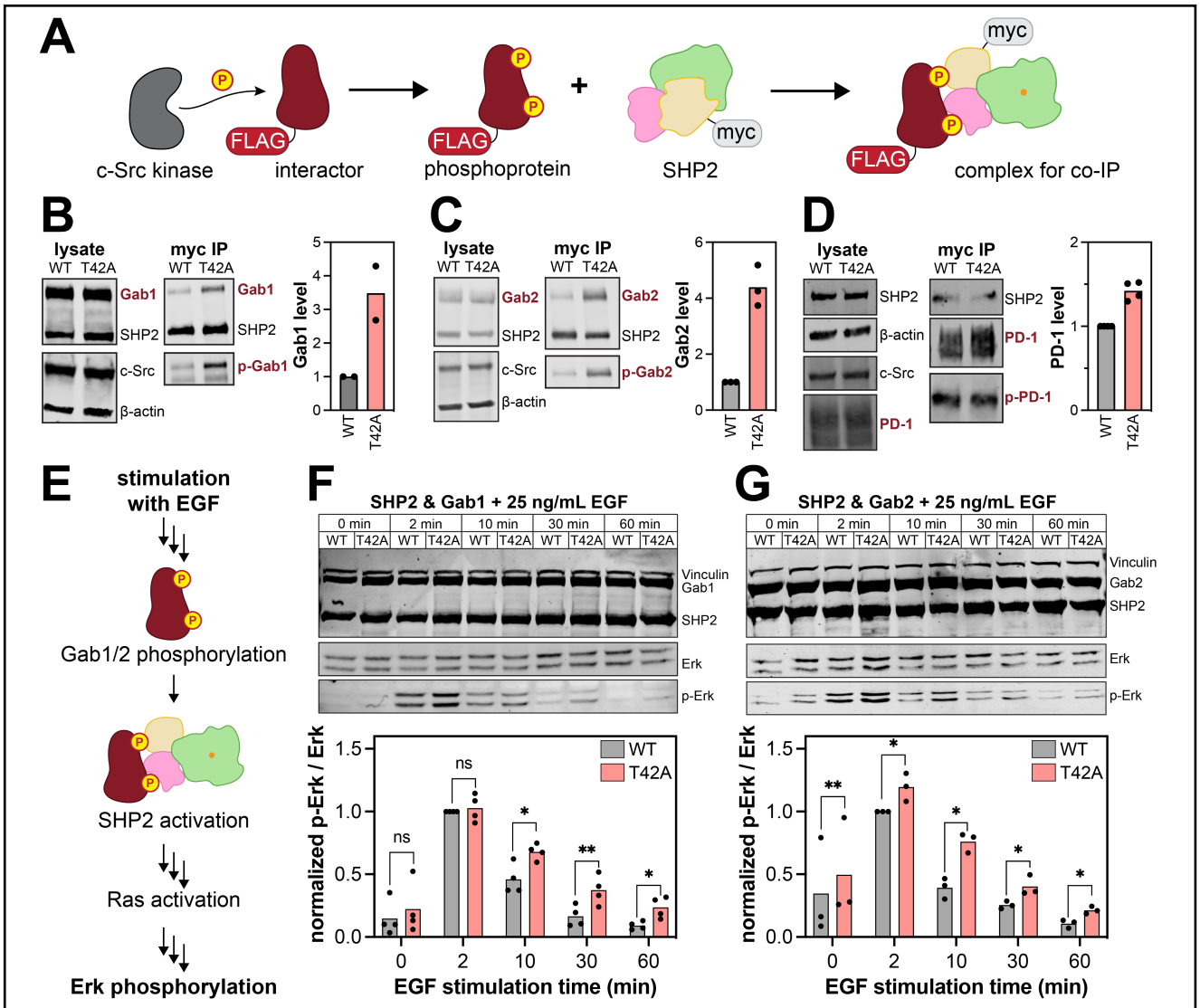

Figure 7. Enhanced cellular interactions and signal transduction by the SHP2 T42A mutation. (**A**) Scheme depicting the co-immunoprecipitation experiments with SHP2 and either Gab1, Gab2, or PD-1 in HEK293 cells. SHP2 co-immunoprecipitation results with (**B**) Gab1, (**C**) Gab2, and (**D**) PD-1. Co-immunoprecipitation of Gab1/Gab2 was detected using an α-FLAG antibody and PD-1 was detected using a PD-1-specific antibody. Co-immunoprecipitation levels of each protein in T42A samples relative to wild-type are normalized for expression level and shown as bar graphs. (**E**) Schematic depiction of EGF stimulation and phospho-Erk signaling experiments in the presence of co-expressed SHP2 and either Gab1 or Gab2. (**F**) Comparison of phospho-Erk levels in response to EGF stimulation in cells expressing Gab1 and either $SHP2^{WT}$ or $SHP2^{T42A}$. (**G**) Comparison of phospho-Erk levels in response to EGF stimulation in cells expressing Gab2 and either $SHP2^{WT}$ or $SHP2^{T42A}$. For panels (F) and (G), the bar graphs below the blots indicate phospho-Erk levels, normalized to total Erk levels, relative to the highest p-Erk signal in $SHP2^{WT}$ time course (2 minutes).

Discussion

The protein tyrosine phosphatase SHP2 is involved in a broad range of signaling pathways and has critical roles in development and cellular homeostasis. The most prevalent and well-studied mutations in SHP2 disrupt auto-inhibitory interactions between the N-SH2 and PTP domain, thereby hyperactivating the enzyme and enhancing downstream signaling. These mutations partly or fully decouple SHP2 activation from SHP2 localization – in this context, SHP2 no longer requires recruitment to phosphoproteins via its SH2 domains for full activation. By contrast, mutations outside of the N-SH2/PTP interdomain interface operate through alternative pathogenic mechanisms, and can have distinct outcomes on cellular signaling.[11] In this study, we have characterized some of the mutations in this category, focusing on mutations in the phosphopeptide binding pockets of the SH2

domains. Our mutations of interest cause a wide range of disease phenotypes: Noonan Syndrome, juvenile myelomonocytic leukemia, acute lymphocytic leukemia, melanoma, and non-syndromic heart defects.[5,11,15,16] Most of these mutations do not have extensive biochemical data quantitatively characterizing their effects on phosphopeptide binding specificity or downstream cell signaling.

Here, we report a sequence-specific enhancement of binding affinity to tyrosine-phosphorylated ligands caused by the T42A mutation in the N-SH2 domain of SHP2. Previous studies have reported increased binding affinity by the T42A mutation, but the change in sequence specificity was not known, nor was the effect of biasing SHP2 activation toward certain ligands.[14,38,60] Our new insights into N-SH2^{T42A} specificity may reflect enhancements in measurement accuracy and library size for our SH2 specificity profiling platform relative to previous methods.[31] Critically, our findings are supported by biochemical and biophysical experiments with a large panel of physiologically relevant peptides and proteins, whereas most studies analyzing SHP2 SH2 mutants have focused on a single peptide at a time. Interestingly, the appearance of +2 Glu and -2 Val preferences in our N-SH2^{T42A} specificity screens were seen before using a peptide microarray approach,[14] but these findings were subtle and not reported as an appreciable change in specificity due to the large overlap in sequence preferences between SHP2^{WT} and SHP2^{T42A}. Notably, one affinity purification mass spectrometry study comparing SHP2 N-SH2^{WT} and N-SH2^{T42A} found mutation-dependent changes in their interaction networks in mammalian cells,[60] further corroborating our biochemical findings. The precise signaling effects of these altered interaction networks have not yet been elucidated.

Much remains unknown about the other mutations explored in this study. To our knowledge, no biochemical or cell biological characterization has been reported for the L43F mutation.[15] Here, we showed a mild increase in basal activity of full-length SHP2^{L43F}, and a slight increase in binding affinity of phosphopeptides to N-SH2^{L43F}. Further studies are needed to elucidate the pathogenic mechanism of this mutant. The T52S mutation has previously been reported to exhibit a change in binding affinity to Gab2, consistent with our data.[61] We did not observe any significant changes in sequence specificity, with the exception of a preference for smaller Ala over bulkier residues at the +1 position. Although we did not conduct molecular dynamics simulations of N-SH2^{T52S}, in our simulations with N-SH2^{WT} and N-SH2^{T42A}, we observed that the methyl group of Thr 52 interacts with the side chain on the +1 residue of the peptide, which could explain how T52S alters the +1 preference.

The R138Q mutation, a rare somatic mutation found in melanoma, severely disrupts phosphopeptide binding to the C-SH2 domain. The molecular basis for the pathogenic effects of the E139D mutation, which occurs in both Noonan Syndrome and JMML, remains elusive. While many studies have addressed its binding affinity and specificity, their results are ambiguous.[14,17,60] Our high-throughput screens indicate a weakening of binding for some peptides to C-SH2^{E139D} when compared to C-SH2^{WT}. Consistent with previous work, our basal activity measurements with full-length SHP2 show that the E139D mutation is activating, suggesting an undefined regulatory role for the C-SH2 domain that may be unrelated to ligand binding.[17,20] In the auto-inhibited state of SHP2, the C-SH2 domain does not make extensive contacts with the PTP domain, and it is unclear if ligand binding to the C-SH2 domain alone can allosterically activate SHP2.[10] The prevailing idea is that the C-SH2 domain binds to phosphoproteins, thereby localizing SHP2 to a nearby phosphotyrosine residue that can allosterically activate SHP2 through engagement of the N-SH2 domain.[30] Thus, the full function of the C-SH2 domain in SHP2 regulation remains to be uncovered.

In this report, we have demonstrated functional and cellular consequences of the T42A mutation in SHP2. Specifically, this mutation causes SHP2 to bind tighter to certain phosphoproteins and is more strongly activated as a result. This enhanced binding and activation translates to downstream effects, such as increased MAPK signaling. What remains unknown is how the biased interaction specificity of the T42A mutant impacts SHP2 signaling when compared with other mutations that simply alter binding affinity but not specificity. Our findings suggest that SHP2^{T42A} will be hyperresponsive to certain upstream signals (e.g. phosphorylated Gab1 and Gab2), but not to others. Further studies will be needed to more broadly assess T42A-induced changes in cell signaling, in order to fully understand the pathogenic mechanism of this variant.

Most disease-associated mutations that alter the functions of cell signaling proteins do so by disrupting their intrinsic regulatory capabilities – for SHP2, pathogenic mutations cluster at the auto-inhibitory interface between the N-SH2 and PTP domains and hyperactivate the enzyme by disrupting interdomain interactions. There is increasing evidence that mutations can also rewire signaling

pathways by changing protein-protein interaction specificity.[62–64] This has been demonstrated most clearly for protein kinases, where mutations have been identified that alter substrate specificity.[62] For protein kinases, it is noteworthy that not all specificity-determining residues are located directly in the ligand/substrate binding pocket, raising the possibility that distal mutations may allosterically alter sequence specificity.[65,66] A similar paradigm has been suggested for SH2 domains, where distal mutations may rewire interaction specificity, however the position corresponding to Thr 42 in SHP2 has not been implicated as a determinant of specificity.[65] The biochemical and structural analyses presented in this paper reveal an unexpected outcome of the pathogenic T42A mutation, where ligand selectivity is altered over 10 Å from the mutation site. Our results highlight the importance of considering the structural plasticity of signaling proteins when evaluating specificity and suggest that the functional consequences of many disease-associated mutations could be misclassified if evaluated solely based on their locations in static protein structures.

Materials and Methods

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Neel H. Shah (neel.shah@columbia.edu).

Materials availability

Plasmids and DNA libraries generated in this study will made freely available upon request. There are no restrictions to the availability of reagents generated in this study.

Data and code availability

All of the processed data, including catalytic efficiencies, binding affinities, EC_{50} values, and enrichment scores from the high-throughput specificity screens are provided as supplementary table files. The raw FASTQ sequencing files from specificity screens, source data from MD simulations, and processed MD trajectories are available as of the date of publication as a Dryad repository (DOI: 10.5061/dryad.msbcc2g41). All original code has been deposited at Github and is publicly available as of the date of publication (https://github.com/nshahlab/2022_Li-et-al_peptide-display). Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Experimental model and study participant details

Human embryonic kidney (HEK) 293 cells (female) were a gift from Brent Stockwell's Lab and were not further authenticated. HEK 293 were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS) and 1% penicillin/streptomycin at 37 °C with 5% CO_2. HEK 293 cells were tested every 4 months for mycoplasmal infection. Chemically competent C43(DE3) cells and Electrocompetent MC1061 cells were purchased from Lucigen (LCG Biosearch Technologies) and grown in LB medium. BL21(DE3) cells were purchased from ThermoFisher Scientific and grown in LB medium.

Method details

Purification of SH2 domains

The SHP2 full-length, wild-type gene used as the template for all SHP2 constructs in this study was cloned from the pGEX-4TI SHP2 WT plasmid, which was a generous gift from Ben Neel (Addgene plasmid #8322).[67] SHP2 SH2 domains were cloned into a His_6-SUMO-SH2-Avi construct.[31] C43(DE3) cells were transformed with plasmids encoding both the respective SH2 domain and the biotin ligase BirA. Cells were grown in LB supplemented with 50 µg/mL kanamycin and 100 µg/mL streptomycin at 37 °C until cells reached an optical density at 600 nm (OD_{600}) of 0.5. IPTG (1 mM) and biotin (250 µM) were added to induce protein expression and ensure biotinylation of SH2 domains, respectively. Protein expression was carried out at 18 °C overnight. Cells were centrifuged and subsequently resuspended in lysis buffer (50 mM Tris pH 7.5, 300 mM NaCl, 20 mM imidazole, 10% glycerol, and freshly added 2 mM β-mercaptoethanol). The cells were lysed using sonication (Fisherbrand Sonic Dismembrator), and spun down at 14,000 rpm for 45 minutes. The supernatant was applied to a 5 mL Ni-NTA column (Cytiva). The resin was washed with 10 column volumes lysis buffer and wash buffer (50 mM Tris pH 7.5, 50 mM NaCl, 20 mM imidazole, 10% glycerol, and freshly added 2 mM β-mercaptoethanol). The protein was eluted off the Ni-NTA column in elution buffer (50 mM Tris pH 7.5, 50 mM NaCl, 500 mM imidazole, 10% glycerol) and brought onto a 5mL HiTrap Q Anion exchange column (Cytiva). The column was washed using Anion A buffer (50 mM Tris pH 7.5, 50 mM NaCl, 1 mM TCEP). Protein elution off the column was induced through a salt gradient between Anion A buffer and Anion B buffer (50 mM Tris pH 7.5, 1 M NaCl, 1 mM TCEP). The eluted protein was cleaved at the His_6-SUMO tag by addition of 0.05mg/mL His_6-tagged Ulp1 protease at 4ºC overnight. This cleavage cocktail was flowed through a 2 mL Ni-NTA gravity column (ThermoFisher) to isolate the cleaved protein away from uncleaved protein and Ulp1. Finally, the

cleaved protein was purified by size-exclusion chromatography on a Superdex 75 16/600 gel filtration column (Cytiva) equilibrated with SEC buffer (20 mM HEPES pH 7.4, 150 mM NaCl, and 10% glycerol). Pure fractions were pooled and concentrated, and flash frozen in liquid N_2 for long-term storage at −80 °C.

Purification of full-length SHP2 proteins

Full-length SHP2 variants were cloned into a pET28-His-TEV plasmid from the pGEX-4TI SHP2 WT plasmid.[67] BL21(DE3) cells were transformed with the respective plasmids, and were grown in LB supplemented with 100 µg/mL kanamycin at 37 °C until cells reached an OD_{600} of 0.5. IPTG (1 mM) was added to induce protein expression, which was carried out at 18 °C overnight. Cells were centrifuged and subsequently resuspended in lysis buffer (50 mM Tris pH 7.5, 300 mM NaCl, 20 mM imidazole, 10% glycerol, and freshly added 2 mM β-mercaptoethanol). The cells were lysed using sonication (Fisherbrand Sonic Dismembrator), and spun down at 14,000 rpm for 45 minutes. The supernatant was applied to a 5 mL Ni-NTA column (Cytiva). The resin was washed with 10 column volumes lysis buffer and wash buffer (50 mM Tris pH 7.5, 50 mM NaCl, 20 mM imidazole, 10% glycerol, and freshly added 2 mM B-MERCAPTOETHANOL). The protein was eluted off the Ni-NTA column in elution buffer (50 mM Tris pH 7.5, 50 mM NaCl, 500 mM imidazole, 10% glycerol) and brought onto a 5mL HiTrap Q Anion exchange column (Cytiva). The column was washed using Anion A buffer (50 mM Tris pH 7.5, 50 mM NaCl, 1 mM TCEP). Protein elution off the column was induced through a salt gradient between Anion A buffer and Anion B buffer (50 mM Tris pH 7.5, 1 M NaCl, 1 mM TCEP). The eluted protein was cleaved at the His_6-TEV tag by addition of 0.10 mg/mL of His_6-tagged TEV protease at 4 °C overnight. This cleavage cocktail was flowed through a 2 mL Ni-NTA gravity column (ThermoFisher) to separate the cleaved protein from uncleaved protein and TEV protease. Finally, the cleaved protein was purified by size-exclusion chromatography on a Superdex 200 16/600 gel filtration column (Cytiva) equilibrated with SEC buffer (20 mM HEPES pH 7.5, 150 mM NaCl, and 10% glycerol). Pure fractions were pooled and concentrated, and flash frozen in liquid N_2 for long-term storage at −80 °C.

Synthesis and purification of peptides

Several of the peptides used in this study were purchased from a commercial vendor (SynPeptide). The remaining peptides used for *in vitro* kinetic and binding assays were synthesized using 9-fluorenylmethoxycarbonyl (Fmoc) solid-phase peptide chemistry. All syntheses were carried out using the Liberty Blue automated microwave-assisted peptide synthesizer from CEM under nitrogen atmosphere, with standard manufacturer-recommended protocols. Peptides were synthesized on MBHA Rink amide resin solid support (0.1 mmol scale). Each Nα-Fmoc amino acid (6 eq, 0.2 M) was activated with diisopropylcarbodiimide (DIC, 1.0 M) and ethyl cyano(hydroxyamino)acetate (Oxyma Pure, 1.0 M) in dimethylformamide (DMF) prior to coupling. The coupling cycles for phosphotyrosine and the amino acid directly after it were done at 75 °C for 15 s, then 90 °C for 230 s. All other coupling cycles were done at 75 °C for 15 s, then 90 °C for 110 s. Deprotection of the Fmoc group was performed in 20% (v/v) piperidine in DMF (75 °C for 15 s then 90 °C for 50 s), except for the amino acid directly after the phosphotyrosine which had an additional initial deprotection (25 °C for 300 s). The resin was washed (4x) with DMF following Fmoc deprotection and after Nα-Fmoc amino acid coupling. All peptides were acetylated at their N-terminus with 10% (v/v) acetic anhydride in DMF and washed (4x) with DMF.

After peptide synthesis was completed, including N-terminal acetylation, the resin was washed (3x each) with dichloromethane (DCM) and methanol (MeOH), and dried under reduced pressure overnight. The peptides were cleaved and the side chain protecting groups were simultaneously deprotected in 95% (v/v) trifluoroacetic acid (TFA), 2.5% (v/v) triisopropylsilane (TIPS), and 2.5% water, in a ratio of 10 µL cleavage cocktail per mg of resin. The cleavage-resin mixture was incubated at room temperature for 90 minutes, with agitation. The cleaved peptides were precipitated in cold diethyl ether, washed in ether, pelleted, and dried under air. The peptides were redissolved in a 50% (v/v) water/acetonitrile solution and filtered from the resin.

The crude peptide mixture was purified using reverse-phase high performance liquid chromatography (RP-HPLC) on either a semi-preparatory C18 column (Agilent, ZORBAX 300SB-C18, 9.4 x 250 mm, 5 µm) with an Agilent HPLC system (1260 Infinity II), or a preparatory C18 column

(XBridge Peptide BEH C18 Prep Column, 19 x 150 mm, 5 µm) with a Waters prep-HPLC system (Prep 150 LC System). Flow rate for purification was kept at 4 mL/min (semi-preparative) or 17mL/min (preparative) with solvents A (water, 0.1% (v/v) TFA) and B (acetonitrile, 0.1% (v/v) TFA). Peptides were generally purified over a 40 minute (semi-preparative) or 13 minute (preparative) linear gradient from solvent A to solvent B, with the specific gradient depending on the peptide sample. Peptide purity was assessed with an analytical column (Agilent, ZORBAX 300SB-C18, 4.6 x 150 mm, 5 µm) at a flow rate of 1 mL/min over a 0-70% B gradient in 30 minutes. All peptides were determined to be ≥95% pure by peak integration. The identities of the peptides were confirmed by mass spectroscopy (Waters Xevo G2-XS QTOF). Pure peptides were lyophilized and redissolved in 100 mM Tris, pH 8.0, as needed for experiments.

Synthesis and purification of fluorescent peptides

The fluorescent peptides were prepared as described above for the unlabeled peptides, except for the coupling of aminohexanoic acid (AHX) and the fluorescein isothiocyanate (FITC) at the N-terminus, instead of an N-terminal acetyl. AHX (6 eq, 0.2 M) was activated with diisopropylcarbodiimide (DIC, 1.0 M) and ethyl cyano(hydroxyamino)acetate (Oxyma Pure, 1.0 M) in dimethylformamide (DMF) prior to coupling. The coupling cycle was done at 75 °C for 35 s then 90 °C for 575 s. Deprotection was performed as described previously.

After peptide synthesis was completed, including N-terminal AHX-labeling, the resin was washed (3x each) with dichloromethane (DCM) and methanol (MeOH) and dried under reduced pressure overnight. Then, a portion of the resin (0.025 mmol) was prepared for FITC labeling by swelling in DMF with agitation for 30 min. Excess DMF was removed and to the resin was added FITC (0.075 mmol, 3 eq.) and DIPEA (0.15mmol, 6 eq.) in DMF. This reaction was incubated at room temperature with agitation for 2 hours. After FITC labeling, the resin was washed (3x each) with dichloromethane (DCM) and methanol (MeOH) and dried under reduced pressure overnight. Finally, the peptides were cleaved and purified as previously described.

Basal SHP2 catalytic activity measurements

Initial rate measurements for the SHP2-catalyzed dephosphorylation of 6,8-difluoro-4-methylumbelliferyl phosphate (DiFMUP) were conducted at 37 ºC in DiFMUP buffer (60 mM HEPES pH 7.2, 150 mM NaCl, 1mM EDTA, 0.05% Tween-20). Reactions of 50 µL were set up in a black polystyrene flat bottom half area 96-well plate. A substrate concentration series of 31.25 µM, 62.5 µM, 125 µM, 250 µM, 500 µM, 1000 µM, 2000 µM and 4000 µM was used to determine k_{cat} and K_M. Reactions were started by addition of appropriate amount of SHP2 wild-type and mutants (wild-type: 2.5 nM; T42A: 2.5 nM; L43F: 2.5 nM; T52S: 4 nM; E76K: 0.5 nM; R138Q: 3 nM, E139D: 2.5 nM). Emitted fluorescence at 455nm was recorded every 25 seconds in a span of 50 minutes using a BioTek Synergy Neo2 multi-mode reader.

Initial rate measurement for the SHP2-catalyzed dephosphorylation of *p*-nitrophenyl phosphate (pNPP) were conducted at 37 ºC in pNPP buffer (10 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM TCEP). Reactions of 75uL were set up in a black polystyrene flat bottom half area 96-well plate. A substrate concentration series of 12.8, 6.4, 3.2, 1.6, 0.8, 0.4, 0.2 and 0.1mM pNPP was used to determine K_{cat} and K_M. Reactions were started by addition of appropriate amount of SHP2 wild-type and mutants (wild-type: 250 nM;T42A: 250 nM; L43F: 250 nM; T52S: 250 nM; E76K: 100 nM; R138Q: 250 nM; E139D: 250 nM).

In all cases, the linear region of the reaction progress curve was determined by visual inspection and fit to a line. These slopes were converted from absorbance or fluorescence units as a function of time to product formation as a function of time using standard curves measured with the reaction products (*p*-nitrophenol and 6,8-difluoro-7-hydroxy-4-methylcoumarin). Finally, these rates were corrected for enzyme concentration by dividing the values the concentration of enzyme used in the experiment to yield V_0 / [enzyme] in units of (s^{-1}). These corrected rates were plotted as a function of substrate concentration and fit to the Michaelis-Menten equation using non-linear regression to determine catalytic parameters. Experiments were generally repeated at least three times, and the average and standard deviation of all individual replicates are reported.

Fluorescence polarization binding assays

SH2 domains were thawed in room temperature water and their absorbance at 280 nm was measured to determine concentration. SH2 domains were serially diluted 15 times in assay buffer (60 mM HEPES pH 7.2, 75 mM KCl, 75 mM NaCl,1 mM EDTA, 0.05% Tween-20), with a 2x starting concentrations generally in the low micromolar range. One well did not contain any SH2 domain. The fluorescent peptide was diluted to 2x the desired concentration and mixed in 1:1 ratio with the different concentrations of SH2 domain (specific fluorescent peptide concentrations can be found in Table S2). The mixture was transferred to a black 96-well plate and incubated for 15 minutes at room temperature. Parallel and perpendicular measurements were taken using the 485/30 polarization cube on the BioTek Neo2 Plate Reader. Data was analyzed and fitted to a quadratic binding equation to determine the K_D for the fluorescent peptide, according to previously established methods.[31,68] Next, a peptide of interest was serially diluted 15 times in assay buffer, with the highest concentration being in the high micromolar range (e.g. 400 μM, for a final concentration of 200 μM). In parallel, a fluorescent peptide was mixed with SH2 domain in assay buffer at 2x the desired final concentration (see Table S2). The fluorescent peptide/SH2 mixture was mixed 1:1 with the diluted peptides in a black 96-well plate and incubated for 15 minutes at room temperature. Fluorescence polarization was measured as previously described for initial K_D measurements. Competition binding data were fit to a cubic binding equation as described previously.[31,68]

Cloning of the PD-1 ITIM scanning mutagenesis library

A scanning mutagenesis library derived from PD-1 residues 218 to 228 was cloned as described previously for other peptide display libraries in the eCPX system.[31,69] A series of oligonucleotides spanning the peptide-coding sequence was synthesized with a different NNS codon in place of each wild-type codon (11 oligonucleotides in total). These degenerate primers were pooled then used to amplify a library of linear DNA encoding mutant ITIM fusions to the eCPX scaffold protein. This library was then cloned into the pBAD33 vector used for surface display.

SH2 specificity profiling using bacterial peptide display

Electrocompetent MC1061 cells were transformed with ~100 ng of the respective library. After 1 hour recovery in 1 mL LB, cells were further diluted into 250 mL LB + 0.1% chloramphenicol. 1.8 mL of overnight culture was used to inoculate 100 mL LB + 0.1% chloramphenicol, and grown until OD_{600} reached 0.5. 20 mL of cell suspension was induced at 25 °C using a final concentration of 0.4% arabinose until the OD_{600} ~ 1 (after approximately 4 hours). The cells were spun down at 4000 rpm for 15 minutes, and the pellet was resuspended in PBS so that the OD_{600} ~1.5. The cells were stored in the fridge and used within a week.

For each sample, 75 μL of Dynabeads™ FlowComp™ Flexi Kit were washed twice in 1 mL SH2 buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 1 mM TCEP, and 0.2% BSA) on a magnetic rack. The beads were then resuspended in 75 μL of SH2 buffer. SH2 domains were thawed quickly and 20 μM of protein was added to the beads. SH2 buffer was added up to 150 μL, and the suspension was incubated for 1 hour at 4 °C while rotating. After 1 hour, the suspension was placed on a magnetic rack and washed twice with 1 mL SH2 buffer.

150 μL of prepared cells per sample were spun down for 4000 rpm for 5 minutes at 4 °C. Kinase screen buffer was prepared (50 mM Tris, 10 mM magnesium chloride, 150 mM sodium chloride; add 2 mM sodium orthovanadate and 1 mM TCEP fresh) and the cells of each sample were resuspended in 100 μL kinase screen buffer. Kinases c-Src, c-Abl, AncSZ, Eph1B were added to a final concentration of 2.5 μM each, creatine phosphate was added to a final concentration of 5 mM, and phosphokinase was added to a final concentration of 50 μg/mL. The suspension was incubated at 37 °C for 5 minutes before ATP was added to a final concentration of 1 mM. This mixture was incubated at 37 °C for 3 hours. After 3 hours, EDTA was added to a final concentration of 25 mM to quench the reaction. The input library control sample was not phosphorylated. These cells were spun down at 4000 rpm for 15 minutes at 4 °C. The cells were then resuspended in 100 μL of SH2 buffer + 0.1% BSA. Phosphorylation of the cells was confirmed by labeling with the PY20-PerCP-eFluor 710 pan-phosphotyrosine antibody followed by analysis via flow cytometry.

100 μL of phosphorylated cells were mixed with 75 μL SH2-beads for 1 hour at 4 °C while

rotating. After 1 hour, samples were placed on a magnetic rack, supernatant was removed and 1 mL SH2 buffer was added to each sample. This was rotated for 30 minutes at 4 °C to wash the beads. After this wash, the beads were placed on a magnetic rack, the supernatant was removed and 50 µL MilliQ was added.

All SH2-selected samples and the input library control were resuspended in 50 µL MilliQ water, vortexed, and boiled for 10 minutes at 100 °C. The boiled lysate was used as the DNA template in a PCR reaction using the TruSeq-eCPX-Fwd and TruSeq-eCPX-Rev primers. The mixture resulting from this PCR was used directly into a second PCR to append Illumina sequencing adaptors and unique 5' and 3' indices to each sample (D700 and D500 series primers). The resulting PCR mixtures were run on a gel, the band of the expected size was extracted and purified, and its concentration was determined using QuantiFluor® dsDNA System (Promega). Samples were pooled at equal molar ratios and sequenced by paired-end Illumina sequencing on a MiSeq or NextSeq instrument using a 150 cycle kit. The number of samples per run, and the loading density on the sequencing chip, were adjusted to obtain at least 1-2 million reads for each index/sample.

Deep sequencing data were processed and analyzed as described previously.[31,69] First, paired-end reads were merged using FLASH.[70] Then, adapter sequences and any constant regions of the library flanking the variable peptide-coding region were removed using Cutadapt.[71] Finally, these trimmed files were analyze using in-house Python scripts in order to count the abundance of each peptide in the library, as described previously (https://github.com/nshahlab/2022_Li-et-al_peptide-display).[31] The resulting raw counts ($n_{peptide}$) were normalized to the total number of reads in the sample (n_{total}) to yield a frequency ($f_{peptide}$) for each peptide in the library (equation 1). The enrichment score for each peptide ($E_{peptide}$) was calculated by taking the ratio of the frequency of that peptide in the enriched sample versus an input (unenriched) sample (equation 2). In the case of the PD-1 ITIM scanning mutagenesis libraries, we report the log_2-transformed enrichment of a variant normalized to that for the wild-type ITIM sequence (equation 3).

(1) $f_{peptide} = \frac{n_{peptide}}{n_{total}}$ (2) $E_{peptide} = \frac{f_{peptide,enriched}}{f_{peptide,input}}$ (3) $\Delta E_{variant} = log_2 \frac{E_{variant}}{E_{wild-type}}$

Cell experiment reagents

The SHP2 gene was cloned from the pGEX-4TI SHP2 WT plasmid from Ben Neel (Addgene plasmid #8322).[67] The PD-1 gene was cloned from the PD-1-miSFIT-4x plasmid, which was a gift from Tudor Fulga (Addgene plasmid #124678).[72] The mouse Gab1 and Gab2 genes were cloned from the FRB-GFP-Gab2(Y604F/Y634F) and FRB-GFP-Gab1(Y628F/Y660F) plasmids, which were a gift from Andrei Karginov (Addgene plasmid #188658 and #188659).[73] The mouse c-Src gene was expressed from the pCMV5 mouse Src plasmid, a generous gift from Joan Brugge and Peter Howley (Addgene plasmid #13663). The C-terminal regulatory tail (residues 528-535) was deleted to generate hyperactive c-Src. For transient transfection, genes of interest were cloned into a pEF vector (a gift from the Arthur Weiss lab).

Co-immunoprecipitation experiments

2.2 x 10^6 HEK 293 cells were seeded in a 10 cm plate. The next day, cells were transfected using 5 µg of each plasmid (SHP2, Src, interacting protein of interest), and 45 µg PEI in 1.5 mL DMEM. The transfection medium was refreshed after 16 hours and replaced with complete medium. 48 hours later, the cells were harvested by scraping in PBS. Cells were washed 3 times in PBS, and lysed in 500 µL lysis buffer (20 mM Tris-HCl, pH 8.0, 137 mM NaCl, 2 mM EDTA, 10% glycerol, and 0.5% NP-40 + protease inhibitors + phosphatase inhibitors) for 30 minutes while rotating at 4 °C. Cells were spun at 17.7 rpm for 15 minutes at 4 °C. Supernatant was transferred to a clean Eppendorf tube and stored at -20 °C.

Protein concentration was determined using a bicinchoninic acid (BCA) assay and absorbance was measured at 562 nm using a BioTek Synergy Neo2 multi-mode reader. 300 µg of protein was incubated overnight with 30 µL of magnetic anti-myc beads in a total volume of 380 µL while rotating at 4 °C. The next day, the beads were washed 3 times on a magnetic rack using 1 mL lysis buffer. Then, 1x Laemmli buffer was added and beads were boiled at 100ºC for 8 minutes. For whole cell lysates, 15 µg protein was loaded onto a gel. For IP samples, 15 µL of boiled supernatant was used.

Gel was transferred onto a nitrocellulose membrane using TurboBlot (BioRad) and the membrane was blocked using 5% bovine serum albumin (BSA) in Tris-buffered saline (TBS) for 1 hour at room temperature. Membranes were rinsed with TBS with 0.1% Tween-20 (TBS-T) and incubated with primary antibodies in TBST + 5% BSA overnight at 4ºC (Src 1:1000, β-actin 1:5000, Myc 1:5000, FLAG 1:5000, pTyr 1:2000). Membranes were washed and incubated with secondary antibodies (IRDye Goat anti-Rabbit 680 and IRDye Goat anti-Mouse 800). Blots were imaged on a LiCor Odyssey.

Molecular dynamics simulations

We built and simulated several systems comprising of the SHP2 N-terminal SH2 domains bound to different ligands, as well as the SH2 domain in the apo state. For most of the systems, the SH2 domain was taken from the crystal structure 6ROY.[30] This structure, without a ligand bound, was used as the starting structure for simulations of the SH2 domain in the apo state. For simulations of the SH2 domain bound to PD-1 pTyr 223, the ligand was taken from the PDB structure of 6ROY, and the missing residues -6 (Pro), -5 (Val), -4 (Phe) and -3 (Ser) were built in using PyMOL.[74] For simulations of the SH2 domain bound to IRS-1 pTyr 896, the structures of the SH2 domain and the bound ligand were taken from the PDB structure of 1AYB.[40] Here the N-terminal residues 1-4 of the SH2 domain were mutated to MTSR (from MRRW) to be consistent with the sequence of the SH2 domain in the rest of the simulations. For the same reason, Cys was built in at position 104 at the C-terminal end. Residues -6 (Phe), -5 (Lys), -4 (Ser), and -3 (Pro) in the ligand were missing in the crystal structure and were built in using PyMOL. For simulations of the SH2 domain bound to Gab2 pTyr 614, the structure of the SH2 domain and ligand were taken from the PDB structure of 6ROY. The starting structure of the SH2 domain was built in the same way as that used in the simulations of the SH2 domain bound to PD-1. The ligand was built by mutating residues -6, -5, -4, +1, and +2 of the ligand in simulations of PD-1 bound to the SH2 domain to Ser, Thr, Gly, Leu, and Ala, respectively. For the simulations of the SH2 domain bound to Imhof-9, the PDB structure 3TL0 was used.[41] For the starting structure of the SH2 domain, missing residues 1-4 (MTSR) at the N-terminal end and residue 104 (Cys) at the C-terminal end were built in using PyMOL. For the starting structure of the ligand, missing residues -6 to -3 (KKAA) were built in, residue +4 was mutated from Tyr to Leu and missing residues +4 to +6 (MFP) were built in using PyMOL. For the simulations of the SH2 domain bound to MILR1 pTyr 338, the crystal structure from 6ROY was used. The SH2 domain was built in the same manner as in simulations of the SH2 domain bound to PD-1. For the ligand, residues -6 to -1 of the ligand in 6ROY were mutated to AKSGAV (from PVFSVD), residue +1 was mutated from Gly to Ser, residue +4 was mutated from Asp to Asn, and residue +6 was mutated from Gln to Gly. For each of these systems, a similar system with Thr 42 in the SH2 domain mutated to alanine was built using PyMOL. Crystalline waters from 6ROY were used in all the simulations. In all the systems, the N-terminal and C-terminal ends were capped with acetyl and amide groups respectively, in both the SH2 domains and the ligands.

All the systems were solvated with TIP3P water[75] and ions were added such that the final ionic strength of the system was 100 mM using the tleap package in AmberTools21.[76] The energy of each system was minimized first for 5000 steps while holding the protein chains and crystalline waters fixed, followed by minimization for 5000 steps while allowing all the atoms to move. For each system, three individual trajectories were generated by reinitializing the velocities at the start of the heating stage described below.

The temperature of each system was raised in two stages – first to 100 K over 1 ns and then to 300 K over 1 ns. The protein chains and crystalline waters were held fixed during this heating stage. Each system was then equilibrated for 2 ns, followed by production runs. Three production trajectories, each 1 μs long, were generated for each system. All equilibration runs and production runs were performed at constant temperature (300 K) and pressure (1 bar).

The simulations were carried out with the Amber package[77] using the ff14SB force field for proteins[78] using an integration timestep of 2 fs. The Particle Mesh Ewald approximation was used to calculate long-range electrostatic energies.[79] All hydrogens bonded to heavy atoms were constrained with the SHAKE algorithm.[80] The Langevin thermostat was used to control the temperature with a collision frequency of 1 ps^{-1}. Pressure was controlled while maintaining periodic boundary conditions.

Key measurements were extracted from the MD trajectories using the CPPTRAJ module of

AmberTools22.[81] For the RMSF calculations, the trajectories were sampled every 100 ps and RMSF values were calculated from the Cα atoms of each residue after determining root mean squared deviation relative to the first state in the production run of the simulation. For RMSF calculations of each system, each trajectory was analyzed separately, as seen in Figure S5A. For distance calculations, trajectories were sampled every 1 ns. The distance measurements from all three replicates of each system were combined to determine the distance distributions seen in Figure 5, Figure S5, and Figure S6. In cases where distance calculations involved a redundant atom (e.g. distances to one of the three non-bridging oxygen atoms in the phosphoryl group of phosphotyrosine), all three distance measurements were calculated, then the shortest distance at each frame was determined and used for the distribution plots. For visualization, trajectories were sampled every 10 ns. All structure visualization and rendering in this study was done using PyMOL.[74]

Quantification and statistical analysis

Quantification of Western blots was performed in Image Studio Lite, using the Median Background setting. Parametric, paired, one-tailed t-tests were performed in GraphPad Prism.

Key Resource Table

Key resources, including cell lines, plasmids, oligonucleotide primers, peptides, and proteins, are listed in **Table S6**.

Acknowledgements

We would like to thank the members of the Shah lab for their scientific insights and helpful discussions. We thank Jeanine Amacher and Marko Jovanovic for their constructive feedback on this manuscript. This research was funded by NIH grant R35GM138014 to NHS. CAC is supported by an NSF Graduate Research Fellowship (award # 2036197). This work used the Expanse GPU cluster at the San Diego Supercomputer Center through allocation BIO220139 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by an NSF grants #2138259, #2138286, #2138307, #2137603, and #2138296.

Supplementary Information

All supplementary figures can be found in a single supplementary figures file. Supplementary tables can be found as individual spreadsheet files.

Figure S1. Basal activity measurements of various SHP2 mutants.

Figure S2. Binding affinities for SH2 domains.

Figure S3. Analysis of sequence-features for N-SH2 mutants.

Figure S4. Amino acid residues at key positions in human SH2 domains.

Figure S5. Measured structural parameters from MD simulations of SHP2 N-SH2^{WT} or N-SH2^{T42A}.

Figure S6. The role of Lys 55 on T42A-dependent peptide recognition.

Table S1. Catalytic efficiencies of full-length SHP2 variants against small molecule substrates pNPP and DiFMUP.

Table S2. Binding affinities of all SH2-peptide pairs measured by direct or competition fluorescence polarization experiments.

Table S3. Enrichment scores from peptide display screens for all SH2 domains against the Human pTyr and pTyr-Var Libraries.

Table S4. Enrichment scores for N-SH2^{WT} and N-SH2^{T42A} from the PD-1 pTyr 223 scanning mutagenesis peptide display screens.

Table S5. EC_{50} values for activation of SHP2^{WT}, SHP2^{T42A}, and SHP2^{R138Q} by various peptides.

Table S6. Key resources table.

References

1. Salmond, R.J., and Alexander, D.R. (2006). SHP2 forecast for the immune system: fog gradually clearing. Trends Immunol. *27*, 154–160. 10.1016/j.it.2006.01.007.
2. Lauriol, J., Jaffré, F., and Kontaridis, M.I. (2015). The role of the protein tyrosine phosphatase SHP2 in cardiac development and disease. Semin. Cell Dev. Biol. *37*, 73–81. 10.1016/j.semcdb.2014.09.013.
3. Asmamaw, M.D., Shi, X.-J., Zhang, L.-R., and Liu, H.-M. (2022). A comprehensive review of SHP2 and its role in cancer. Cell Oncol (Dordr) *45*, 729–753. 10.1007/s13402-022-00698-1.
4. Grossmann, K.S., Rosário, M., Birchmeier, C., and Birchmeier, W. (2010). The tyrosine phosphatase shp2 in development and cancer. In Advances in cancer research (Advances in cancer research), pp. 53–89. 10.1016/S0065-230X(10)06002-1.
5. Tartaglia, M., and Gelb, B.D. (2005). Noonan syndrome and related disorders: genetics and pathogenesis. Annu. Rev. Genomics Hum. Genet. *6*, 45–68. 10.1146/annurev.genom.6.080604.162305.
6. Marino, B., Digilio, M.C., Toscano, A., Giannotti, A., and Dallapiccola, B. (1999). Congenital heart diseases in children with Noonan syndrome: An expanded cardiac spectrum with high prevalence of atrioventricular canal. J. Pediatr. *135*, 703–706. 10.1016/s0022-3476(99)70088-0.
7. Tartaglia, M., Niemeyer, C.M., Fragale, A., Song, X., Buechner, J., Jung, A., Hählen, K., Hasle, H., Licht, J.D., and Gelb, B.D. (2003). Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. Nat. Genet. *34*, 148–150. 10.1038/ng1156.
8. Loh, M.L., Vattikuti, S., Schubbert, S., Reynolds, M.G., Carlson, E., Lieuw, K.H., Cheng, J.W., Lee, C.M., Stokoe, D., Bonifas, J.M., et al. (2004). Mutations in PTPN11 implicate the SHP-2 phosphatase in leukemogenesis. Blood *103*, 2325–2331. 10.1182/blood-2003-09-3287.
9. Bentires-Alj, M., Paez, J.G., David, F.S., Keilhack, H., Halmos, B., Naoki, K., Maris, J.M., Richardson, A., Bardelli, A., Sugarbaker, D.J., et al. (2004). Activating mutations of the noonan syndrome-associated SHP2/PTPN11 gene in human solid tumors and adult acute myelogenous leukemia. Cancer Res. *64*, 8816–8820. 10.1158/0008-5472.CAN-04-1923.
10. Hof, P., Pluskey, S., Dhe-Paganon, S., Eck, M.J., and Shoelson, S.E. (1998). Crystal structure of the tyrosine phosphatase SHP-2. Cell *92*, 441–450. 10.1016/s0092-8674(00)80938-1.
11. Tartaglia, M., Martinelli, S., Stella, L., Bocchinfuso, G., Flex, E., Cordeddu, V., Zampino, G., Burgt, I. van der, Palleschi, A., Petrucci, T.C., et al. (2006). Diversity and functional consequences of germline and somatic PTPN11 mutations in human disease. Am. J. Hum. Genet. *78*, 279–290. 10.1086/499925.
12. Kaneko, T., Joshi, R., Feller, S.M., and Li, S.S. (2012). Phosphotyrosine recognition domains: the typical, the atypical and the versatile. Cell Commun. Signal. *10*, 32. 10.1186/1478-811X-10-32.
13. Cohen, P. (2001). The role of protein phosphorylation in human health and disease. . European Journal of Biochemistry *268*, 5001–5010. 10.1046/j.0014-2956.2001.02473.x.
14. Martinelli, S., Torreri, P., Tinti, M., Stella, L., Bocchinfuso, G., Flex, E., Grottesi, A., Ceccarini, M., Palleschi, A., Cesareni, G., et al. (2008). Diverse driving forces underlie the invariant occurrence of the T42A, E139D, I282V and T468M SHP2 amino acid substitutions causing Noonan and LEOPARD syndromes. Hum. Mol. Genet. *17*, 2018–2029. 10.1093/hmg/ddn099.
15. Weismann, C.G., Hager, A., Kaemmerer, H., Maslen, C.L., Morris, C.D., Schranz, D., Kreuder, J., and Gelb, B.D. (2005). PTPN11 mutations play a minor role in isolated congenital heart disease. Am. J. Med. Genet. A *136*, 146–151. 10.1002/ajmg.a.30789.
16. Kratz, C.P., Niemeyer, C.M., Castleberry, R.P., Cetin, M., Bergsträsser, E., Emanuel, P.D., Hasle, H., Kardos, G., Klein, C., Kojima, S., et al. (2005). The mutational spectrum of PTPN11 in juvenile myelomonocytic leukemia and Noonan syndrome/myeloproliferative disease. Blood *106*, 2183–2185. 10.1182/blood-2005-02-0531.

17. Keilhack, H., David, F.S., McGregor, M., Cantley, L.C., and Neel, B.G. (2005). Diverse biochemical properties of Shp2 mutants. Implications for disease phenotypes. J. Biol. Chem. *280*, 30984–30993. 10.1074/jbc.M504699200.

18. Pádua, R.A.P., Sun, Y., Marko, I., Pitsawong, W., Stiller, J.B., Otten, R., and Kern, D. (2018). Mechanism of activating mutations and allosteric drug inhibition of the phosphatase SHP2. Nat. Commun. *9*, 4507. 10.1038/s41467-018-06814-w.

19. LaRochelle, J.R., Fodor, M., Vemulapalli, V., Mohseni, M., Wang, P., Stams, T., LaMarche, M.J., Chopra, R., Acker, M.G., and Blacklow, S.C. (2018). Structural reorganization of SHP2 by oncogenic mutations and implications for oncoprotein resistance to allosteric inhibition. Nat. Commun. *9*, 4508. 10.1038/s41467-018-06823-9.

20. Qiu, W., Wang, X., Romanov, V., Hutchinson, A., Lin, A., Ruzanov, M., Battaile, K.P., Pai, E.F., Neel, B.G., and Chirgadze, N.Y. (2014). Structural insights into Noonan/LEOPARD syndrome-related mutants of protein-tyrosine phosphatase SHP2 (PTPN11). BMC Struct. Biol. *14*, 10. 10.1186/1472-6807-14-10.

21. Koch, C.A., Anderson, D., Moran, M.F., Ellis, C., and Pawson, T. (1991). SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. Science *252*, 668–674. 10.1126/science.1708916.

22. Hidaka, M., Homma, Y., and Takenawa, T. (1991). Highly conserved eight amino acid sequence in SH2 is important for recognition of phosphotyrosine site. Biochem. Biophys. Res. Commun. *180*, 1490–1497. 10.1016/s0006-291x(05)81364-6.

23. Noguchi, T., Matozaki, T., Horita, K., Fujioka, Y., and Kasuga, M. (1994). Role of SH-PTP2, a protein-tyrosine phosphatase with Src homology 2 domains, in insulin-stimulated Ras activation. Mol. Cell. Biol. *14*, 6674–6682. 10.1128/mcb.14.10.6674-6682.1994.

24. Kapoor, G.S., Zhan, Y., Johnson, G.R., and O'Rourke, D.M. (2004). Distinct domains in the SHP-2 phosphatase differentially regulate epidermal growth factor receptor/NF-kappaB activation through Gab1 in glioblastoma cells. Mol. Cell. Biol. *24*, 823–836. 10.1128/MCB.24.2.823-836.2004.

25. Gu, H., Pratt, J.C., Burakoff, S.J., and Neel, B.G. (1998). Cloning of p97/Gab2, the major SHP2-binding protein in hematopoietic cells, reveals a novel pathway for cytokine-induced gene activation. Mol. Cell *2*, 729–740. 10.1016/s1097-2765(00)80288-9.

26. Kuhné, M.R., Zhao, Z., Rowles, J., Lavan, B.E., Shen, S.H., Fischer, E.H., and Lienhard, G.E. (1994). Dephosphorylation of insulin receptor substrate 1 by the tyrosine phosphatase PTP2C. J. Biol. Chem. *269*, 15833–15837. 10.1016/S0021-9258(17)40756-3.

27. Okazaki, T., Maeda, A., Nishimura, H., Kurosaki, T., and Honjo, T. (2001). PD-1 immunoreceptor inhibits B cell receptor-mediated signaling by recruiting src homology 2-domain-containing tyrosine phosphatase 2 to phosphotyrosine. Proc Natl Acad Sci USA *98*, 13866–13871. 10.1073/pnas.231486598.

28. Higashi, H., Tsutsumi, R., Muto, S., Sugiyama, T., Azuma, T., Asaka, M., and Hatakeyama, M. (2002). SHP-2 tyrosine phosphatase as an intracellular target of Helicobacter pylori CagA protein. Science *295*, 683–686. 10.1126/science.1067147.

29. Imhof, D., Wavreille, A.-S., May, A., Zacharias, M., Tridandapani, S., and Pei, D. (2006). Sequence specificity of SHP-1 and SHP-2 Src homology 2 domains. Critical roles of residues beyond the pY+3 position. J. Biol. Chem. *281*, 20271–20282. 10.1074/jbc.M601047200.

30. Marasco, M., Berteotti, A., Weyershaeuser, J., Thorausch, N., Sikorska, J., Krausze, J., Brandt, H.J., Kirkpatrick, J., Rios, P., Schamel, W.W., et al. (2020). Molecular mechanism of SHP2 activation by PD-1 stimulation. Sci. Adv. *6*, eaay4458. 10.1126/sciadv.aay4458.

31. Li, A., Voleti, R., Lee, M., Gagoski, D., and Shah, N.H. (2023). High-throughput profiling of sequence recognition by tyrosine kinases and SH2 domains using bacterial peptide display. eLife *12*. 10.7554/eLife.82345.

32. Shah, N.H., Löbel, M., Weiss, A., and Kuriyan, J. (2018). Fine-tuning of substrate preferences of the Src-family kinase Lck revealed through a high-throughput specificity screen. eLife *7*, e35190. 10.7554/eLife.35190.

33. Rönnstrand, L., Arvidsson, A.K., Kallin, A., Rorsman, C., Hellman, U., Engström, U., Wernstedt, C., and Heldin, C.H. (1999). SHP-2 binds to Tyr763 and Tyr1009 in the PDGF beta-receptor and mediates PDGF-induced activation of the Ras/MAP kinase pathway and chemotaxis. Oncogene *18*, 3696–3702. 10.1038/sj.onc.1202705.

34. Hitomi, K., Tahara-Hanaoka, S., Someya, S., Fujiki, A., Tada, H., Sugiyama, T., Shibayama, S., Shibuya, K., and Shibuya, A. (2010). An immunoglobulin-like receptor, Allergin-1, inhibits immunoglobulin E-mediated immediate hypersensitivity reactions. Nat. Immunol. *11*, 601–607. 10.1038/ni.1886.

35. O'Shea, J.P., Chou, M.F., Quader, S.A., Ryan, J.K., Church, G.M., and Schwartz, D. (2013). pLogo: a probabilistic approach to visualizing sequence motifs. Nat. Methods *10*, 1211–1212. 10.1038/nmeth.2646.

36. Anselmi, M., Calligari, P., Hub, J.S., Tartaglia, M., Bocchinfuso, G., and Stella, L. (2020). Structural Determinants of Phosphopeptide Binding to the N-Terminal Src Homology 2 Domain of the SHP2 Phosphatase. J. Chem. Inf. Model. *60*, 3157–3171. 10.1021/acs.jcim.0c00307.

37. Marasco, M., Kirkpatrick, J., Nanna, V., Sikorska, J., and Carlomagno, T. (2021). Phosphotyrosine couples peptide binding and SHP2 activation via a dynamic allosteric network. Comput. Struct. Biotechnol. J. *19*, 2398–2415. 10.1016/j.csbj.2021.04.040.

38. Toto, A., Malagrinò, F., Visconti, L., Troilo, F., and Gianni, S. (2020). Unveiling the Molecular Basis of the Noonan Syndrome-Causing Mutation T42A of SHP2. Int. J. Mol. Sci. *21*. 10.3390/ijms21020461.

39. Anselmi, M., and Hub, J.S. (2020). An allosteric interaction controls the activation mechanism of SHP2 tyrosine phosphatase. Sci. Rep. *10*, 18530. 10.1038/s41598-020-75409-7.

40. Lee, C.H., Kominos, D., Jacques, S., Margolis, B., Schlessinger, J., Shoelson, S.E., and Kuriyan, J. (1994). Crystal structures of peptide complexes of the amino-terminal SH2 domain of the Syp tyrosine phosphatase. Structure *2*, 423–438. 10.1016/s0969-2126(00)00044-7.

41. Zhang, Y., Zhang, J., Yuan, C., Hard, R.L., Park, I.-H., Li, C., Bell, C., and Pei, D. (2011). Simultaneous binding of two peptidyl ligands by a SRC homology 2 domain. Biochemistry *50*, 7637–7646. 10.1021/bi200439v.

42. Liu, Y., Lau, J., Li, W., Tempel, W., Li, L., Dong, A., Narula, A., Qin, S., and Min, J. (2016). Structural basis for the regulatory role of the PPxY motifs in the thioredoxin-interacting protein TXNIP. Biochem. J. *473*, 179–187. 10.1042/BJ20150830.

43. Hayashi, T., Senda, M., Suzuki, N., Nishikawa, H., Ben, C., Tang, C., Nagase, L., Inoue, K., Senda, T., and Hatakeyama, M. (2017). Differential Mechanisms for SHP2 Binding and Activation Are Exploited by Geographically Distinct Helicobacter pylori CagA Oncoproteins. Cell Rep. *20*, 2876–2890. 10.1016/j.celrep.2017.08.080.

44. Liu, B.A., Shah, E., Jablonowski, K., Stergachis, A., Engelmann, B., and Nash, P.D. (2011). The SH2 domain-containing proteins in 21 species establish the provenance and scope of phosphotyrosine signaling in eukaryotes. Sci. Signal. *4*, ra83. 10.1126/scisignal.2002105.

45. Kaneko, T., Huang, H., Cao, X., Li, X., Li, C., Voss, C., Sidhu, S.S., and Li, S.S.C. (2012). Superbinder SH2 domains act as antagonists of cell signaling. Sci. Signal. *5*, ra68. 10.1126/scisignal.2003021.

46. Martyn, G.D., Veggiani, G., Kusebauch, U., Morrone, S.R., Yates, B.P., Singer, A.U., Tong, J., Manczyk, N., Gish, G., Sun, Z., et al. (2022). Engineered SH2 domains for targeted phosphoproteomics. ACS Chem. Biol. *17*, 1472–1484. 10.1021/acschembio.2c00051.

47. Lassila, J.K., Zalatan, J.G., and Herschlag, D. (2011). Biological phosphoryl-transfer reactions: understanding mechanism and catalysis. Annu. Rev. Biochem. *80*, 669–702. 10.1146/annurev-biochem-060409-092741.

48. Wu, B., Wang, F., Zhang, J., Zhang, Z., Qin, L., Peng, J., Li, F., Liu, J., Lu, G., Gong, Q., et al. (2012). Identification and structural basis for a novel interaction between Vav2 and Arap3. J. Struct. Biol. *180*, 84–95. 10.1016/j.jsb.2012.06.011.

49. RCSB PDB - 7RNV: SH2 domain of guanine nucleotide exchange factor Vav2 in complex with

an actin peptide with phosphorylated tyrosine 53 https://www.rcsb.org/structure/7RNV.

50. Songyang, Z., Shoelson, S.E., McGlade, J., Olivier, P., Pawson, T., Bustelo, X.R., Barbacid, M., Sabe, H., Hanafusa, H., and Yi, T. (1994). Specific motifs recognized by the SH2 domains of Csk, 3BP2, fps/fes, GRB-2, HCP, SHC, Syk, and Vav. Mol. Cell. Biol. *14*, 2777–2785. 10.1128/mcb.14.4.2777-2785.1994.

51. Patsoukis, N., Duke-Cohan, J.S., Chaudhri, A., Aksoylar, H.-I., Wang, Q., Council, A., Berg, A., Freeman, G.J., and Boussiotis, V.A. (2020). Interaction of SHP-2 SH2 domains with PD-1 ITSM induces PD-1 dimerization and SHP-2 activation. Commun. Biol. *3*, 128. 10.1038/s42003-020-0845-0.

52. Tao, Y., Xie, J., Zhong, Q., Wang, Y., Zhang, S., Luo, F., Wen, F., Xie, J., Zhao, J., Sun, X., et al. (2021). A novel partially open state of SHP2 points to a “multiple gear” regulation mechanism. J. Biol. Chem. *296*, 100538. 10.1016/j.jbc.2021.100538.

53. Montagner, A., Yart, A., Dance, M., Perret, B., Salles, J.-P., and Raynal, P. (2005). A novel role for Gab1 and SHP2 in epidermal growth factor-induced Ras activation. J. Biol. Chem. *280*, 5350–5360. 10.1074/jbc.M410012200.

54. Schaeper, U., Gehring, N.H., Fuchs, K.P., Sachs, M., Kempkes, B., and Birchmeier, W. (2000). Coupling of Gab1 to c-Met, Grb2, and Shp2 mediates biological responses. J. Cell Biol. *149*, 1419–1432. 10.1083/jcb.149.7.1419.

55. Time-resolved phosphoproteomics reveals scaffolding and catalysis-responsive patterns of SHP2-dependent signaling | eLife https://elifesciences.org/articles/64251.

56. Xu, X., Hou, B., Fulzele, A., Masubuchi, T., Zhao, Y., Wu, Z., Hu, Y., Jiang, Y., Ma, Y., Wang, H., et al. (2020). PD-1 and BTLA regulate T cell signaling differentially and only partially through SHP1 and SHP2. J. Cell Biol. *219*. 10.1083/jcb.201905085.

57. Chen, Y.-N.P., LaMarche, M.J., Chan, H.M., Fekkes, P., Garcia-Fortanet, J., Acker, M.G., Antonakos, B., Chen, C.H.-T., Chen, Z., Cooke, V.G., et al. (2016). Allosteric inhibition of SHP2 phosphatase inhibits cancers driven by receptor tyrosine kinases. Nature *535*, 148–152. 10.1038/nature18621.

58. Nichols, R.J., Haderk, F., Stahlhut, C., Schulze, C.J., Hemmati, G., Wildes, D., Tzitzilonis, C., Mordec, K., Marquez, A., Romero, J., et al. (2018). RAS nucleotide cycling underlies the SHP2 phosphatase dependence of mutant BRAF-, NF1- and RAS-driven cancers. Nat. Cell Biol. *20*, 1064–1073. 10.1038/s41556-018-0169-1.

59. Dance, M., Montagner, A., Salles, J.-P., Yart, A., and Raynal, P. (2008). The molecular functions of Shp2 in the Ras/Mitogen-activated protein kinase (ERK1/2) pathway. Cell. Signal. *20*, 453–459. 10.1016/j.cellsig.2007.10.002.

60. Müller, P.J., Rigbolt, K.T.G., Paterok, D., Piehler, J., Vanselow, J., Lasonder, E., Andersen, J.S., Schaper, F., and Sobota, R.M. (2013). Protein tyrosine phosphatase SHP2/PTPN11 mistargeting as a consequence of SH2-domain point mutations associated with Noonan Syndrome and leukemia. J. Proteomics *84*, 132–147. 10.1016/j.jprot.2013.04.005.

61. Visconti, L., Malagrinò, F., Pagano, L., and Toto, A. (2020). Understanding the Mechanism of Recognition of Gab2 by the N-SH2 Domain of SHP2. Life (Basel) *10*. 10.3390/life10060085.

62. Creixell, P., Schoof, E.M., Simpson, C.D., Longden, J., Miller, C.J., Lou, H.J., Perryman, L., Cox, T.R., Zivanovic, N., Palmeri, A., et al. (2015). Kinome-wide decoding of network-attacking mutations rewiring cancer signaling. Cell *163*, 202–217. 10.1016/j.cell.2015.08.056.

63. Lundby, A., Franciosa, G., Emdal, K.B., Refsgaard, J.C., Gnosa, S.P., Bekker-Jensen, D.B., Secher, A., Maurya, S.R., Paul, I., Mendez, B.L., et al. (2019). Oncogenic mutations rewire signaling pathways by switching protein recruitment to phosphotyrosine sites. Cell *179*, 543-560.e26. 10.1016/j.cell.2019.09.008.

64. Mo, X., Niu, Q., Ivanov, A.A., Tsang, Y.H., Tang, C., Shu, C., Li, Q., Qian, K., Wahafu, A., Doyle, S.P., et al. (2022). Systematic discovery of mutation-directed neo-protein-protein interactions in cancer. Cell *185*, 1974-1985.e12. 10.1016/j.cell.2022.04.014.

65. Creixell, P., Palmeri, A., Miller, C.J., Lou, H.J., Santini, C.C., Nielsen, M., Turk, B.E., and

Linding, R. (2015). Unmasking determinants of specificity in the human kinome. Cell *163*, 187–201. 10.1016/j.cell.2015.08.057.

66. Bradley, D., Viéitez, C., Rajeeve, V., Selkrig, J., Cutillas, P.R., and Beltrao, P. (2021). Sequence and Structure-Based Analysis of Specificity Determinants in Eukaryotic Protein Kinases. Cell Rep. *34*, 108602. 10.1016/j.celrep.2020.108602.

67. O'Reilly, A.M., Pluskey, S., Shoelson, S.E., and Neel, B.G. (2000). Activated mutants of SHP-2 preferentially induce elongation of Xenopus animal caps. Mol. Cell. Biol. *20*, 299–311. 10.1128/mcb.20.1.299-311.2000.

68. Cushing, P.R., Fellows, A., Villone, D., Boisguérin, P., and Madden, D.R. (2008). The relative binding affinities of PDZ partners for CFTR: a biochemical basis for efficient endocytic recycling. Biochemistry *47*, 10084–10098. 10.1021/bi8003928.

69. Shah, N.H., Wang, Q., Yan, Q., Karandur, D., Kadlecek, T.A., Fallahee, I.R., Russ, W.P., Ranganathan, R., Weiss, A., and Kuriyan, J. (2016). An electrostatic selection mechanism controls sequential kinase signaling downstream of the T cell receptor. eLife *5*, e20105. 10.7554/eLife.20105.

70. Magoč, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics *27*, 2957–2963. 10.1093/bioinformatics/btr507.

71. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet j. *17*, 10. 10.14806/ej.17.1.200.

72. Michaels, Y.S., Barnkob, M.B., Barbosa, H., Baeumler, T.A., Thompson, M.K., Andre, V., Colin-York, H., Fritzsche, M., Gileadi, U., Sheppard, H.M., et al. (2019). Precise tuning of gene expression levels in mammalian cells. Nat. Commun. *10*, 818. 10.1038/s41467-019-08777-y.

73. Fauser, J., Huyot, V., Matsche, J., Szynal, B.N., Alexeev, Y., Kota, P., and Karginov, A.V. (2022). Dissecting protein tyrosine phosphatase signaling by engineered chemogenetic control of its activity. J. Cell Biol. *221*. 10.1083/jcb.202111066.

74. Schrödinger, L., and Delano, W. (2020). PyMOL (http://www.pymol.org/pymol).

75. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. *79*, 926. 10.1063/1.445869.

76. Case, D.A., Aktulga, H.M., Belfon, K., Ben-Shalom, I.Y., Brozell, S.R., Cerutti, D.S., Cheatham, III, T.E., Cisneros, G.A., Cruzeiro, V.W.D., Darden, T.A., et al. (2021). Amber 2021 (University of California, San Francisco).

77. Case, D.A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R.J. (2005). The Amber biomolecular simulation programs. J. Comput. Chem. *26*, 1668–1688. 10.1002/jcc.20290.

78. Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., and Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J. Chem. Theory Comput. *11*, 3696–3713. 10.1021/acs.jctc.5b00255.

79. Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J. Chem. Phys. *98*, 10089. 10.1063/1.464397.

80. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H.J.C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. *23*, 327–341. 10.1016/0021-9991(77)90098-5.

81. Case, D.A., Aktulga, H.M., Belfon, K., Ben-Shalom, I.Y., Berryman, J.T., Brozell, S.R., Cerutti, D.S., Cheatham, III, T.E., Cisneros, G.A., Cruzeiro, V.W.D., et al. (2022). Amber 2022 (University of California, San Francisco).