The pathogenic T42A mutation in SHP2 rewires the interaction specificity of its N-terminal regulatory domain

Anne E. van Vlimmeren[1,2,†], Rashmi Voleti[1,†], Cassandra A. Chartier[1], Ziyuan Jiang[1], Deepti Karandur[3], Preston A. Humphries[4], Wan-Lin Lo[4], 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

[4] Division of Microbiology and Immunology, Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT 84112

[†] 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

Preston A. Humphries: 0000-0001-8736-6432

Wan-Lin Lo: 0000-0002-9074-6847

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

Keywords

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

Abstract

Mutations in the tyrosine phosphatase SHP2 are associated with a variety of human diseases. Most mutations in SHP2 increase its basal catalytic activity by disrupting auto-inhibitory interactions between its phosphatase domain and N-terminal SH2 (phosphotyrosine recognition) domain. By contrast, some disease-associated mutations located in the ligand-binding pockets of the N- or C-terminal SH2 domains do not increase basal activity and likely exert their pathogenicity through alternative mechanisms. We lack a molecular understanding of how these SH2 mutations impact SHP2 structure, activity, and signaling. Here, we characterize five SHP2 SH2 domain ligand-binding pocket mutants through a combination of high-throughput biochemical screens, biophysical and biochemical measurements, and molecular dynamics simulations. We show that, while some of these mutations alter binding affinity to phosphorylation sites, the T42A mutation in the N-SH2 domain is unique in that it also substantially alters ligand-binding specificity, despite being 8-10 Å from the specificity-determining region of the SH2 domain. This mutation exerts its effect on sequence specificity by remodeling the phosphotyrosine binding pocket, altering the mode of engagement of both the phosphotyrosine and surrounding residues on the ligand. The functional consequence of this altered specificity is that the T42A mutant has biased sensitivity toward a subset of activating ligands and enhances downstream signaling. 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.

Significance Statement

The protein tyrosine phosphatase SHP2 is mutated in a variety of human diseases, including several cancers and developmental disorders. Most mutations in SHP2 hyperactivate the enzyme by destabilizing its auto-inhibited state, but several disease-associated mutations do not conform to this mechanism. We show that one such mutation, T42A, alters the ligand binding specificity of the N-terminal regulatory domain of SHP2, causing the mutant phosphatase to be more readily activated by certain upstream signals than the wild-type phosphatase. Our findings reveal a novel mode of SHP2 dysregulation that will improve our understanding of pathogenic signaling. Our study also illustrates how mutations distal to the specificity-determining region of a protein can alter ligand binding specificity.

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). Germline *PTPN11* mutations underlie approximately 50% of all cases of Noonan Syndrome, a congenital disorder which is characterized by a wide range of developmental defects (4–6). In addition, somatic mutations in *PTPN11* are 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 known 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, largely 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). This mutation is thought to make SHP2 more readily activated by upstream phosphoproteins. However, SHP2 still requires 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 SHP2 activity (15, 16). The C-SH2 domain mutant R138Q has been identified in melanoma, whereas the E139D mutation is 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, cause tumorigenesis or aberrant 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 changes the sequence specificity of the N-SH2 domain. Notably, the T42A mutation does not lie in a canonical specificity determining region for SH2 domains (**Figure 1E, F**) (18, 19). Through molecular dynamics simulations and further biochemical experiments, we identify structural changes caused by the T42A mutation that 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 the interaction specificity of the N-SH2 domain.

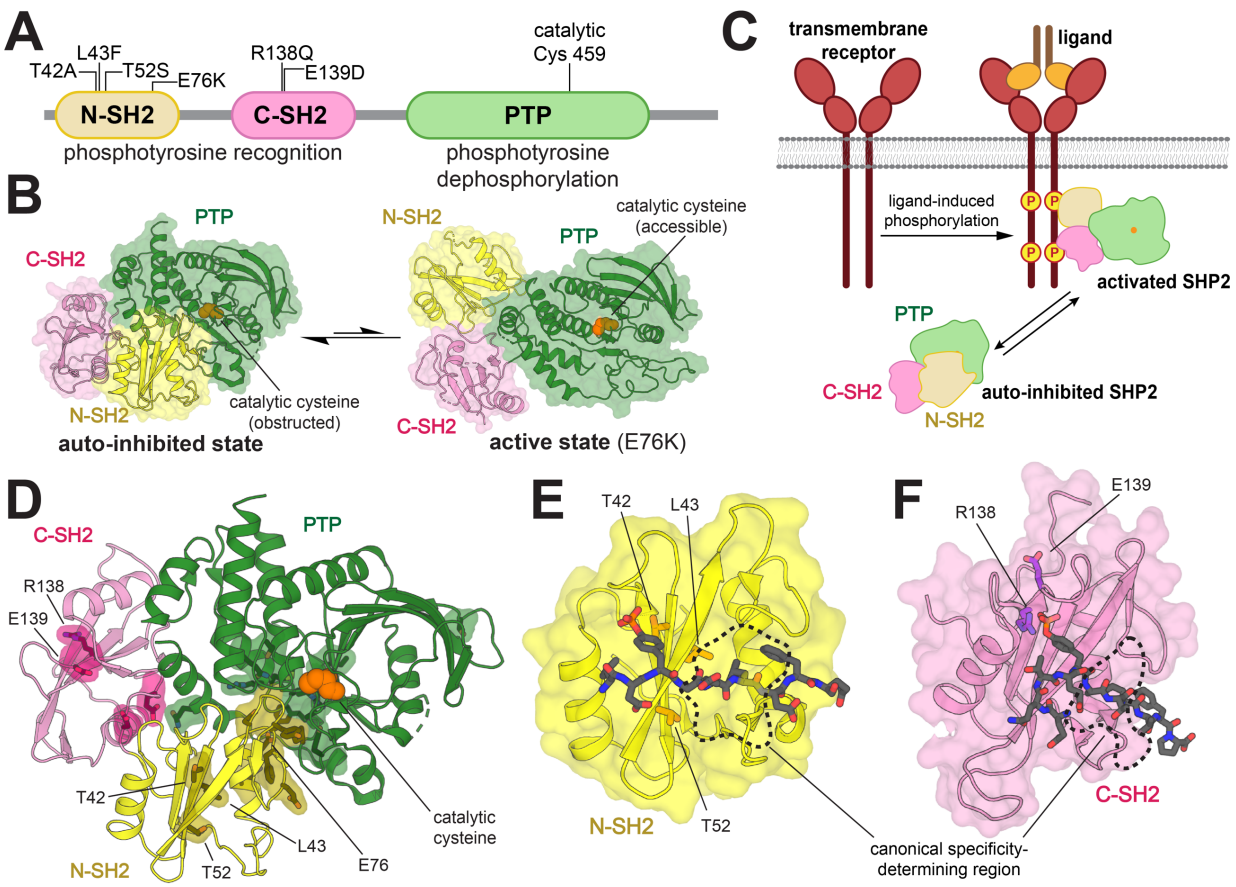

Figure 1. Structure and regulation of SHP2. (**A**) Domain architecture diagram of SHP2. Relevant mutations and the catalytic cysteine (Cys^{459}) are indicated. (**B**) Auto-inhibited state of SHP2 (*left*, PDB: 4DGP) and representative active state of SHP2 (*right*, E76K mutant, PDB: 6CRF). (**C**) The SH2 domains of SHP2 bind to upstream phosphoproteins, inducing a conformational change that activates SHP2. (**D**) Rendering of auto-inhibited SHP2, highlighting several disease-associated mutation sites within (e.g. E76) and outside (e.g. T42) of the N-SH2/PTP interface (PDB: 4DGP). (**E**) Mutation sites in or near the N-SH2 binding pocket (PDB: 6ROY). (**F**) Mutation sites in or near the C-SH2 binding pocket (PDB: 6R5G). The primary specificity-determining regions of the SH2 domains, which dictate +1 to +5 residue preferences, are marked with black dashed lines.

Results

Mutations in the SH2 domains of SHP2 impact both binding affinity and sequence specificity

Mutations in the SH2 domains of SHP2 potentially 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 disease-relevant and close to the pTyr-binding pocket (**Figure 1E**). These mutations do not cause a substantial increase in basal phosphatase activity, in contrast to the E76K mutation, which lies at the auto-inhibitory interface and strongly enhances phosphatase activity (**Figure S1A and Table S1A**) (17, 20, 21). We also 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 S1A and Table S1**), as has been reported previously (17, 22). The R138Q mutation is expected to disrupt phosphopeptide binding, as Arg^{138} directly coordinates the phosphoryl group of phosphotyrosine ligand (23, 24). This mutation had no impact on the basal catalytic activity of SHP2 (**Figure S1A and Table S1**). We also measured the melting temperatures of these mutants, as a proxy for the conformational state of the protein (25). Only the L43F mutant showed a modest decrease in melting temperature, suggestive of a slightly more open conformation than wild-type (**Figure S1B**). By comparison, the E76K mutant had a much lower melting temperature, consistent with its open conformation.

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**) (26). 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**) (26–32). We observed a broad range of effects on binding affinity for N-$SH2^{T42A}$. Compared to N-$SH2^{WT}$, 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. By contrast, 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**). To confirm that the effects were specific to these ligand-binding pocket mutations, we measured the binding affinity of N-$SH2^{E76K}$ to four peptides. All N-$SH2^{E76K}$ affinities were similar to those for N-$SH2^{WT}$ (**Figure S2C, Table S2**).

For C-SH2 domain mutants R138Q and E139D, we first measured binding against two fluorescent phosphopeptides: one derived from a known binding site on PD-1 ($pTyr^{248}$), as well as the designed ligand Imhof-9 (**Figure S2D and Table S2**) (32, 33). As expected, C-$SH2^{R138Q}$ binding to phosphopeptides was severely attenuated (**Figure S2D**), and this mutant was therefore excluded from further binding analyses. C-$SH2^{E139D}$ binding was comparable to C-$SH2^{WT}$ binding against the two fluorescent phosphopeptides (**Figure S2D**) and against the 9 unlabeled peptides used for N-SH2 binding assays (**Figure S2E and Table S2**). 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.

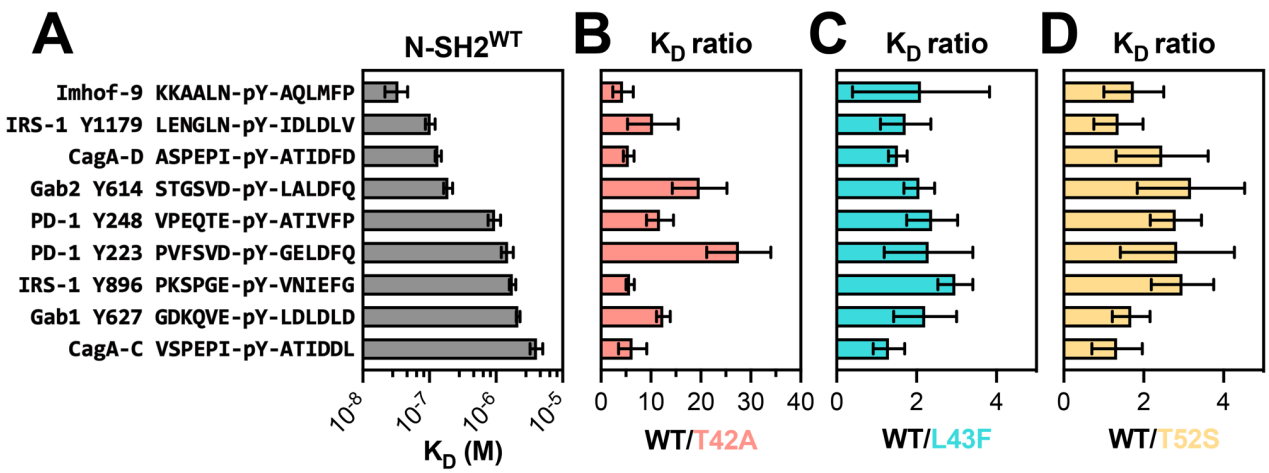

Figure 2. K_D measurements reveal sequence-specific enhancement of binding affinity in N-SH2^{T42A}. (**A**) Measured binding affinities of N-SH2^{WT} against peptides derived from various known SHP2 interactors. (**B**) Fold-change in K_D for N-SH2^{T42A} compared to N-SH2^{WT}, for each of the peptides shown in panel (A). (**C**) Same as (B), but for N-SH2^{L43F}. (**D**) Same as (B), but for N-SH2^{T52S}. For (**A**)-(**D**), N = 3-4 independent protein, peptide, and fluorescent peptide titrations. Source data and p-values 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 specificity of the SH2 mutants relative to their wild-type counterparts in a large-scale, unbiased screen. This method entails bacterial display of a peptide library, enzymatic phosphorylation of the displayed peptides, selective enrichment using SH2-coated beads, and deep sequencing of the peptide-coding DNA (34). For each peptide in a library, an enrichment score is calculated from the deep sequencing data as the frequency of the peptide in the SH2-enriched sample divided by the frequency of the peptide in the unenriched input sample (34). For this study, we used two largely non-overlapping 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 (34). The Human pTyr Library consists of 1,916 sequences derived from known phosphorylation sites in the human proteome, along with another 617 control mutants (35).

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,E**). 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 selected three sequences for fluorescence polarization binding affinity measurements (**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 (36). The third peptide, MILR1 pTyr^{338}, was derived from a known SHP2 binding site on the mast cell immunoreceptor Allegrin-1 (37). 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.

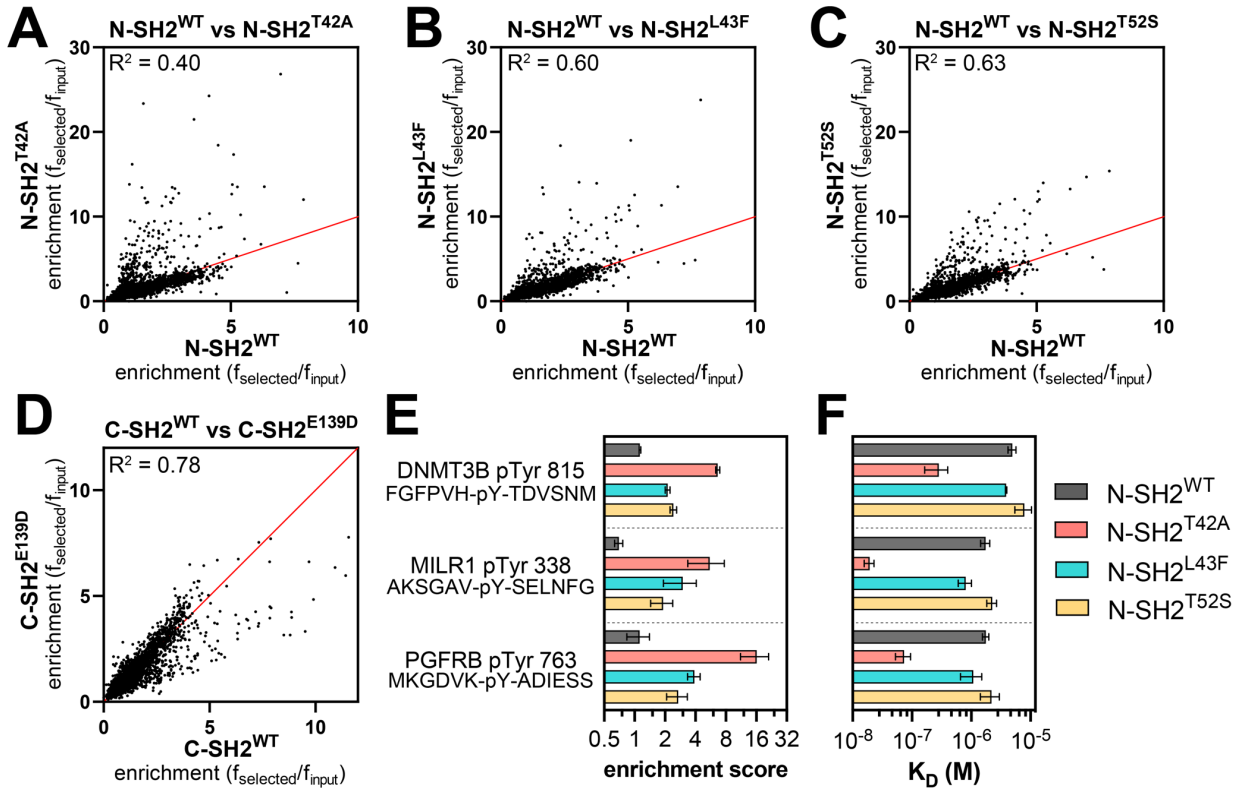

Figure 3. Peptide library screens identify sequences with enhanced N-SH2^{T42A} binding. (**A-C**) Comparison of enrichment scores for N-SH2^{WT} and each N-SH2 mutant. (**D**) Comparison of enrichment scores for C-SH2^{WT} and C-SH2^{E139D}. The red line denotes x = y. pTyr-Var Library screens (N = 2), Human pTyr Library screens (N = 3). 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) (N = 3).

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

Next, we examined the sequence features of the peptides enriched in the screens with each N-SH2 domain. 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 (34). 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**) (38).

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 so, 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 (19, 32). 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} has a unique enhanced preference for hydrophilic residues. One notable difference, 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. 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} using bacterial peptide display. 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 S3F-G 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 S3F-G and Table S4**).

Our results show that the two domains have modestly correlated binding preferences with respect to this scanning mutagenesis library (**Figure S3G**)**.** 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 S3F**)**.** For example, the substitution of the +1 Gly to Ala or Thr is favored by N-SH2^{WT}, consistent with previous studies (19), 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 S3F**). Taken together, our experiments with the human phosphosite libraries 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.

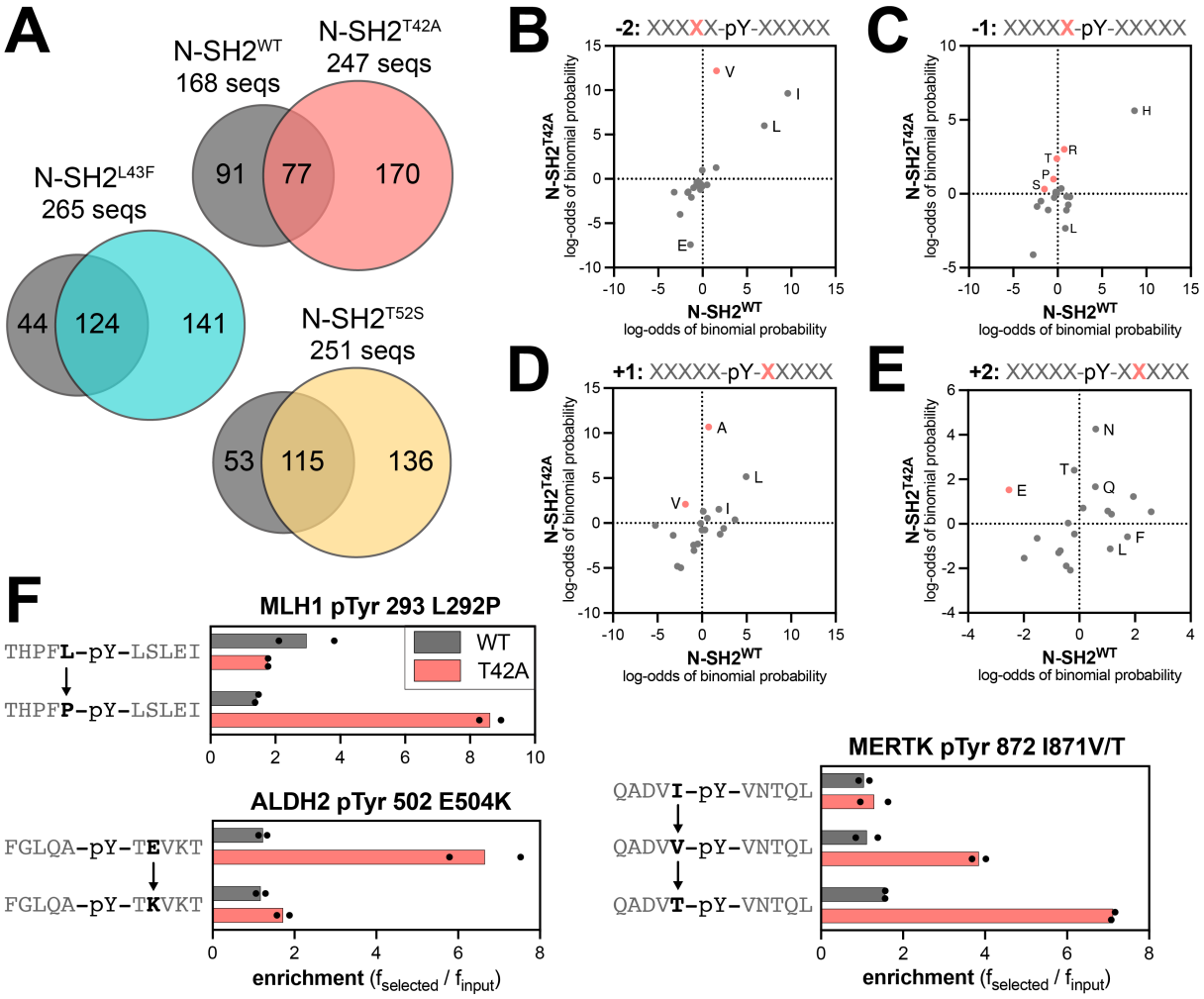

Figure 4. Analysis of enriched sequence features reveals specificity differences at several positions. (**A**) Overlap in sequences enriched by each N-SH2 domain above an enrichment score cutoff of 3.2. (**B**) Log-transformed probabilities of amino acid enrichment at the -2 position relative to the 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}. N = 2 independent library transformations and screens.

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 (14, 17, 39–41). 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**) (33, 42–45). 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} (14, 17, 39–41), 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 (Ala, Val, or Leu) at the position corresponding to Thr^{42} (**Figure S4A**) (46), but the impact of this residue on sequence specificity has not been systematically explored. Here, we used molecular dynamics (MD) simulations to examine how the T42A mutation impacts SHP2 N-SH2 peptide engagement (39, 47, 48).

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 run. 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. Closer inspection of BC loop interactions in the 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} made a persistent hydrogen bond with a non-bridging oxygen on the phosphoryl group, which constrains the orientation of the phosphotyrosine residue (**Figure 5B and Figure S5B,C**). By contrast, in almost every N-SH2^{T42A} simulation, where the phosphoryl group was not tethered to Thr^{42}, the phosphotyrosine residue relaxed into a new orientation characterized by a distinct hydrogen bond network (**Figure 5C,** *additional structural details in* **Figure S5B-H**).

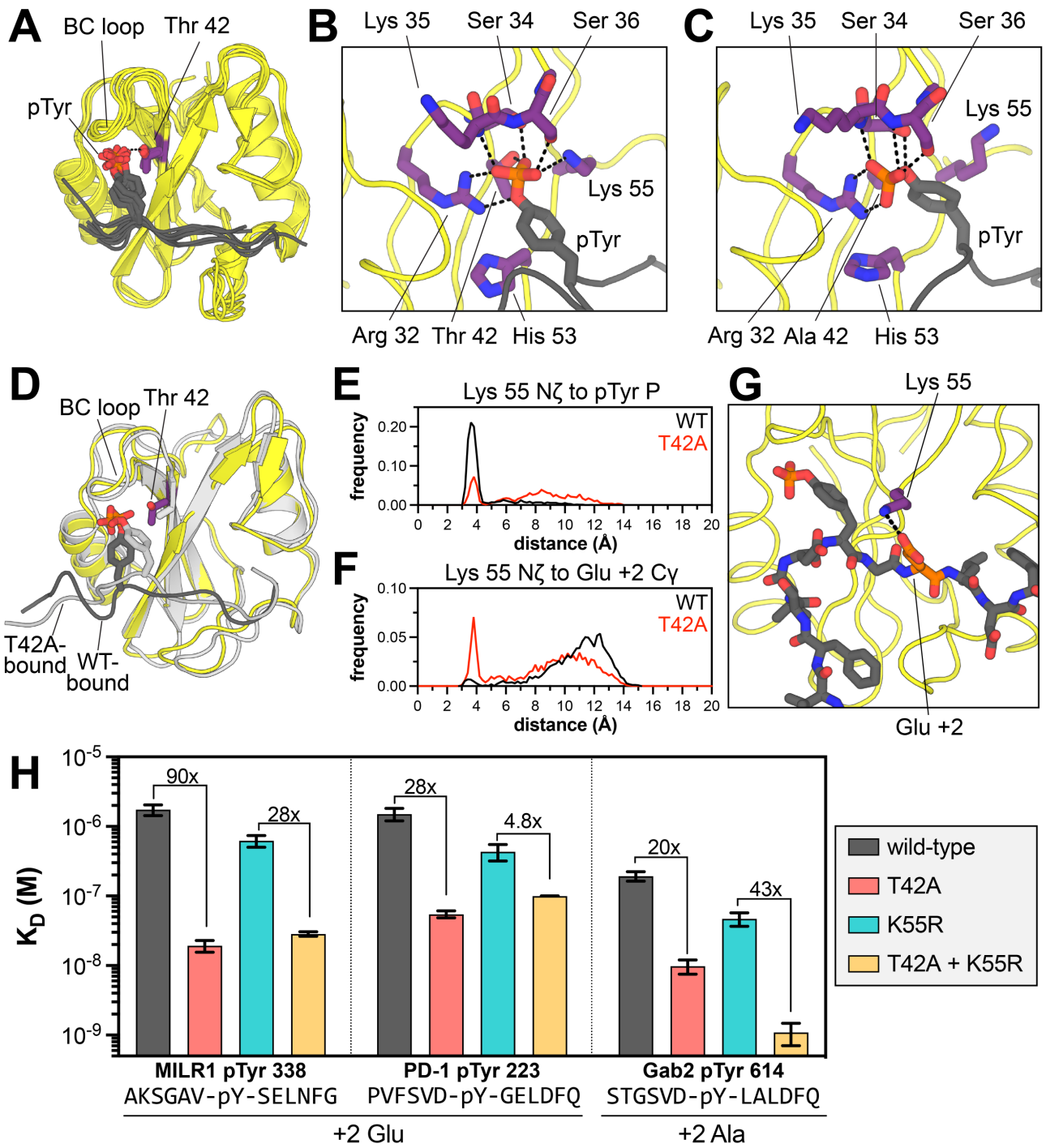

Figure 5. Structural impact of the T42A mutation on phosphotyrosine and proximal sequence recognition. (**A**) Hydrogen bonding of Thr^{42} in SHP2 N-SH2^{WT} to the phosphoryl group of phosphopeptide ligands in several crystal structures (PDB: 6ROY, 1AYA, 1AYB, 3TL0, 5DF6, 5X94, and 5X7B). (**B**) Representative structures of (**B**) N-SH2^{WT} and (**C**) N-SH2^{T42A} bound to the PD-1 pTyr^{223} (ITIM) peptide at the ends of respective trajectories. (**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. N = 3-5 independent titrations. All fold-changes and respective p-values can be found in Table S2.

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 was positioned slightly deeper into the ligand binding pocket of the mutant SH2 domain. The phenyl ring moved distinctly closer to residue 42 in the mutant simulations, presumably engaging in stabilizing hydrophobic interactions (**Figure 5D and Figure S5F**) (47, 48). Overall, the peptide main chain residues near the phosphotyrosine appeared closer to the body of the SH2 domain (**Figure 5D**). This likely alters side chain packing at the interface and may explain 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, was frequently shorter in N-SH2^{T42A} simulations than N-SH2^{WT} simulations (**Figure S5H**).

One striking difference between the N-SH2^{WT} and N-SH2^{T42A} simulations in the peptide-bound state was the positioning and movement of Lys^{55}. In crystal structures and in our N-SH2^{WT} simulations, the Lys^{55} ammonium group interacted with the phosphotyrosine phosphoryl group or engaged the phenyl ring in a cation-π interaction (**Figure 5B,E and Figure S6A**) (33, 42–45). In the N-SH2^{T42A} simulations, the phosphoryl group rotated away from Lys^{55} and engaged the BC loop and Arg^{32}, thereby liberating the Lys^{55} side chain (**Figure 5C,E and Figure S6A**). This shift altered 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 paired with the Asp^{40} side chain (**Figure S6C**). For the N-SH2^{T42A} simulations with the PD-1 pTyr^{223} peptide, we observed substantial 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**) (46). 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**) (49, 50). Furthermore, Vav-family SH2 domains are known to prefer a +2 Glu on their ligands (51), further corroborating a role for SHP2 Lys^{55} in substrate selectivity in an T42A-context.

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 phosphopeptides with and without a +2 Glu (**Figure 5H**). Many SH2 domains have an arginine at position 55 (**Figure S4B**). This 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 these interactions with arginine would likely be tighter than analogous interactions with Lys^{55} and may persist in a T42A context. Indeed, the K55R mutation, on its own, enhanced binding to all measured phosphopeptides, with an effect ranging from 2-fold to 7-fold (**Figure 5H, Figure S6E** *gray vs. cyan,* **and Table S2**). As discussed earlier, the T42A mutation on its own enhanced binding anywhere from 4-fold to 90-fold (**Figure 5H, Figure S6E** *gray vs. red***,** and **Figure 2**).

For peptides with a +2 Glu, the effect of the T42A mutation was large in isolation, but was substantially diminished in presence of a K55R mutation (**Figure 5H, Figure S6F, and Table S2**). For example, for the MILR1 pTyr 338 peptide, the 90-fold enhancement of binding affinity caused by the T42A mutation was attenuated to 28-fold in a K55R background. For the PD-1 pTyr 223 peptide, which also has a +2 Glu, the T42A mutation alone enhanced binding 28-fold, but the effect of this mutation dropped to 4.3-fold for in the presence of K55R. The Gab2 pTyr 614 peptide has a similar sequence to the PD-1 pTyr 223 peptide, but with an Ala at the +2 position. For this peptide, the T42A mutation enhanced binding 20-fold, but in contrast to the PD-1 and MILR1 peptides, this effect was actually enhanced to 43-fold in the presence of the K55R mutation (**Figure 5H, Figure S6F, and Table S2**). For peptides lacking a +2 Glu that showed a small enhancement in binding affinity upon the T42A mutation, such as IRS-1 pTyr 1179 and Imhof-9, the T42A effect was not impacted by the

K55R mutation (**Figure S6E**).

These experiments strongly suggest that Thr^{42} and Lys^{55} are energetically coupled (**Figure S6F**), and that the dramatic enhancement in binding affinity of N-$SH2^{T42A}$ for peptides with a +2 Glu is, at least in part, dependent on the presence of Lys^{55}. The sequence similarity between the PD-1 pTyr 223 and Gab2 pTyr 614 peptides suggests that other T42A-specific sequence preferences, independent of Lys^{55} and a +2 Glu, can also contribute to enhanced binding. These include the T42A preference for a -2 Val described above (**Figure 4B**). 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, 39), 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.

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 an important role in activating SHP2 by localizing the N-SH2 domain to a binding site for which the N-SH2 otherwise has a weak affinity (52).

Previous reports showed that $SHP2^{T42A}$ had enhanced activity when compared to wild-type SHP2 under saturating phosphopeptide concentration (14, 17). However, these experiments did not compare activation across multiple peptides. Thus, we measured activation of $SHP2^{T42A}$ using our panel of phosphopeptides, allowing us to ascertain peptide-specific effects on the activation of $SHP2^{T42A}$. For some peptides, the T42A mutation strongly shifted the EC_{50} 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 $N\text{-}SH2^{T42A}$ over $N\text{-}SH2^{WT}$ (**Figure 6G**, *large bubbles*) also showed the largest enhancement in activation (**Figure 6G**, *distance from dotted line*). These results demonstrate, for the first time, that the T42A mutation can sensitize SHP2 to specific activating ligands over others by altering N-SH2 binding affinity and specificity.

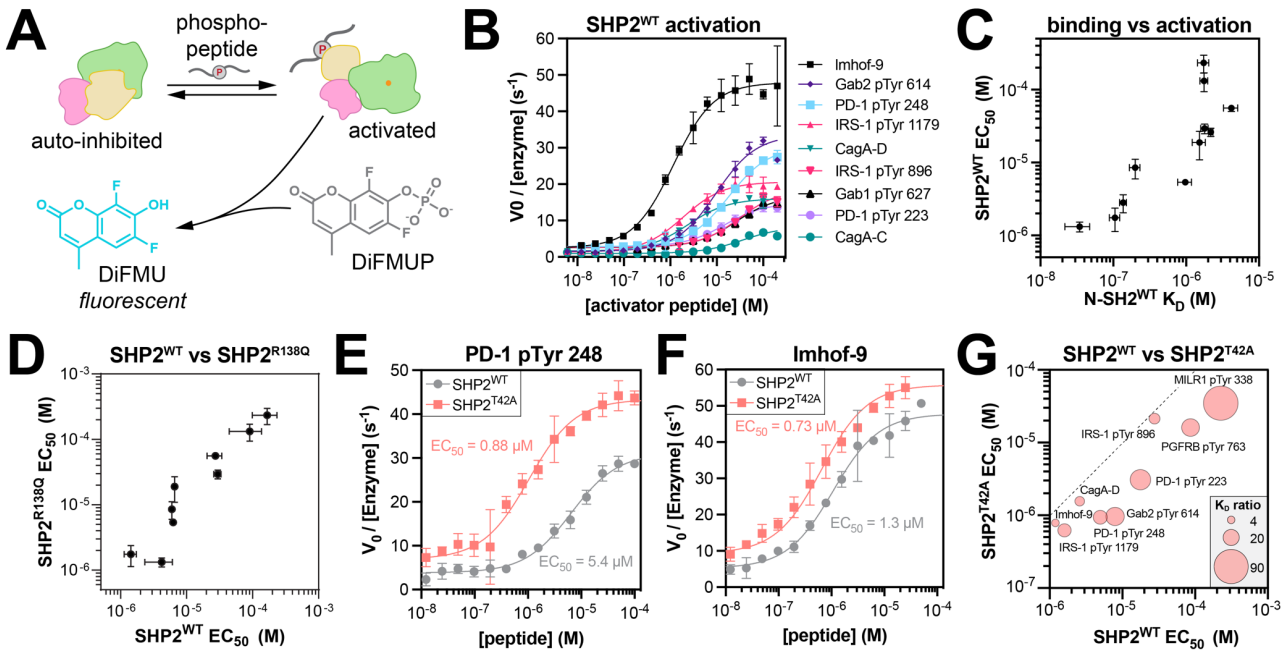

Figure 6. T42A-dependent changes in the activation of full-length SHP2. (**A**) Measurement of SHP2 activation by phosphopeptides. (**B**) Representative activation curves for $SHP2^{WT}$ (N = 3-17). (**C**) Correlation between the EC_{50} of $SHP2^{WT}$ activation by phosphopeptides and the K_D of those phosphopeptides for the N-$SH2^{WT}$ domain (N = 3-4 for K_D values, N = 3-17 for EC_{50} values). (**D**) Activation EC_{50} values for $SHP2^{WT}$ versus $SHP2^{R138Q}$ (N = 3-17 for $SHP2^{WT}$, N = 3-5 for $SHP2^{R138Q}$). (**E**) Comparison of $SHP2^{WT}$ and $SHP2^{T42A}$ activation by PD-1 pTyr 248 (N = 3-4). (**F**) Comparison of $SHP2^{WT}$ and $SHP2^{T42A}$ activation by Imhof-9 (N = 6-17). (**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-$SH2^{WT}$ vs N-$SH2^{T42A}$. 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 and p-values can be found in Table S5.

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 SHP2-interacting protein of interest and a constitutively active variant of the tyrosine kinase c-Src, to ensure phosphorylation of our interacting protein of interest. We chose Gab1, Gab2, and PD-1 as our proteins of interest, as these proteins play important roles in SHP2-relevant signaling pathways (28, 33, 53–56). For all three proteins, we found that $SHP2^{T42A}$ co-immunoprecipitated more with the interacting protein than $SHP2^{WT}$ (**Figure 7A and Figure S7A,B**). Anti-phosphotyrosine staining confirmed that these immunoprecipitated proteins were phosphorylated. We also examined whether the enhanced co-immunoprecipitation observed for $SHP2^{T42A}$ was indeed the result of increased intrinsic N-SH2 binding affinity, as opposed to an increased occupancy of the open conformation of the full-length protein, which exposes the SH2 domains. To test this, we compared $SHP2^{E76K}$, which predominantly occupies the open confirmation, to the $SHP2^{T42A+E76K}$ double mutant. More Gab2 was co-immunoprecipitated with $SHP2^{T42A+E76K}$ than with $SHP2^{E76K}$ (**Figure 7A**), suggesting that T42A and E76K enhance binding through non-redundant mechanisms.

Next, we assessed whether enhanced SHP2 binding caused by the T42A mutation impacts downstream 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 (59). Thus, we transfected $SHP2^{WT}$ or $SHP2^{T42A}$ into HEK 293 cells, along with Gab1 or Gab2. Cells were stimulated with epidermal growth factor (EGF), and Erk phosphorylation was analyzed as a marker of MAPK activation downstream of the EGF receptor. For the Gab2 experiment, we included $SHP2^{E76K}$, which should have high levels of phospho-Erk, even in absence of stimulation. We observed a stimulation-time-dependent increase then decrease in phospho-Erk levels for all three mutants. Notably, the response was stronger and longer for $SHP2^{T42A}$ than for $SHP2^{WT}$ samples (**Figure 7B and Figure S7C,D**). A previous study using the isolated $N\text{-}SH2^{T42A}$ and $N\text{-}SH2^{WT}$ domains showed that the T42A mutation increased co-immunoprecipitation of growth factor signaling proteins from HeLa cell lysates, but effects on signaling were not directly assessed (60). Our results directly demonstrate that tighter binding to phosphoproteins by full-length $SHP2^{T42A}$ can enhance downstream Ras/MAPK signaling.

Since SHP2 is a ubiquitously expressed protein, we investigated the functional effects of the T42A mutation in different cellular context, namely in T cells. We prepared SHP2 knock-out Jurkat T cells (**Figure S7E**), and reconstituted them with either $SHP2^{WT}$, $SHP2^{T42A}$ or $SHP2^{E76K}$. The wild-type and mutant SHP2-expressing cells were labeled with different concentrations of the fluorescent dye CellTrace Violet and pooled (**Figure S7F**). Then, the cells were stimulated with a CD3-specific antibody to activate T cell receptor-CD3 complexes, and signal transduction was evaluated by measuring phospho-Erk levels using flow cytometry (**Figure 7C**). In agreement with our EGF stimulation experiments in HEK 293 cells, phospho-Erk levels in T cells were higher with $SHP2^{T42A}$ and $SHP2^{E76K}$ when compared with $SHP2^{WT}$ (**Figure 7D and Figure S7G**). Next, we studied the effects of the T42A mutation on full T cell activation, as measured by expression of the activation marker CD69 (**Figure 7E and Figure S7H**). $SHP2^{T42A}$ showed enhanced T cell activation at lower doses of CD3 stimulation. Interestingly, $SHP2^{E76K}$ had an intermediate effect between $SHP2^{WT}$ and $SHP2^{T42A}$. SHP2 can participate in signaling pathways that both promote T cell activation (e.g. by activating the MAPK pathway) and suppress T cell activation (e.g. through the co-inhibitory receptor PD-1) (61, 62). Thus, we hypothesize that the reduced activation by $SHP2^{E76K}$ when compared with $SHP2^{T42A}$ may be a consequence of $SHP2^{E76K}$ tapping into both positive and negative regulatory pathways, without the need for SH2 ligand binding. By contrast, $SHP2^{T42A}$ may only be able to engage in negative regulatory pathways when an appropriate SH2 binding partner is presented.

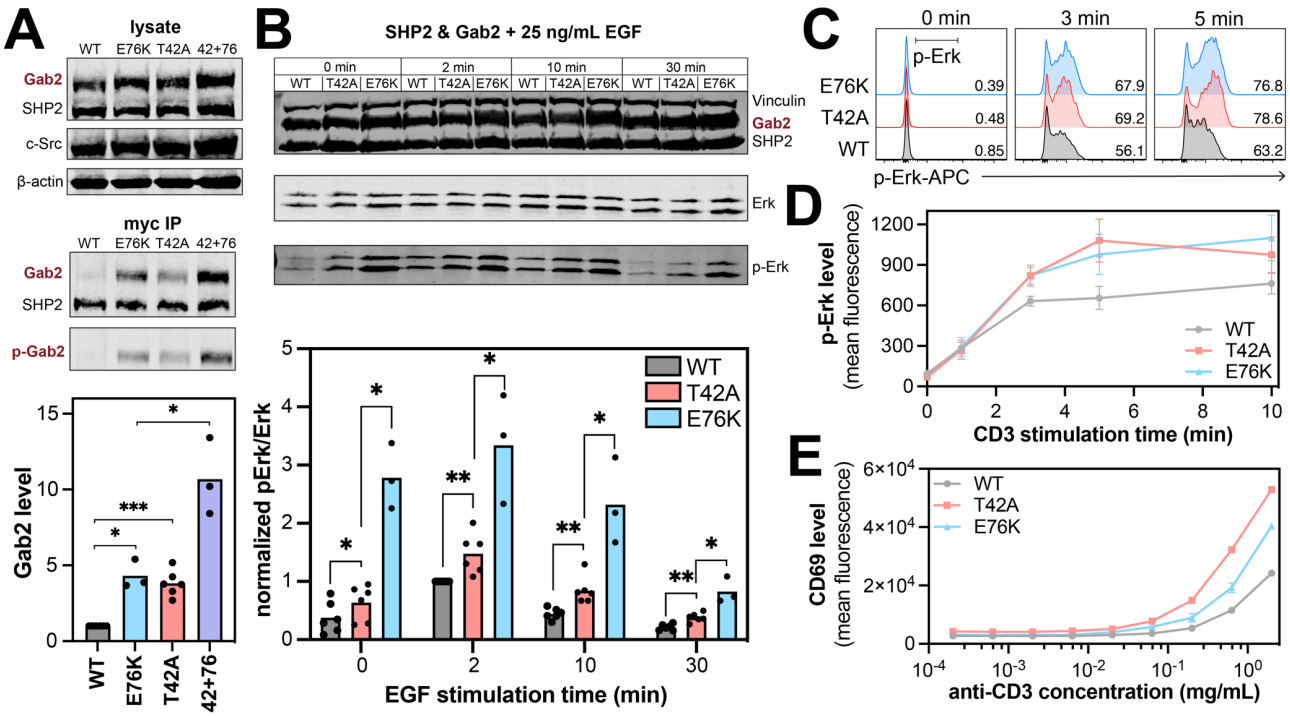

Figure 7. Enhanced protein-protein interactions and downstream signaling by SHP2^{T42A}. (**A**) Gab2 co-immunoprecipitation with SHP^{WT}, SHP2^{E76K}, SHP2^{T42A} and SHP2$^{T42A+E76K}$ (42+76). Relative Gab2 co-immunoprecipitation levels are normalized for expression and shown in the bar graph. (**B**) Phospho-Erk (p-Erk) levels in response to EGF stimulation in cells expressing Gab2 with either SHP2^{WT}, SHP2^{T42A} SHP2^{E76K} (N = 3-6). The bar graphs below the blots show p-Erk levels, normalized to total Erk levels, relative to the highest p-Erk signal in the SHP2^{WT} time course (2 minutes). For all bar graphs, a paired, one-tailed t-test was used to test for significance (* denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$). (**C**) Representative histograms for Erk phosphorylation responses in wild-type, T42A, or E76K SHP2-expressing Jurkat variants stimulated with anti-CD3 antibody for the indicated time. The black bar in the first panel depicts the gate used to define the pErk+ population. (**D**) Geometric mean fluorescence intensity for phosho-Erk levels in SHP2^{WT}, SHP2^{T42A} or SHP2^{E76K}-expressing SHP2-deficient Jurkat variants (N = 3 stimulation replicates). (**E**) Geometric mean fluorescence intensity for CD69 levels in SHP2^{WT}, SHP2^{T42A} or SHP2^{E76K}-expressing SHP2-deficient Jurkat variants (N = 2 stimulation replicates). The data in panels **D** and **E** are representative of 5-6 biological replicates.

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 well-studied mutations in SHP2 disrupt auto-inhibitory interactions between the N-SH2 and PTP domain, thereby hyperactivating the enzyme and enhancing 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, 40, 60). Our 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 (34). 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. Of note, the appearance of +2 Glu and -2 Val preferences in our N-SH2^{T42A} specificity screens were present in a sequence logo generated from previous peptide microarray data (14), but as those findings were subtle, the N-SH2^{T42A} binding profile was reported as similar to N-SH2^{WT} in that study.

We have also 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. We show that this enhanced binding and activation translates to downstream effects, such as increased intensity and duration of MAPK signaling in two different cell lines. This is in agreement with an affinity purification mass spectrometry study comparing SHP2 N-SH2^{WT} and N-SH2^{T42A} which found mutation-dependent changes in their interaction networks in mammalian cells (60). In that study, N-SH2^{T42A} showed increased interactions with growth factor signaling proteins, including those involved in MAPK signaling. Our comparisons of SHP2^{T42A} with SHP2^{E76K} suggest that these mutations alter SHP2 function through distinct mechanisms. 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.

Much remains unknown about the other mutations explored in this study. To our knowledge, no biochemical or cell biological characterization has previously been reported for the L43F mutation (15). Here, we showed a mild increase in basal activity of full-length SHP2^{L43F} and a corresponding decrease in melting temperature, suggesting that this mutation slightly destabilizes the auto-inhibited state. L43F also causes a slight increase in binding affinity to phosphopeptides. The T52S mutation has previously been reported to change binding affinity to Gab2, consistent with our data (63). We did not observe any large changes in sequence specificity, with the exception of a preference for smaller Ala over bulkier residues at the +1 position. 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 severely disrupts phosphopeptide binding to the C-SH2 domain. The molecular basis for the pathogenic effects of the E139D mutation remains elusive. While many studies have addressed its binding affinity and specificity, their results are ambiguous (14, 17, 60). Our high-throughput screens indicate weakened binding of 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, 22).

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 (64–66). This has been demonstrated most clearly for protein kinases, where mutations have been identified that alter substrate specificity (64). It is noteworthy that not all specificity-determining residues in kinases are located directly in the ligand/substrate binding pocket, raising the possibility that distal mutations may allosterically alter sequence specificity (67, 68). A similar paradigm has been suggested for SH2 domains, where distal mutations may rewire interactions, however the position corresponding to Thr^{42} in SHP2 has not been implicated as a determinant of specificity (67). 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. Notably, mutations at the analogous Thr residue in other SH2 domains have also been implicated in disease, substantiating the significance of this residue (69–71). More broadly, our results illustrate 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

Key resources, including cell lines, plasmids, oligonucleotide primers, peptides, and proteins, are listed in **Table S6**. Detailed materials and methods can be found in the supplementary information appendix. These include details of protein and peptide production, binding and activity measurements, bacterial peptide display screens, cell culture experiments, and molecular dynamics simulations.

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. **Funding:** This research was funded by NIH grant GM138014 to NHS, NIH grant AI175301 to WLL, and an Innovation Award from the Praespero Foundation to WLL. 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.

Author contributions: AEV conceived, designed, performed, analyzed, and interpreted the experiments; performed statistical analysis; and wrote the manuscript. RV conceived, designed, performed, analyzed, and interpreted the experiments and wrote the manuscript. CAC, ZJ, PAH, and WLL performed and analyzed the experiments and edited the manuscript. DK performed and analyzed the MD simulations and edited the manuscript. NHS designed, analyzed, and interpreted the experiments, analyzed the MD simulations, and wrote the manuscript.

Competing interests: The authors declare no competing interests.

Data and materials 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 a Dryad repository (DOI: 10.5061/dryad.msbcc2g41). 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.

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.

Figure S7. Cellular consequences of the T42A mutation.

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. R.J. Salmond, D.R. Alexander, SHP2 forecast for the immune system: fog gradually clearing. *Trends Immunol.* **27**, 154–160 (2006).
2. J. Lauriol, F. Jaffré, M.I. Kontaridis, The role of the protein tyrosine phosphatase SHP2 in cardiac development and disease. *Semin. Cell Dev. Biol.* **37**, 73–81 (2015).
3. M.D. Asmamaw, X.-J. Shi, L.-R. Zhang, H.-M. Liu, A comprehensive review of SHP2 and its role in cancer. *Cell Oncol (Dordr)* **45**, 729–753 (2022).
4. K.S. Grossmann, M. Rosário, C. Birchmeier, W. Birchmeier, "The tyrosine phosphatase shp2 in development and cancer." in *Advances in Cancer Research* (Advances in cancer research, 2010), pp. 53–89.
5. M. Tartaglia, B.D. Gelb, Noonan syndrome and related disorders: genetics and pathogenesis. *Annu. Rev. Genomics Hum. Genet.* **6**, 45–68 (2005).
6. B. Marino, M.C. Digilio, A. Toscano, A. Giannotti, B. Dallapiccola, Congenital heart diseases in children with Noonan syndrome: An expanded cardiac spectrum with high prevalence of atrioventricular canal. *J. Pediatr.* **135**, 703–706 (1999).
7. M. Tartaglia, *et al.*, Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. *Nat. Genet.* **34**, 148–150 (2003).
8. M.L. Loh, *et al.*, Mutations in PTPN11 implicate the SHP-2 phosphatase in leukemogenesis. *Blood* **103**, 2325–2331 (2004).
9. M. Bentires-Alj, *et al.*, Activating mutations of the noonan syndrome-associated SHP2/PTPN11 gene in human solid tumors and adult acute myelogenous leukemia. *Cancer Res.* **64**, 8816–8820 (2004).
10. P. Hof, S. Pluskey, S. Dhe-Paganon, M.J. Eck, S.E. Shoelson, Crystal structure of the tyrosine phosphatase SHP-2. *Cell* **92**, 441–450 (1998).
11. M. Tartaglia, *et al.*, Diversity and functional consequences of germline and somatic PTPN11 mutations in human disease. *Am. J. Hum. Genet.* **78**, 279–290 (2006).
12. T. Kaneko, R. Joshi, S.M. Feller, S.S. Li, Phosphotyrosine recognition domains: the typical, the atypical and the versatile. *Cell Commun. Signal.* **10**, 32 (2012).
13. P. Cohen, The role of protein phosphorylation in human health and disease. . *European Journal of Biochemistry* **268**, 5001–5010 (2001).
14. S. Martinelli, *et al.*, 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 (2008).
15. C.G. Weismann, *et al.*, PTPN11 mutations play a minor role in isolated congenital heart disease. *Am. J. Med. Genet. A* **136**, 146–151 (2005).
16. C.P. Kratz, *et al.*, The mutational spectrum of PTPN11 in juvenile myelomonocytic leukemia and Noonan syndrome/myeloproliferative disease. *Blood* **106**, 2183–2185 (2005).
17. H. Keilhack, F.S. David, M. McGregor, L.C. Cantley, B.G. Neel, Diverse biochemical properties of Shp2 mutants. Implications for disease phenotypes. *J. Biol. Chem.* **280**, 30984–30993 (2005).
18. T. Kaneko, *et al.*, Loops govern SH2 domain specificity by controlling access to binding pockets. *Sci. Signal.* **3**, ra34 (2010).
19. M. Anselmi, *et al.*, Structural Determinants of Phosphopeptide Binding to the N-Terminal Src Homology 2 Domain of the SHP2 Phosphatase. *J. Chem. Inf. Model.* **60**, 3157–3171 (2020).
20. R.A.P. Pádua, *et al.*, Mechanism of activating mutations and allosteric drug inhibition of the phosphatase SHP2. *Nat. Commun.* **9**, 4507 (2018).

21. J.R. LaRochelle, *et al.*, Structural reorganization of SHP2 by oncogenic mutations and implications for oncoprotein resistance to allosteric inhibition. *Nat. Commun.* **9**, 4508 (2018).

22. W. Qiu, *et al.*, Structural insights into Noonan/LEOPARD syndrome-related mutants of protein-tyrosine phosphatase SHP2 (PTPN11). *BMC Struct. Biol.* **14**, 10 (2014).

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

24. M. Hidaka, Y. Homma, T. Takenawa, Highly conserved eight amino acid sequence in SH2 is important for recognition of phosphotyrosine site. *Biochem. Biophys. Res. Commun.* **180**, 1490–1497 (1991).

25. A. Serbina, A.C. Bishop, Quantitation of autoinhibitory defects in pathogenic SHP2 mutants by differential scanning fluorimetry. *Anal. Biochem.* doi: 10.1016/j.ab.2023.115300 115300 (2023).

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

27. G.S. Kapoor, Y. Zhan, G.R. Johnson, D.M. O'Rourke, 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 (2004).

28. H. Gu, J.C. Pratt, S.J. Burakoff, B.G. Neel, 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 (1998).

29. M.R. Kuhné, *et al.*, Dephosphorylation of insulin receptor substrate 1 by the tyrosine phosphatase PTP2C. *J. Biol. Chem.* **269**, 15833–15837 (1994).

30. T. Okazaki, A. Maeda, H. Nishimura, T. Kurosaki, T. Honjo, 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 (2001).

31. H. Higashi, *et al.*, SHP-2 tyrosine phosphatase as an intracellular target of Helicobacter pylori CagA protein. *Science* **295**, 683–686 (2002).

32. D. Imhof, *et al.*, 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 (2006).

33. M. Marasco, *et al.*, Molecular mechanism of SHP2 activation by PD-1 stimulation. *Sci. Adv.* **6**, eaay4458 (2020).

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

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

36. L. Rönnstrand, *et al.*, 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 (1999).

37. K. Hitomi, *et al.*, An immunoglobulin-like receptor, Allergin-1, inhibits immunoglobulin E-mediated immediate hypersensitivity reactions. *Nat. Immunol.* **11**, 601–607 (2010).

38. J.P. O'Shea, *et al.*, pLogo: a probabilistic approach to visualizing sequence motifs. *Nat. Methods* **10**, 1211–1212 (2013).

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

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

41. M. Anselmi, J.S. Hub, An allosteric interaction controls the activation mechanism of SHP2 tyrosine phosphatase. *Sci. Rep.* **10**, 18530 (2020).

42. C.H. Lee, *et al.*, Crystal structures of peptide complexes of the amino-terminal SH2 domain of the Syp tyrosine phosphatase. *Structure* **2**, 423–438 (1994).

43. Y. Zhang, *et al.*, Simultaneous binding of two peptidyl ligands by a SRC homology 2 domain. *Biochemistry* **50**, 7637–7646 (2011).

44. Y. Liu, *et al.*, Structural basis for the regulatory role of the PPxY motifs in the thioredoxin-interacting protein TXNIP. *Biochem. J.* **473**, 179–187 (2016).

45. T. Hayashi, *et al.*, Differential Mechanisms for SHP2 Binding and Activation Are Exploited by Geographically Distinct Helicobacter pylori CagA Oncoproteins. *Cell Rep.* **20**, 2876–2890 (2017).

46. B.A. Liu, *et al.*, The SH2 domain-containing proteins in 21 species establish the provenance and scope of phosphotyrosine signaling in eukaryotes. *Sci. Signal.* **4**, ra83 (2011).

47. T. Kaneko, *et al.*, Superbinder SH2 domains act as antagonists of cell signaling. *Sci. Signal.* **5**, ra68 (2012).

48. G.D. Martyn, *et al.*, Engineered SH2 domains for targeted phosphoproteomics. *ACS Chem. Biol.* **17**, 1472–1484 (2022).

49. B. Wu, *et al.*, Identification and structural basis for a novel interaction between Vav2 and Arap3. *J. Struct. Biol.* **180**, 84–95 (2012).

50. RCSB PDB - 7RNV: SH2 domain of guanine nucleotide exchange factor Vav2 in complex with an actin peptide with phosphorylated tyrosine 53, Available at https://www.rcsb.org/structure/7RNV. Accessed 21 April 2023.

51. Z. Songyang, *et al.*, 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 (1994).

52. N. Patsoukis, *et al.*, Interaction of SHP-2 SH2 domains with PD-1 ITSM induces PD-1 dimerization and SHP-2 activation. *Commun. Biol.* **3**, 128 (2020).

53. A. Montagner, *et al.*, A novel role for Gab1 and SHP2 in epidermal growth factor-induced Ras activation. *J. Biol. Chem.* **280**, 5350–5360 (2005).

54. U. Schaeper, *et al.*, Coupling of Gab1 to c-Met, Grb2, and Shp2 mediates biological responses. *J. Cell Biol.* **149**, 1419–1432 (2000).

55. Time-resolved phosphoproteomics reveals scaffolding and catalysis-responsive patterns of SHP2-dependent signaling | eLife, Available at https://elifesciences.org/articles/64251. Accessed 22 September 2021.

56. X. Xu, *et al.*, PD-1 and BTLA regulate T cell signaling differentially and only partially through SHP1 and SHP2. *J. Cell Biol.* **219**, (2020) doi:10.1083/jcb.201905085.

57. Y.-N.P. Chen, *et al.*, Allosteric inhibition of SHP2 phosphatase inhibits cancers driven by receptor tyrosine kinases. *Nature* **535**, 148–152 (2016).

58. R.J. Nichols, *et al.*, RAS nucleotide cycling underlies the SHP2 phosphatase dependence of mutant BRAF-, NF1- and RAS-driven cancers. *Nat. Cell Biol.* **20**, 1064–1073 (2018).

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

60. P.J. Müller, *et al.*, 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 (2013).

61. T.V. Nguyen, Y. Ke, E.E. Zhang, G.-S. Feng, Conditional deletion of Shp2 tyrosine phosphatase in thymocytes suppresses both pre-TCR and TCR signals. *J. Immunol.* **177**, 5990–5996 (2006).

62. T. Yokosuka, *et al.*, Programmed cell death 1 forms negative costimulatory microclusters that directly inhibit T cell receptor signaling by recruiting phosphatase SHP2. *J. Exp. Med.* **209**, 1201–1217 (2012).

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

64. P. Creixell, *et al.*, Kinome-wide decoding of network-attacking mutations rewiring cancer signaling. *Cell* **163**, 202–217 (2015).

65. A. Lundby, *et al.*, Oncogenic mutations rewire signaling pathways by switching protein recruitment to phosphotyrosine sites. *Cell* **179**, 543-560.e26 (2019).

66. X. Mo, *et al.*, Systematic discovery of mutation-directed neo-protein-protein interactions in cancer. *Cell* **185**, 1974-1985.e12 (2022).

67. P. Creixell, *et al.*, Unmasking determinants of specificity in the human kinome. *Cell* **163**, 187–201 (2015).

68. D. Bradley, *et al.*, Sequence and Structure-Based Analysis of Specificity Determinants in Eukaryotic Protein Kinases. *Cell Rep.* **34**, 108602 (2021).

69. E.D. Renner, *et al.*, Novel signal transducer and activator of transcription 3 (STAT3) mutations, reduced T(H)17 cell numbers, and variably defective STAT3 phosphorylation in hyper-IgE syndrome. *J. Allergy Clin. Immunol.* **122**, 181–187 (2008).

70. M.J. Kiel, *et al.*, Integrated genomic sequencing reveals mutational landscape of T-cell prolymphocytic leukemia. *Blood* **124**, 1460–1472 (2014).

71. Q. Xiang, *et al.*, Autosomal dominant hyper IgE syndrome from a single centre in Chongqing, China (2009-2018). *Scand. J. Immunol.* **91**, e12885 (2020).