High-throughput biochemical phenotyping of SHP2 variants reveals the molecular basis of diseases and allosteric drug inhibition

Albert A Lee[1], Daniel A Mokhtari[2,3], Emily D Egan[4], Stephen C Blacklow[4,5,6], Daniel Herschlag[‡,2], Polly M Fordyce[‡,1,7,8,9].

[1]Department of Genetics, Stanford University, Stanford, CA 94305
[2]Department of Biochemistry, Stanford University, Stanford, CA 94305
[3]Velocity Bio, San Carlos, CA 94070
[4]Department of Biological Chemistry and Molecular Pharmacology, Blavatnik Institute, Harvard Medical School, Cambridge, MA 02115
[5]Department of Cancer Biology, Dana Farber Cancer Institute, Boston, MA 02215
[6]Broad Institute of Harvard and MIT, Cambridge, MA 02142
[7]Department of Bioengineering, Stanford University, Stanford, CA 94305
[8]Sarafan ChEM-H, Stanford University, Stanford, CA 94305
[9]Chan Zuckerberg Biohub San Francisco, San Francisco, CA 94110
[‡]Co-corresponding authors.

Abstract

Interpreting clinical and functional consequences of genetic variants remains challenging due to limited quantitative biochemical data at scale. We applied high-throughput microfluidic enzyme kinetics to profile 190 clinical variants of SHP2, a protein tyrosine phosphatase linked to developmental disorders and cancers. Through >300,000 reaction progress curves, we derived kinetic and thermodynamic parameters quantifying variant effects on catalysis, autoinhibition, stability, phosphopeptide binding, and drug responses. This multidimensional dataset reveals that dysregulated autoinhibition, rather than altered stability or catalysis, predominantly determines SHP2-associated pathogenesis. Thermodynamic modeling reveals that clinical-stage allosteric inhibitors preferentially stabilize a previously underappreciated, partially active conformation over the fully inactive state, leading to variant-dependent drug responses. Our high-throughput biochemical framework establishes a general approach to decipher the biochemical logic connecting protein variants to clinical outcomes.

Introduction

Understanding how disease-associated mutations affect protein function, drive pathology, and can be therapeutically addressed is a central goal of biomedical science (*1*–*5*). Beyond its clinical importance, studying disease mutations has long fueled key discoveries of fundamental principles of molecular and biological function (*6*, *7*). While next-generation sequencing has unveiled a vast reservoir of human protein-coding variants (*8*, *9*), comprehensive biochemical and biophysical characterization of these variants at scale remains limited. As a result, most variants are only cataloged at the level of sequence and clinical observation, with their functional and clinical significance remaining largely unresolved (*4*, *10*). While mutational scanning assays can provide clinically relevant information by estimating variant impact on function using cellular fitness, these composite readouts typically do not directly measure the distinct physical parameters required to understand complex disease mechanisms. Here, we address this gap by systematically measuring biochemical and biophysical properties of 190 clinically documented variants of the disease-relevant protein tyrosine phosphatase (PTP) SHP2, providing a functional atlas that links mutations to disease phenotypes, therapeutic responses, and fundamental aspects of SHP2 biochemistry.

SHP2 is a ubiquitously expressed PTP encoded by the *PTPN11* gene and plays an essential role in regulating cellular processes, including metabolism, cell growth, differentiation, and immune response (*11*–*15*). Consistent with this central role, mutations in SHP2 are a major cause of multiple diseases (*16*–*21*). Germline SHP2 mutations cause autosomal dominant developmental disorders and are responsible for 50% of Noonan syndrome (NS) and 90% of Noonan syndrome with multiple lentigines (NSML) cases, disorders affecting 1 in 1000 live births (*17*, *22*, *23*). Somatic SHP2 mutations are associated with various types of cancer, including 35% of juvenile myelomonocytic leukemia (JMML) cases (*21*).

Structurally, SHP2 consists of two SH2 domains (nSH2 and cSH2), a catalytic PTP domain, and an unstructured C-terminal tail (**Fig. 1A**). SHP2 regulation is canonically described via a two-conformational-state model in which the protein interconverts between a closed, inactive conformation and an open, active conformation. In the autoinhibited conformation, the nSH2 domain sterically occludes the PTP active site and prevents catalysis (*24*, *25*); in the open conformation, the SH2 domains are rotated away such that the catalytic site is fully exposed (*26*). The open state can be further stabilized by the SH2 domains binding to phosphotyrosine (pTyr) motifs on activated receptors or scaffold proteins (*24*, *27*–*31*). The importance of SHP2 in proliferative signaling makes it a compelling target for cancer therapy (*15*, *32*, *33*), and at least 14 SHP2 inhibitors have progressed to clinical trials (*32*–*37*). Nearly all clinical-stage SHP2 inhibitors exploit SHP2's native autoinhibitory mechanism. These compounds bind to a tunnel-like pocket at the interface of the SH2 and PTP domains, distal to the active site (*34*, *35*); binding is thought to stabilize the closed conformation, thereby potently and selectively inhibiting SHP2 relative to other PTPs.

Despite progress in understanding SHP2 pathology and function, over half of the >500 *PTPN11* missense variants in the ClinVar database are variants of uncertain significance (VUS), complicating genetic test interpretation and clinical management (*10*). Moreover, current data remain insufficient to define the biochemical basis of overlapping and divergent clinical phenotypes of SHP2 variants. For instance, NS and NSML share clinical features and cancer predisposition (*38*–*40*), yet NS mutations enhance SHP2 basal catalytic activity, while NSML mutations reduce it (*41*–*44*). Prior studies show that some NSML variants also reduce autoinhibition (*25*, *44*), a defect shared with NS variants, suggesting that disease phenotypes may depend on an interplay between catalytic and conformational effects. However, no systematic comparison has defined how catalytic and conformational effects combine to account for NS and NSML phenotypes. Similarly, it is unclear why both hyperactive and hypoactive SHP2 variants are associated with cancer (*15*, *33*, *45*), and why some mutations are cancer-specific while others overlap with developmental disorders (*46*, *47*). In-depth knowledge of the biochemical consequences of disease-causing variants at scale is needed to resolve these paradoxes.

Our understanding of SHP2 sensitivity towards clinical-stage inhibitors is also critically limited, as the inhibitor sensitivity of most SHP2 variants remains unknown. Moreover, while detailed studies of several cancer-

associated mutants, including the canonical E76K, show that they shift the conformational ensemble toward the open state and confer resistance to the allosteric inhibitor SHP099 (*26*, *48*), it is unclear if this mechanism applies broadly across different variants and drug scaffolds. Addressing this limitation requires a systematic map of variant-specific drug sensitivity, coupled with quantitative biochemical analysis to resolve the mechanistic basis of these changes.

Here, we performed deep biochemical phenotyping on a library of 190 clinically documented SHP2 variants via HT-MEK, a microfluidic platform for high-throughput and quantitative biochemistry (*49*, *50*). We isolated and quantified distinct biochemical mechanisms by which mutations alter SHP2 function, including catalytic activity, stability, autoinhibition, and its response to physiological activators and clinical inhibitors, yielding ten kinetic and thermodynamic parameters that collectively define the biochemical phenotype of each variant. Analysis of this dataset revealed distinct biochemical signatures that predict pathogenicity and stratify variants by disease class, providing a biochemical basis that reconciles paradoxes in SHP2-associated pathologies. Profiling responses to three clinical-stage inhibitors uncovered a ~100-fold range in sensitivity across variants and revealed an unexpected trend that contradicts the canonical two-conformational-state (closed/open) model, in which mild loss of autoinhibition leads to highest drug sensitivity. These findings led us to a three-conformational-state model of allosteric inhibition that accurately accounts for all observed measurements of SHP2 inhibition and suggests strategies to enhance therapeutic potency. This work provides an in-depth mechanistic framework for SHP2 pathology and therapeutic strategies, establishing a scalable approach for causally linking genotype to molecular function, disease, and therapeutic response that can be applied to many additional clinically relevant proteins.

Results

HT-MEK enables high-throughput expression, purification, and activity assays for SHP2 variants

To systematically characterize impacts of human allelic variation on SHP2 biochemical function, we generated a library of 189 missense variants documented in ClinVar within SHP2's structured domains (residues 1-525, hereafter SHP2 or $SHP2_{FL}$) (**Fig. 1A; Supplementary Table 1**) (*10*). These mutations are distributed throughout the linear sequence and across the structured domains of SHP2 (**Figs. 1A**&**B**). The library includes variants classified as pathogenic (46), benign (5, including WT), of uncertain significance (VUS; 69), or with conflicting/less certain annotations (70) (**Fig. 1C**).

Mutations of SHP2 can alter different aspects of biochemical function, including thermodynamic stability, intrinsic catalytic activity (activity of the isolated PTP domain), autoinhibition, regulation by activating pTyr peptide, or inhibition by drugs (**Fig. 1D**). To separately quantify the impacts on each of these properties for all variants with high accuracy and precision, we used High-Throughput Microfluidic Enzyme Kinetics (HT-MEK) (*49*), a microfluidic platform that enables recombinant expression, purification, and quantitative functional characterization of 1000s of enzymes in parallel (**Fig. 1E**). In HT-MEK, microfluidic devices containing 1798 reaction chambers are aligned with printed plasmid arrays. Introduction of cell-free expression reagents enables parallel expression of a unique C-terminally EGFP-tagged enzyme variant per chamber. The expressed proteins are then immobilized on device surfaces pre-patterned with anti-EGFP antibody, enabling purification and quantification. Integrated on-chip valves protect the immobilized enzymes during reagent exchange and permit reactions to be initiated and terminated with precise temporal control, allowing iterative assays for high-throughput measurement of thermodynamic and kinetic parameters (**Fig. 1E**).

To test if HT-MEK can accurately measure kinetic constants of SHP2-mediated catalysis, we quantified hydrolysis of the fluorogenic substrate 6,8-difluoro-4-methylumbelliferyl phosphate (DiFMUP) by a control library consisting of $SHP2_{FL}$, a truncated construct containing only the catalytic domain ($SHP2_{CD}$), four active site variants ($D425E_{CD}$, $R465K_{CD}$, $Q506A_{CD}$, $T466A_{CD}$), and a catalytically dead variant ($C459S_{CD}$) (**Figs. S1**&**S2**). Measured activities for $C459S_{CD}$ were indistinguishable from background activity (**Figs. 1F**&**S3**). HT-MEK-derived Michaelis-Menten parameters for these constructs were consistent (within ~2-fold) with values

obtained from traditional plate-based assays (**Figs. 1F**&**G**) and within the range of literature values (**Fig. S4**) (*51*, *52*).

SHP2 variants in different structural regions substantially increase or decrease basal activity

Next, we applied HT-MEK to measure k_{cat}, K_M, and k_{cat}/K_M for DiFMUP hydrolysis for 190 $SHP2_{FL}$ allelic variants (**Fig. 2A**). Data were consistent across 10 devices (**Figs. S5**&**S6**), and we aggregated measurements across all replicates (median n = 19) to report median values for each variant (**Figs. 2B**, **S7**, **S8**).

Observed changes in k_{cat}/K_M within the $SHP2_{FL}$ construct ("basal activity") ranged from a 30-fold increase to a 30-fold decrease in k_{cat}/K_M (**Fig. 2B**). Impacts of different substitutions at the same position varied in magnitude but generally shifted k_{cat}/K_M in the same direction (**Fig. S8**). Most variants (120/189) significantly increased activity, while 26 significantly decreased activity and 43 did not show a statistically significant change from WT (via bootstrap hypothesis testing) (**Fig. 2B**). The large number of replicates allowed us to confidently detect even subtle effects, including k_{cat}/K_M decreases and increases of 26% and 27%, respectively (**Fig. 2B**). Mutations that increased the basal activity clustered around the interface between the PTP and the nSH2 domain (**Fig. 2C**), consistent with reduced autoinhibition. Mutations that reduced the basal activity were exclusively found in the PTP domain and concentrated around the active site, consistent with reduced intrinsic activity of the PTP domain itself (**Fig. 2C**), a model that we directly tested below.

SHP2 catalytic domain variants predominantly reduce intrinsic activity

Changes in basal activity upon mutation could result from altered phosphate ester hydrolysis of the catalytic PTP domain ("intrinsic activity") (**Fig. 2D**), altered amount of natively folded protein, or altered autoinhibitory regulation. To eliminate effects from changes in autoinhibition, we measured the impact of the 94 unique mutations within the isolated catalytic PTP domain ($SHP2_{CD}$) on DiFMUP hydrolysis (k_{cat}, K_M, and k_{cat}/K_M) (**Figs. S9-S12**). In contrast to $SHP2_{FL}$, variants in the isolated domain were predominantly inactivating. Thirty-six variants significantly reduced k_{cat}/K_M (1.3- to >800-fold reductions), while none increased k_{cat}/K_M more than two-fold (**Fig. 2E**). As expected, deactivating mutations clustered around the active site (**Fig. 2F**). However, three deactivating mutations were found >20 Å from the catalytic cysteine (Q347R, I479T, P491H). Notably, I479 and P491 reside on helices flanking a conserved loop known to be allosterically coupled to the active site conformation in the homolog PTP1B (**Fig. S13**) (*53*, *54*), highlighting HT-MEK's ability to detect allosteric mutations and potential allosteric sites. The reduced activity of the remaining distal variant, Q347R, is instead fully accounted for by loss of natively folded protein (see below).

Global thermodynamic destabilization of the SHP2 native state does not account for most observed activity changes

Enzymes must fold to be catalytically active, and it has been suggested that folding stability defects account for a majority of disease associated alleles (*55–57*). Measuring loss of enzymatic activity as a function of increasing chaotropic destabilizing agent (*e.g.* urea) can quantify impacts of mutations on folding stability (ΔG_{fold}) for enzymes that undergo a reversible, two-state folding transition (**Fig. S14A**). While plate-based measurements of $SHP2_{CD}$ DiFMUP hydrolysis as a function of urea show a fast, reversible two-state transition (**Fig. S14B**), HT-MEK measurements revealed a slow, irreversible loss of activity in addition to the reversible transition, likely due to interactions between the unfolded protein and the device surface (see **Supplementary Text 1**). Accounting for both processes allowed high-throughput quantification of $SHP2_{CD}$ variants' apparent native state stability ($\Delta G_{native,app}$) from HT-MEK data that agree with those measured in solution (estimated error < 10%) (**Figs. S14-S17**, **Supplementary Text 1**).

Of 95 $SHP2_{CD}$ variants, 88 were sufficiently catalytically active to allow reliable quantification of urea-dependent loss of activity. Of these, 25/88 significantly reduced stability and none were significantly stabilized ($p < 0.01$) (**Fig. S14F**). Destabilizing mutations clustered on one face of the catalytic pocket, adjacent to the phosphotyrosine recognition loop and Q loop (**Fig. S14G**). Only five variants (G268C, C333S, Q347R, R501K,

G503V) were less than 50% natively folded under our standard assay conditions (in the absence of urea) (**Fig. S14H**); these variants showed matching fold-decreases in natively-folded fraction and intrinsic activity (**Fig. S14I**), indicating that activity loss is fully accounted for by loss of native protein, including the observed distal effect of Q347R. Only three of the remaining 83 variants lose >20% of natively folded protein under urea-free assay conditions. Thus, the observed changes in intrinsic activity for most variants reflect direct effects on catalysis, not loss of natively folded protein (**Fig. S14I**).

Interpreting urea-dependent differences in activity for $SHP2_{FL}$ is challenging, as multidomain $SHP2_{FL}$ is not expected to follow simple two-state unfolding (**Supplementary Text 1**). Nevertheless, low concentrations of urea (1 M) primarily destabilize the PTP domain while leaving SH2 domains intact (*58*), making it possible to infer the stability of the PTP domain in $SHP2_{FL}$ (**Fig. S18A**). We estimated the PTP domain within all $SHP2_{FL}$ variants is >88% natively folded under normal assay conditions). In general, the PTP domain within $SHP2_{FL}$ is more stable than when isolated, as expected from stabilization via interdomain contacts (**Figs. S18C**&**D**). Overall, the slight change in fraction native (a median effect of 1.7% in $SHP2_{FL}$) cannot explain the much larger observed impacts on measured k_{cat}/K_M values for $SHP2_{FL}$ variants (**Fig. S18E**). Thus, these impacts are likely due to perturbations within the native state (*i.e.* altered autoinhibition).

Comparative substrate profiling identifies mutations impacting peptide recognition

DiFMUP is a convenient probe of catalytic activity but it lacks the full set of interactions formed by physiological peptide substrates (*52*). To determine whether mutations perturb recognition of peptide substrates, we compared DiFMUP hydrolysis with dephosphorylation of a phosphopeptide derived from a known SHP2 target (EGFRpY992, DADEpYLIPQQG) (*44*, *59*). In this assay, we quantified hydrolysis by detecting inorganic phosphate release via fluorescence increase upon binding to a labeled phosphate-binding protein (*60*). The EGFRpY992 hydrolysis k_{cat}/K_M values for both $SHP2_{FL}$ and $SHP2_{CD}$ constructs closely matched those for DiFMUP (r^2 = 0.81, **Figs. 2G**, **S19-S22**), indicating that most variants do not alter peptide recognition and suggesting that DiFMUP is a reliable proxy for general catalytic activity. Mutations at positions 279, 282, 503, 506, and 507 significantly reduced k_{cat}/K_M (residual |*z*-score| > 3) for EGFRpY992 relative to DiFMUP; all of these residues are located near the peptide-binding cleft, consistent with an effect on peptide substrate recognition (**Fig. 2H**) (*61*).

Most variants increase basal activity by disrupting autoinhibition

To deconvolve and quantify impacts of mutations on autoinhibition, we compared DiFMUP k_{cat}/K_M values for each variant within the $SHP2_{FL}$ and $SHP2_{CD}$ constructs. This comparison revealed four mechanistic categories (**Fig. 3A**): (1) WT-like variants with no change in either construct; (2) variants with reduced $SHP2_{CD}$ and a proportional reduction in $SHP2_{FL}$ activity, indicating decreased intrinsic activity; (3) variants with increased $SHP2_{FL}$ activity but no change in $SHP2_{CD}$ activity (including variants with mutations in the SH2 domains), consistent with impaired autoinhibition; and (4) variants with a greater decrease in $SHP2_{CD}$ than $SHP2_{FL}$ activity (or reduced $SHP2_{CD}$ activity with mildly increased $SHP2_{FL}$ basal activity), suggesting decreases in both intrinsic activity and autoinhibition.

The extent of autoinhibition determines the fraction of SHP2 molecules in the open, active conformation. As $SHP2_{CD}$ activity is approximately the same as the activity of the $SHP2_{FL}$ open state (*26*, *48*), we quantified autoinhibition by calculating the ratio of k_{cat}/K_M values for $SHP2_{CD}$ and $SHP2_{FL}$ for each mutation (**Fig. 3B**). By this measure, only 7.3% of WT SHP2 is in the open, active state (fraction active, f^0_{active}) (95% CI: 6.5–8.1%). Two variants previously shown to lessen autoinhibition, E76D and E76K, are substantially more open, with 26% (95% CI: 22–30%) and 100% (95% CI: 70–100%) f^0_{active}, respectively (**Fig. 3C**). All three values are consistent with prior NMR-based estimates (*48*).

Overall, the calculated f^0_{active} for SHP2 variants ranged from <5% to >99%, with most variants (121/185) significantly increasing f^0_{active} (**Fig. 3D**). These mutations predominantly clustered at the nSH2-PTP domain interface (**Fig. 3E**), consistent with the established role of this region in maintaining autoinhibition (*24*).

Mutations beyond this interface also significantly altered f^0_{active}, with most occurring within the SH2 domains (**Fig. 3E**) where they presumably perturb the SH2 domain conformational ensemble to weaken the autoinhibitory interaction (*24*, *28*, *30*). Since the PTP domain active site lies at the nSH2-PTP autoinhibitory interface, many mutations near the active site impair intrinsic catalysis while simultaneously relieving autoinhibition ("mixed effect" in **Fig. 3A**; **Fig. 3F**) such that they have little net impact on basal activity (**Figs. 3G**&**S23**). These mutants would be misclassified as "WT-like" when measuring basal catalysis alone, highlighting the value of acquiring multidimensional data.

Conformational state and activator affinity determine the sensitivity of SHP2 variants to activation

SHP2 autoinhibition is regulated by the binding of pTyr-containing peptide (activator) to the tandem SH2 domains in the open, active conformation, typically described by a two-state conformational selection model (**Fig. 3H**) (*27*). To quantify impacts of mutations on activator sensitivity, we measured rates of DiFMUP hydrolysis as a function of increasing concentrations of ppIRS1, a synthetic bisphosphorylated peptide derived from IRS1 that binds both the nSH2 and cSH2 domains of SHP2 (**Fig. 3I**) (*32*). Fitting activation dose-response curves for 188 SHP2 variants (median replicate n = 9) yielded half-maximal effective peptide concentrations (EC_{50} values) that varied by >100-fold (**Figs. 3J**, **S24-S26**).

Based on a model in which one ppIRS1 binds per SHP2 molecule (**Fig. 3H**), EC_{50} (fitted with Hill slope of 1) is quantitatively determined by: (1) the intrinsic fraction active of SHP2 in the absence of activating ligand (f^0_{active}), and (2) the affinity of the SH2 domains for the activator (K^A_d) (**Eqn. 1**, see **Supplementary Text 2**):

Eqn. 1: $$EC_{50} = \frac{1}{f^0_{active}} K^A_d$$

As most variants have mutations located far from the activator binding site, we expected that they would minimally impact affinity (K^A_d) such that the EC_{50} would simply be inversely related to f^0_{active}. Consistent with this expectation, EC_{50} for most variants (176/188) were generally captured by their f^0_{active} with a single globally-fitted K^A_d ($K^A_d = 2.77 \pm 0.16\ nM$, r^2 = 0.63; **Fig. 3K**); a subset of variants (12/188) deviated significantly from this expectation more than 3-fold (p < 0.01). Consistent with altered ligand binding affinity, 9/12 of the deviating mutations appeared within the SH2 domains and this group included T42A, a variant known to increase ppIRS1 binding affinity (**Figs. 3L**&**S27**) (*62*). Our measurements of activity, stability, autoinhibition, and activator sensitivity are all consistent with the canonical understanding of SHP2 biochemistry and regulation. This suite of data established a foundation for decoding the biochemical basis of SHP2 pathogenesis, as explored below.

Pathogenic variants converge on impaired autoinhibition and activator hypersensitivity

There is a basic expectation that variant-induced changes in biochemical functions ultimately drive disease phenotypes (*63*–*65*). To first test if measured biochemical parameters can distinguish pathogenic from benign SHP2 variants, we hierarchically clustered 179 variants based on five *z*-score-normalized biochemical features: basal activity, intrinsic activity, fraction active, stability, and activator EC_{50} (**Fig. 4A**). ClinVar-assigned pathogenic and benign variants (**Supplementary Table 1**) clustered separately, supporting the association of distinct biochemical profiles with pathogenicity. Notably, basal activity and stability—commonly emphasized in variant characterization and often assumed to drive disease (*55*–*57*)—were less effective predictors of SHP2 pathogenicity (**Fig. S28**). Instead, altered fraction active and activator EC_{50} were most strongly predictive of pathogenicity by leave-one-out cross-validation; indeed, using just these two features performed identically to using all five biochemical parameters (**Fig. S28**).

Overall, 42/45 pathogenic variants have significantly higher f^0_{active} than benign variants, consistent with impaired autoinhibition predominantly driving pathogenicity. The remaining three (T42A, I56V, and T468M) were significantly more sensitive to ligand-induced activation (**Fig. 4B**). Thus, pathogenic SHP2 variants converge on a shared biochemical phenotype: dysregulated autoinhibition achieved via either destabilization of

the autoinhibited state or enhanced sensitivity to activating ligands, regardless of their catalytic activity. This convergence suggests that the open conformation may drive pathogenesis primarily through its role as a signaling scaffold (*15*, *66–69*). Notably, AlphaMissense (*70*), a machine learning predictor for pathogenicity, identifies many variants with benign-level f^0_{active} and activator EC_{50} as pathogenic (**Fig. S29**). This tendency to overpredict pathogenicity has also been observed in prior evaluations (*71–73*).

Next, using the two most informative measurements identified above (f^0_{active} and activator EC_{50}), we applied a *k*-nearest neighbor classifier to predict likely clinical impacts of 139 SHP2 variants with less certain clinical annotations. This approach predicted pathogenicity for 1/3 "likely benign" variants (33%), 19/60 VUS (32%), 15/19 "conflicting" variants (79%), 10/14 "likely pathogenic" variants (71%), and 32/33 "pathogenic/likely pathogenic" variants (97%) (**Figs. 4B**&**S30**). Thus, a substantial proportion of VUS may have pathogenic potential and warrant further clinical investigation. Since the start of our study in 2022, only one variant (T73I) has been upgraded to "pathogenic" in ClinVar (originally classified as conflicting interpretations of pathogenicity); our model's prediction is consistent with this reclassification, highlighting the prognostic value of biochemical phenotyping (**Fig. 4B**).

Quantitative differences in biochemical signatures differentiate disease subtypes

Next, to assess if different disease subtypes have distinct biochemical signatures, we compared measured biochemical parameters for cases classified as NS, NSML, or cancer within the expert-curated NSEuroNet (www.nseuronet.com) and COSMIC databases (*74*), respectively (**Fig. 4C**). These databases record the number of cases and specific disease subtypes associated with each variant, enabling statistical analyses of biochemical features across clinical phenotypes. Compared to NS/NSML cases, cancer cases tend to carry variants with a higher fraction active and activator sensitivity (both $p < 0.01$) (**Figs. 4D**&**S31A**). This distinction may arise because strongly activated germline variants are embryonically lethal and therefore observed only as non-inherited somatic mutations (*75*, *76*); conversely, many weakly activated variants observed in developmental disorders may lack sufficient oncogenic drive to cause cancer. Consistent with this trend, NS/NSML patients carrying variants with higher fraction active show an increased leukemia risk (**Fig. S32**). Within cancers, blood cancer cases carry variants with greater fraction active and activator sensitivity than do solid cancer cases (both $p < 0.01$) (**Figs. 4D**&**S33A**), with JMML cases carrying variants showing the largest increases (**Fig. S33B**).

Previous observations of hypoactive variants in solid cancer cases have led to the proposal that SHP2, a well-established proto-oncogene, may also act as a tumor suppressor, suggesting dual roles in cancer (*15*, *45*). While we also identified a subset of cancer cases involving variants with decreased basal activity (3% of blood, 20% of solid cancer cases) (**Fig. S34**), our comprehensive phenotyping offers an alternative interpretation. Regardless of basal activity, most cancer cases carry variants with significantly increased fraction active or activator sensitivity (1244/1260 cases with variants of increased, and 75/87 cases with variants of reduced basal activity) (**Fig. S34**), suggesting that increased open-state propensity, common to variants with both increased and decreased basal activity, is the primary biochemical determinant of SHP2-related oncogenesis.

It has been reported that NS variants associate with elevated basal activity, while NSML variants associate with diminished basal activity due to lower intrinsic activity (*41–44*). Here, a larger dataset (72 NS and 11 NSML variants, as assigned in NSEuroNet) confirms that NS and NSML variants form distinct clusters in biochemical space, consistent with these prior observations (**Figs. 4E**&**S35**). Three variants (N200Y, R501K and Q510L) deviated from this pattern; further investigation revealed that N200Y was misannotated as NSML-associated and instead causes NS (NSEuroNet, personal communication). In addition, we suggest that Q510L and R501K, currently annotated as NS, are more likely NSML based on their biochemical characteristics (**Figs. 4E**&**S35**); supporting this hypothesis, the related variant Q510E is associated with NSML.

Overall, an increased propensity to open, driven by either higher fraction active or activator sensitivity, appears to drive pathogenesis for developmental disorders and cancer. A greater magnitude of this effect predisposes to cancer, and additional reductions in intrinsic activity further stratify specific disease subtypes (**Fig. 4F**). Thus,

our high-throughput, multi-dimensional biochemical pipeline can effectively identify functional abnormalities associated with disease types and propose likely disease outcomes for allelic variants.

Shifts in SHP2 conformational equilibrium alter allosteric inhibitor sensitivity

Allelic variation can impact the potency of drug inhibition and thus the choice of clinical treatment regimens. Here, we quantified the impacts of SHP2 variants on inhibition by three allosteric drugs advanced to clinical trials (TNO155, RMC-4630, and GDC-1971) (*33*–*36*). Crystal structures show that these compounds bind at the interface between the nSH2-cSH2 linker, cSH2 domain, cSH2-PTP linker, and PTP domain, with the nSH2 domain blocking the active site in a manner closely resembling the autoinhibited state (**Fig. 5A**) (*34*, *35*). These structures thus support a conformational selection mechanism in which these drugs stabilize the autoinhibited state to inhibit SHP2 (**Fig. S36A**) (*48*).

Because cellular inhibition occurs in the presence of activating ligands (*e.g.,* ppIRS1) that shift SHP2 toward the open, active conformation and reduce inhibitor potency (**Fig. S36B**) (*26*), we measured DiFMUP hydrolysis across 9 drug concentrations and 4 ppIRS1 concentrations. Fitting dose-response curves with 1:1 binding isotherms (with a Hill slope fixed at 1) yielded an IC_{50} at each ppIRS1 concentration per drug per variant (**Figs. 5B**, **S37-S50**); these values were consistent with plate reader measurements and correlated with drug binding affinity measured directly by microscale thermophoresis (**Fig. S37D**).

In the absence of ppIRS1, measured IC_{50} values varied by ~100-fold across SHP2 variants for each allosteric inhibitor, with many pathogenic variants exhibiting increased drug sensitivity relative to WT (**Figs. 5C**&**S50**). Under the canonical two-conformational-state inhibition model (**Fig. S36A**), changes in inhibitor IC_{50} could reflect differences in f^0_{active} and/or the affinity of the drug for the closed state (K^I_d) (**Supplementary Text 2**):

Eqn. 2: $$IC_{50} = \left(\frac{1}{1-f^0_{active}}\right)K^I_d$$

Many mutations with impacts on IC_{50} were far from the drug binding site (**Fig. S51**) and measured IC_{50} values were well-correlated for the three drugs despite their different structures (**Fig. S52**), suggesting most of the observed differences in IC_{50} arise from changes to f^0_{active} and not ligand binding affinity in the closed state (K^I_d).

Consistent with prior literature and an antagonistic binding model, increasing ppIRS1 reduced potency of all three allosteric inhibitors across all tested SHP2 variants (**Figs. 5D**&**S53**), and variants with a greater fraction active exhibited a stronger dependence of IC_{50} on ppIRS1 levels (**Fig. S53**) (*26*). Thus, in cancer cells with higher concentrations of SHP2-activating ligands (*77*, *78*), SHP2 will be less sensitive to these allosteric inhibitors, and this effect will be even more pronounced for cancer cells bearing variants with higher fraction active. Overall, the potency of allosteric inhibitors depends strongly on the identity of the variant and conditions within the cell. These data therefore define biomarkers (*i.e.* variants and activator levels) that can identify patients most likely to respond to therapy, and provide mechanistic support for combination strategies that suppress upstream activation to enhance therapeutic efficacy (*79*).

The canonical two-conformational-state model fails to explain allosteric inhibition by clinical drugs

Beyond providing a biochemical phenotype for each SHP2 variant, the suite of kinetic and thermodynamic parameters measured here provides a unique opportunity to rigorously test mechanistic models of SHP2 function. As detailed previously, SHP2 allosteric inhibition is typically represented using a two-conformational-state model, in which the enzyme can be either open (active) or closed (autoinhibited), and drugs bind and stabilize the closed state (**Fig. S36A**, **Eqn. 2**, **Supplementary Text 2**) (*48*). However, two experimental observations are at odds with this model. First, this model predicts that SHP2 activity should be completely inhibited at high drug concentrations (**Fig. S54**). Contrary to this prediction, substantial residual SHP2 activity persists at high drug concentrations when using either DiFMUP (**Figs. S37**&**S55**) or the larger peptide substrate EGFRpY992 (**Fig. S56**). Second, this model predicts that more open variants should be less sensitive to drugs if the drug affinity in the closed state (K^I_d) is unchanged (**Fig. S57A**). Contrary to this

prediction, IC_{50} values for TNO155, GDC-1971, and RMC-4630 initially decreased with increasing f_{active}^{0} (**Figs. 5E**&**S57B**). This discrepancy was consistent across all three drugs in HT-MEK assays, plate-reader activity assays, and direct binding measurements (**Fig. S58**). Moreover, this discrepancy remained even when considering only variants >15 Å away from the inhibitor in the crystal structure, eliminating those mutations most likely to impact K_d^I (**Figs. 5E**&**S59**). These findings establish that the current two-conformational-state framework is an inadequate description of SHP2 allosteric inhibition and point to the need for a revised mechanistic model.

Accounting for variant-specific differences in inhibition requires a partially active intermediate state that preferentially binds drugs

The simplest model capable of explaining all observations is a three-conformational-state model that introduces an intermediate state (E_i) in addition to the canonical closed (E_c) and open (E_o) states (**Figs. 6A-C**, **S54B**, **S60**; **Supplementary Texts 3**&**4**). Such a partially-open, intermediate conformation has also been proposed by recent single-molecule and small-angle X-ray scattering analyses (*80*, *81*). The suite of biochemical measurements obtained here provides support for the functional relevance of this state and critical new insights into its properties. To reproduce the observed incomplete inhibition and relationship between IC_{50} and f_{active}^{0}, this intermediate state must: (1) possess residual catalytic activity ($r \cdot k, 0 < r < 1$), and (2) bind drug more strongly than the closed state ($\alpha < 1$) (**Supplementary Text 3**). We note that the model presented here represents a minimal model, and the true model may be yet more complicated (**Supplementary Text 5**).

Biasing SHP2 towards the intermediate state potentiates allosteric inhibition

The three-conformational-state inhibition model uniquely predicts that SHP2 will be: (1) more potently inhibited by allosteric drugs if shifted into the intermediate state, and (2) less potently inhibited if shifted into the open state. Prior single-molecule studies indicate that engaging a single SH2 domain with pTyr peptides stabilizes an intermediate conformation, whereas simultaneous engagement of both SH2 domains favors the fully open conformation. If this intermediate conformation corresponds to the highest drug-affinity state proposed in our three-conformational-state model, engagement of a single SH2 domain should enhance inhibition potency, while engagement of both should reduce it.

To test these predictions (**Figs. 6D**&**S61**), we mutated the key arginine in either nSH2 (R32A) or cSH2 (R138A) to impair binding to pTyr peptides so that the activator ppIRS1 is expected to only engage a single SH2 domain in each mutant. In both cases, increasing ppIRS1 concentration lowered the IC_{50} comparably (**Fig. 6E**), consistent with the prediction that engagement of a single SH2 domain populates an intermediate state with enhanced drug affinity. By contrast, when ppIRS1 could engage both SH2 domains (*i.e.* WT), the equilibrium shifted towards the open state, increasing IC_{50} (**Fig. 6E**). This increase saturated at high ppIRS1 concentrations, consistent with the three-conformational-state model prediction.

The three-conformational-state model further suggests that observed variant-specific increases in drug sensitivity (lower IC_{50} values) result from increased occupancy of the intermediate state (which has higher affinity for inhibitors) (**Fig. 6A**&**F**). The singly phosphorylated IRS1-pY1172 peptide binds nSH2 ~250-fold more strongly than cSH2 (*82*). Therefore, for the predominantly closed WT SHP2, increasing IRS1-pY1172 concentration should initially only engage nSH2, stabilize the intermediate state, and lower IC_{50}; as IRS1-pY1172 concentration increase further, additional engagement of cSH2 should drive SHP2 into the fully open state and increase IC_{50} (**Fig. 6G**). Consistent with these predictions, increasing IRS1-pY1172 initially decreased, and then increased, IC_{50} for WT SHP2 (**Fig. 6H**). By contrast, increasing IRS1-pY1172 concentration for a drug-sensitive mutant (*i.e.* D61G), which is expected to predominantly occupy the intermediate state, should only increase IC_{50} (**Fig. 6G**). As expected, increasing IRS1-pY1172 exclusively increased IC_{50} for D61G (**Fig. 6H**).

Notably, SHP099 (the first-in-class allosteric SHP2 inhibitor) behaved differently from TNO155, RMC-4630, and GDC-1971 (*26*, *32*). For SHP099, variants with a higher open fraction than WT consistently displayed higher

IC_{50} (**Fig. S62A**). Thus, the model predicts that SHP099 does not preferentially bind the intermediate state over the closed state ($\alpha \geq 1$) (**Supplementary Text 3**). Consequently, enrichment in the intermediate state via nSH2 engagement provides no advantage for SHP099 binding, and increasing IRS1-pY1172 should only increase SHP099 IC_{50} (**Fig. S62B**). Measured SHP099 IC_{50} values in the presence of increasing IRS1-pY1172 matched the prediction (**Fig. S62C**), providing additional evidence supporting the three-conformational-state inhibition model.

Cellular evidence supports in vitro predictions of SHP2 inhibition

To test if *in vitro* drug profiling predicts SHP2 inhibition in cells, we treated U2OS cells containing transiently-transfected SHP2 variants with inhibitors and quantified SHP2 activity by assessing pERK levels via Western blot. Overexpression (typically >20-fold over the endogenous level) of SHP2 WT did not alter inhibitor sensitivity, whereas expression of the drug-resistant mutant E76K caused a substantial reduction of sensitivity, indicating that the assay can detect variant-induced shifts in sensitivity (**Figs. S63-S66**). We found that cellular IC_{50} values correlated with *in vitro* IC_{50} values measured under mildly activating conditions (0 or 5 nM ppIRS1) (Pearson *r* (log-log) = 0.84-0.93; **Figs. S67**&**S68**), though the variability of the cellular assay limited our ability to resolve statistically significant differences. As predicted, variants with intermediate f_{active}^{0} were more or similarly sensitive to TNO155, RMC-4630, and GDC-1971 compared to WT, consistent with stronger inhibitor binding to the intermediate state (**Fig. S69**).

Discussion

The number of identified human allelic variants is growing rapidly (*1–4*), with most classified as VUS (*10*). Even for variants annotated as benign or pathogenic, we typically lack quantitative information about their biochemical consequences. As shared biochemical alterations drive similar clinical phenotypes (*63*), linking variants to their biochemical effects can improve genetic test interpretation, clinical management, and the development of targeted therapeutics (*83*, *84*), even for rare or previously unobserved variants. Here, we systematically quantify how 190 SHP2 allelic variants alter ten thermodynamic and kinetic parameters that describe SHP2 stability, catalysis, and regulation, yielding (to our knowledge) the largest quantitative biochemical dataset for allelic variants to date.

Quantifying variant effects across multiple distinct biochemical functions overcomes key limitations of current high-throughput experimental and computational approaches to profile variants. Multiplexed assays of variant effects (MAVEs) can evaluate thousands of variants in cells (*65*, *85*, *86*), and prior SHP2 deep mutational scanning yielded enrichment scores that correlate with our measured k_{cat}/K_M values (**Fig. S70**) (*45*). However, the design of MAVEs often assumes that catalytic output is the primary predictor of pathology; for SHP2, our data show this assumption does not hold. Moreover, cellular assays report convolved outputs and lack the systematic control needed to directly quantify more complex properties such as autoinhibition and activator affinity, yet these proved to be the principal determinants of SHP2 pathogenesis. Beyond experimental assays, machine learning models such as AlphaMissense provide binary pathogenicity predictions but lack mechanistic interpretability (62, 63, 80), thus providing no biochemical understanding of why a given variant is pathogenic or ability to discriminate between pathogenic modalities. High-throughput biochemical data extend these approaches: defining the biochemical parameters most relevant to disease can inform the design of cellular assays that better approximate these parameters at greater scale, while quantitative biochemical datasets provide ground truth data to interpret existing and train next-generation computational models.

Systematic biochemical dissection of SHP2 variants enhances disease classification and establishes a foundation for disease mechanistic hypotheses. NS and NSML exhibit overlapping phenotypes that frequently confound clinical diagnosis (*39*, *40*); our data, consistent with prior reports but now expanded across a broader variant space (*25*, *44*), suggest that dysregulated autoinhibition drives their common phenotypes while deficits in intrinsic catalysis account for NSML-specific symptoms such as lentigines. In cancer cases, nearly all somatic missense variants—including those previously interpreted as consistent with a tumor-suppressive role (*15*, *33*, *45*)—exhibit dysregulated autoinhibition. Among these, variants with reduced catalytic activity are

tolerated almost exclusively in solid cancers, suggesting hematologic malignancies may additionally require catalytic output at or exceeding wild-type levels. As dysregulated autoinhibition is the shared feature across all pathogenic classes regardless of catalytic activity, we speculate that SHP2 variants primarily promote disease through catalysis-independent scaffolding functions (*15*, *66*–*69*).

Beyond classification, thermodynamic and kinetic constants for hundreds of variants enabled systematic evaluation of mechanistic models for SHP2 function. Here, modeling of variant inhibition data revealed that three clinical-stage allosteric inhibitors preferentially bind a partially active intermediate state, whereas the first-generation inhibitor SHP099 does not. Although crystal structures depict all SHP2 inhibitors bound to the fully closed conformation (*32*, *34*, *35*), these static snapshots may not represent the complete conformational ensemble sampled under physiological conditions (*87*–*89*). The observed shift in binding properties from SHP099 to the more potent clinical-stage inhibitors illustrates how drug optimization can inadvertently shift the targeted conformation of allosteric inhibitors in ways not apparent from static structures. Integrating a panel of conformationally diverse variants into lead optimization can benchmark compound activity against mechanistic models and enable early identification of unintended shifts in conformational selectivity before they advance undetected into clinical development.

While allosteric inhibitors have the potential for greater selectivity and the ability to modulate (rather than block) activity (*90*), we highlight four challenges associated with allosteric SHP2 inhibition. First, many cancer variants preferentially adopt the open conformation, rendering them intrinsically less sensitive to compounds that bind the closed or intermediate conformations. Second, as allosteric inhibition relies on structure-wide conformational coupling, resistance-conferring mutations can arise throughout the protein (*90*–*92*). Third, preferential binding to an intermediate state with residual activity means that allosteric SHP2 inhibitors will not completely suppress activity, even at saturating concentrations. Fourth, drug binding to SHP2 competes with binding of endogenous activators, creating a conformational “tug-of-war”. This tug-of-war inverts the therapeutic window such that inhibitors exert weaker effects in cancer cells with elevated activators and stronger effects in normal tissues, potentially diminishing on-target potency while heightening the risk of on-target, off-tumor toxicity. As similar limitations have been observed in other drug targets including tyrosine kinases and Ras (*93*–*95*), these issues may reflect fundamental constraints of allosteric inhibition as a therapeutic strategy.

Restricting activator engagement to a single SH2 domain biases the conformational ensemble toward the intermediate state, enhancing potency for clinical allosteric inhibitors. This suggests a potential therapeutic strategy in which compounds that selectively block one SH2 domain could be combined with allosteric inhibitors to maximize potency in hyperactivated disease contexts. Alternatively, inhibitors that target the PTP domain could bypass this conformational tug-of-war entirely, or even exploit it synergistically if the target conformation is only accessible in the open state. Such inhibitors could act through orthosteric inhibition or by targeting allosteric sites within the PTP domain that operate independently of SH2-mediated autoinhibition (*96*, *97*); our data identified one such candidate site (**Fig. S13**).

We have demonstrated that systematic, multidimensional biochemical profiling of clinical SHP2 variants can expose profound and consequential mechanistic complexity—in both disease etiology and drug response—that remained elusive in one of the most studied protein tyrosine phosphatases in disease biology. SHP2 is unlikely to be exceptional in this regard. Extending this biochemistry-centered approach broadly across other clinically important proteins will be essential to transform the growing catalog of human variants from a challenge into a mechanistic and therapeutic opportunity for advancing precision medicine.

References

1. M. Claussnitzer, J. H. Cho, R. Collins, N. J. Cox, E. T. Dermitzakis, M. E. Hurles, S. Kathiresan, E. E. Kenny, C. M. Lindgren, D. G. MacArthur, K. N. North, S. E. Plon, H. L. Rehm, N. Risch, C. N. Rotimi, J. Shendure, N. Soranzo, M. I. McCarthy, A brief history of human disease genetics. *Nature* **577**, 179–189 (2020).

2. L. Backwell, J. A. Marsh, Diverse Molecular Mechanisms Underlying Pathogenic Protein Mutations: Beyond the Loss-of-Function Paradigm. *Annual Review of Genomics and Human Genetics* **23**, 475–498 (2022).

3. M. Taipale, Disruption of protein function by pathogenic mutations: common and uncommon mechanisms. *Biochem. Cell Biol.* **97**, 46–57 (2019).

4. T. Lappalainen, D. G. MacArthur, From variant to function in human disease genetics. *Science* **373**, 1464–1468 (2021).

5. A. R. Moore, S. C. Rosenberg, F. McCormick, S. Malek, RAS-targeted therapies: is the undruggable drugged? *Nat Rev Drug Discov* **19**, 533–552 (2020).

6. L. Pauling, H. A. Itano, S. J. Singer, I. C. Wells, Sickle Cell Anemia, a Molecular Disease. *Science* **110**, 543–548 (1949).

7. M. Malumbres, M. Barbacid, RAS oncogenes: the first 30 years. *Nat Rev Cancer* **3**, 459–465 (2003).

8. S. Chen, L. C. Francioli, J. K. Goodrich, R. L. Collins, M. Kanai, Q. Wang, J. Alföldi, N. A. Watts, C. Vittal, L. D. Gauthier, T. Poterba, M. W. Wilson, Y. Tarasova, W. Phu, R. Grant, M. T. Yohannes, Z. Koenig, Y. Farjoun, E. Banks, S. Donnelly, S. Gabriel, N. Gupta, S. Ferriera, C. Tolonen, S. Novod, L. Bergelson, D. Roazen, V. Ruano-Rubio, M. Covarrubias, C. Llanwarne, N. Petrillo, G. Wade, T. Jeandet, R. Munshi, K. Tibbetts, A. O'Donnell-Luria, M. Solomonson, C. Seed, A. R. Martin, M. E. Talkowski, H. L. Rehm, M. J. Daly, G. Tiao, B. M. Neale, D. G. MacArthur, K. J. Karczewski, A genomic mutational constraint map using variation in 76,156 human genomes. *Nature* **625**, 92–100 (2024).

9. J. M. Engreitz, H. A. Lawson, H. Singh, L. M. Starita, G. C. Hon, H. Carter, N. Sahni, T. E. Reddy, X. Lin, Y. Li, N. V. Munshi, M. H. Chahrour, A. P. Boyle, B. C. Hitz, A. Mortazavi, M. Craven, K. L. Mohlke, L. Pinello, T. Wang, A. Kundaje, F. Yue, S. Cody, N. P. Farrell, M. I. Love, L. A. Muffley, M. J. Pazin, F. Reese, E. Van Buren, K. K. Dey, M. Kircher, J. Ma, P. Radivojac, B. Balliu, B. A. Williams, D. Huangfu, C. Y. Park, T. Quertermous, J. Das, M. A. Calderwood, D. M. Fowler, M. Vidal, L. Ferreira, S. D. Mooney, V. Pejaver, J. Zhao, S. Gazal, E. Koch, S. K. Reilly, S. Sunyaev, A. E. Carpenter, J. D. Buenrostro, C. S. Leslie, R. E. Savage, S. Giric, C. Luo, K. Plath, A. Barrera, M. Schubach, A. R. Gschwind, J. E. Moore, N. Ahituv, S. S. Yi, I. Hallgrimsdottir, K. J. Gaulton, S. Sakaue, S. Booeshaghi, E. Mattei, S. Nair, L. Pachter, A. T. Wang, J. Shendure, V. Agarwal, A. Blair, T. Chalkiadakis, F. M. Chardon, P. M. Dash, C. Deng, N. Hamazaki, P. Keukeleire, C. Kubo, J.-B. Lalanne, T. Maass, B. Martin, T. A. McDiarmid, M. Nobuhara, N. F. Page, S. Regalado, J. Sims, A. Ushiki, S. M. Best, G. Boyle, N. Camp, S. Casadei, E. Y. Da, M. Dawood, S. C. Dawson, S. Fayer, A. Hamm, R. G. James, G. P. Jarvik, A. E. McEwen, N. Moore, S. Pendyala, N. A. Popp, M. Post, A. F. Rubin, N. T. Smith, J. Stone, M. Tejura, Z. R. Wang, M. K. Wheelock, I. Woo, B. D. Zapp, D. Amgalan, A. Aradhana, S. M. Arana, M. C. Bassik, J. R. Bauman, A. Bhattacharya, X. S. Cai, Z. Chen, S. Conley, S. Deshpande, B. R. Doughty, P. P. Du, J. A. Galante, C. Gifford, W. J. Greenleaf, K. Guo, R. Gupta, S. Isobe, E. Jagoda, N. Jain, H. Jones, H. Y. Kang, S. H. Kim, Y. Kim, S. Klemm, R. Kundu, S. Kundu, M. Lago-Docampo, Y. C. Lee-Yow, R. Levin-Konigsberg, D. Y. Li, D. Lindenhofer, X. R. Ma, G. K. Marinov, G. E. Martyn, C. V. McCreery, E. Metzl-Raz, J. P. Monteiro, M. T. Montgomery, K. S. Mualim, C. Munger, G. Munson, T. C. Nguyen, T. Nguyen, B. T. Palmisano, A. Pampari, M. Rabinovitch, M. Ramste, J. Ray, K. R. Roy, O. M. Rubio, J. M. Schaepe, G. Schnitzler, J. Schreiber, D. Sharma, M. U. Sheth, H. Shi, V. Singh, R. Sinha, L. M. Steinmetz, J. Tan, A. Tan, J. Tycko, R. C. Valbuena, V. V. P. Amiri, M. J. F. M. van Kooten, A. Vaughan-Jackson, A. Venida, C. S. Weldy, M. D. Worssam, F. Xia, D. Yao, T. Zeng, Q. Zhao, R. Zhou, Z. S. Chen, B. A. Cimini, G. Coppin, A. G. Coté, M. Haghighi, T. Hao, D. E. Hill, J. Lacoste, F. Laval, C. Reno, F. P. Roth, S. Singh, K. Spirohn-Fitzgerald, M.

Taipale, T. Teelucksingh, M. Tixhon, A. Yadav, Z. Yang, W. L. Kraus, D. A. Armendariz, A. E. Dederich, A. Gogate, L. El Hayek, S. C. Goetsch, K. Kaur, H. B. Kim, M. K. McCoy, M. Z. Nzima, C. A. Pinzón-Arteaga, B. A. Posner, D. A. Schmitz, S. Sivakumar, A. Sundarrajan, L. Wang, Y. Wang, J. Wu, L. Xu, J. Xu, L. Yu, Y. Zhang, H. Zhao, Q. Zhou, H. Won, J. L. Bell, K. A. Broadaway, K. N. Degner, A. S. Etheridge, B. H. Koller, W. Mah, W. Mu, K. D. Ritola, J. D. Rosen, S. A. Schoenrock, R. A. Sharp, D. Bauer, G. Lettre, R. Sherwood, B. Becerra, L. J. Blaine, E. Che, M. J. Francoeur, E. N. Gibbs, N. Kim, E. M. King, B. P. Kleinstiver, E. Lecluze, Z. Li, Z. M. Patel, Q. V. Phan, J. Ryu, M. L. Starr, T. Wu, C. A. Gersbach, G. E. Crawford, IGVF Consortium, Writing group (ordered by contribution), Steering Committee Co-Chairs (alphabetical by last name), Code of Conduct Committee (alphabetical by last name), Working Group and Focus Group Co-Chairs (alphabetical by last name), Catalog, Characterization, M. Computational Analysis and Prediction, Project Design, Mapping, Networks, Standards and Pipelines, Cardiometabolic, Cellular Programs and Networks, Coding Variants, CRISPR, Defining and Systematizing Function, Enumerating Variants, Evolution, Imaging, Immune, Impact on Diverse Populations, iPSC, MPRA, Noncoding Variants, Neuro, Phenotypic Impact and Function, QTL/Statgen, Single Cell, Mpi. (alphabetical by last name) Characterization Awards (contact PI other members (alphabetical by last name)), UM1HG011966, UM1HG011969, UM1HG011972, UM1HG011989, UM1HG011996, UM1HG012003, UM1HG012010, UM1HG012053, Deciphering the impact of genomic variation on function. *Nature* **633**, 47–57 (2024).

10. M. J. Landrum, J. M. Lee, M. Benson, G. R. Brown, C. Chao, S. Chitipiralla, B. Gu, J. Hart, D. Hoffman, W. Jang, K. Karapetyan, K. Katz, C. Liu, Z. Maddipatla, A. Malheiro, K. McDaniel, M. Ovetsky, G. Riley, G. Zhou, J. B. Holmes, B. L. Kattman, D. R. Maglott, ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Research* **46**, D1062–D1067 (2018).

11. R. M. Freeman, J. Plutzky, B. G. Neel, Identification of a human src homology 2-containing protein-tyrosine-phosphatase: a putative homolog of Drosophila corkscrew. *Proceedings of the National Academy of Sciences* **89**, 11239–11243 (1992).

12. E. Hui, J. Cheung, J. Zhu, X. Su, M. J. Taylor, H. A. Wallweber, D. K. Sasmal, J. Huang, J. M. Kim, I. Mellman, R. D. Vale, T cell costimulatory receptor CD28 is a primary target for PD-1–mediated inhibition. *Science* **355**, 1428–1433 (2017).

13. E. E. Zhang, E. Chapeau, K. Hagihara, G.-S. Feng, Neuronal Shp2 tyrosine phosphatase controls energy balance and metabolism. *Proceedings of the National Academy of Sciences* **101**, 16064–16069 (2004).

14. T. M. Saxton, M. Henkemeyer, S. Gasca, R. Shen, D. J. Rossi, F. Shalaby, G.-S. Feng, T. Pawson, Abnormal mesoderm patterning in mouse embryos mutant for the SH2 tyrosine phosphatase Shp-2. *The EMBO Journal* **16**, 2352–2364 (1997).

15. M. D. Asmamaw, X.-J. Shi, L.-R. Zhang, H.-M. Liu, A comprehensive review of SHP2 and its role in cancer. *Cell Oncol.* **45**, 729–753 (2022).

16. M. Tartaglia, E. L. Mehler, R. Goldberg, G. Zampino, H. G. Brunner, H. Kremer, I. van der Burgt, A. H. Crosby, A. Ion, S. Jeffery, K. Kalidas, M. A. Patton, R. S. Kucherlapati, B. D. Gelb, Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nat Genet* **29**, 465–468 (2001).

17. M. Tartaglia, B. D. Gelb, NOONAN SYNDROME AND RELATED DISORDERS: Genetics and Pathogenesis. *Annual Review of Genomics and Human Genetics* **6**, 45–68 (2005).

18. M. C. Digilio, E. Conti, A. Sarkozy, R. Mingarelli, T. Dottorini, B. Marino, A. Pizzuti, B. Dallapiccola, Grouping of Multiple-Lentigines/LEOPARD and Noonan Syndromes on the PTPN11 Gene. *The American Journal of Human Genetics* **71**, 389–394 (2002).

19. E. Legius, C. Schrander-Stumpel, E. Schollen, C. Pulles-Heintzberger, M. Gewillig, J.-P. Fryns, PTPN11 mutations in LEOPARD syndrome. *Journal of Medical Genetics* **39**, 571–574 (2002).

20. R. J. Chan, G.-S. Feng, PTPN11 is the first identified proto-oncogene that encodes a tyrosine phosphatase. *Blood* **109**, 862–867 (2006).

21. M. Tartaglia, C. M. Niemeyer, A. Fragale, X. Song, J. Buechner, A. Jung, K. Hählen, H. Hasle, J. D. Licht, B. D. Gelb, Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. *Nat Genet* **34**, 148–150 (2003).

22. B. D. Gelb, M. Tartaglia, "Noonan Syndrome with Multiple Lentigines" in *GeneReviews®*, M. P. Adam, J. Feldman, G. M. Mirzaa, R. A. Pagon, S. E. Wallace, A. Amemiya, Eds. (University of Washington, Seattle, Seattle (WA), 1993; http://www.ncbi.nlm.nih.gov/books/NBK1383/).

23. A. Sarkozy, M. C. Digilio, G. Zampino, B. Dallapiccola, M. Tartaglia, B. D. Gelb, "LEOPARD syndrome: Clinical aspects and molecular pathogenesis" in *Noonan Syndrome and Related Disorders - A Matter of Deregulated Ras Signaling* (2009; https://scholars.mssm.edu/en/publications/leopard-syndrome-clinical-aspects-and-molecular-pathogenesis-2/), pp. 55–65.

24. 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).

25. Z.-H. Yu, J. Xu, C. D. Walls, L. Chen, S. Zhang, R. Zhang, L. Wu, L. Wang, S. Liu, Z.-Y. Zhang, Structural and Mechanistic Insights into LEOPARD Syndrome-Associated SHP2 Mutations. *Journal of Biological Chemistry* **288**, 10472–10482 (2013).

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

27. S. Pluskey, T. J. Wandless, C. T. Walsh, S. E. Shoelson, Potent Stimulation of SH-PTP2 Phosphatase Activity by Simultaneous Occupancy of Both SH2 Domains. *Journal of Biological Chemistry* **270**, 2897–2900 (1995).

28. C. H. Lee, D. Kominos, S. Jacques, B. Margolis, J. Schlessinger, S. E. Shoelson, J. Kuriyan, Crystal structures of peptide complexes of the amino-terminal SH2 domain of the Syp tyrosine phosphatase. *Structure* **2**, 423–438 (1994).

29. M. Marasco, A. Berteotti, J. Weyershaeuser, N. Thorausch, J. Sikorska, J. Krausze, H. J. Brandt, J. Kirkpatrick, P. Rios, W. W. Schamel, M. Köhn, T. Carlomagno, Molecular mechanism of SHP2 activation by PD-1 stimulation. *Science Advances* **6**, eaay4458 (2020).

30. D. Barford, B. G. Neel, Revealing mechanisms for SH2 domain mediated regulation of the protein tyrosine phosphatase SHP-2. *Structure* **6**, 249–254 (1998).

31. S. Sugimoto, T. J. Wandless, S. E. Shoelson, B. G. Neel, C. T. Walsh, Activation of the SH2-containing protein tyrosine phosphatase, SH-PTP2, by phosphotyrosine-containing peptides derived from insulin receptor substrate-1. *Journal of Biological Chemistry* **269**, 13614–13622 (1994).

32. Y.-N. P. Chen, M. J. LaMarche, H. M. Chan, P. Fekkes, J. Garcia-Fortanet, M. G. Acker, B. Antonakos, C. H.-T. Chen, Z. Chen, V. G. Cooke, J. R. Dobson, Z. Deng, F. Fei, B. Firestone, M. Fodor, C. Fridrich, H. Gao, D. Grunenfelder, H.-X. Hao, J. Jacob, S. Ho, K. Hsiao, Z. B. Kang, R. Karki, M. Kato, J. Larrow, L. R. La Bonte, F. Lenoir, G. Liu, S. Liu, D. Majumdar, M. J. Meyer, M. Palermo, L. Perez, M. Pu, E. Price, C. Quinn, S. Shakya, M. D. Shultz, J. Slisz, K. Venkatesan, P. Wang, M. Warmuth, S. Williams, G. Yang, J. Yuan, J.-H. Zhang, P. Zhu, T. Ramsey, N. J. Keen, W. R. Sellers, T. Stams, P. D. Fortin, Allosteric inhibition of SHP2 phosphatase inhibits cancers driven by receptor tyrosine kinases. *Nature* **535**, 148–152 (2016).

33. A. Scheiter, L.-C. Lu, L. H. Gao, G.-S. Feng, Complex Roles of PTPN11/SHP2 in Carcinogenesis and Prospect of Targeting SHP2 in Cancer Therapy. *Annual Review of Cancer Biology* **8**, 15–33 (2024).

34. M. J. LaMarche, M. Acker, A. Argintaru, D. Bauer, J. Boisclair, H. Chan, C. H.-T. Chen, Y.-N. Chen, Z. Chen, Z. Deng, M. Dore, D. Dunstan, J. Fan, P. Fekkes, B. Firestone, M. Fodor, J. Garcia-Fortanet, P. D. Fortin, C. Fridrich, J. Giraldes, M. Glick, D. Grunenfelder, H.-X. Hao, M. Hentemann, S. Ho, A. Jouk, Z. B. Kang, R. Karki, M. Kato, N. Keen, R. Koenig, L. R. LaBonte, J. Larrow, G. Liu, S. Liu, D. Majumdar, S. Mathieu, M. J. Meyer, M. Mohseni, R. Ntaganda, M. Palermo, L. Perez, M. Pu, T. Ramsey, J. Reilly, P. Sarver, W. R. Sellers, M. Sendzik, M. D. Shultz, J. Slisz, K. Slocum, T. Smith, S. Spence, T. Stams, C. Straub, V. Jr. Tamez, B.-B. Toure, C. Towler, P. Wang, H. Wang, S. L. Williams, F. Yang, B. Yu, J.-H. Zhang, S. Zhu, Identification of TNO155, an Allosteric SHP2 Inhibitor for the Treatment of Cancer. *J. Med. Chem.* **63**, 13578–13594 (2020).

35. A. M. Taylor, B. R. Williams, F. Giordanetto, E. H. Kelley, A. Lescarbeau, K. Shortsleeves, Y. Tang, W. P. Walters, A. Arrazate, C. Bowman, E. Brophy, E. W. Chan, G. Deshmukh, J. B. Greisman, T. L. Hunsaker, D. R. Kipp, P. Saenz Lopez-Larrocha, D. Maddalo, I. J. Martin, P. Maragakis, M. Merchant, M. Murcko, H. Nisonoff, V. Nguyen, V. Nguyen, O. Orozco, C. Owen, L. Pierce, M. Schmidt, D. E. Shaw, S. Smith, E. Therrien, J. C. Tran, J. Watters, N. J. Waters, J. Wilbur, L. Willmore, Identification of GDC-1971 (RLY-1971), a SHP2 Inhibitor Designed for the Treatment of Solid Tumors. *J. Med. Chem.* **66**, 13384–13399 (2023).

36. Z. Guo, Y. Duan, K. Sun, T. Zheng, J. Liu, S. Xu, J. Xu, Advances in SHP2 tunnel allosteric inhibitors and bifunctional molecules. *European Journal of Medicinal Chemistry* **275**, 116579 (2024).

37. C. Ma, D. Kang, P. Gao, W. Zhang, X. Wu, Z. Xu, H. Han, L. Zhang, Y. Cai, Y. Wang, Y. Wang, W. Long, Discovery of JAB-3312, a Potent SHP2 Allosteric Inhibitor for Cancer Treatment. *J. Med. Chem.* **67**, 13534–13549 (2024).

38. C. P. Kratz, L. Franke, H. Peters, N. Kohlschmidt, B. Kazmierczak, U. Finckh, A. Bier, B. Eichhorn, C. Blank, C. Kraus, J. Kohlhase, S. Pauli, G. Wildhardt, K. Kutsche, B. Auber, A. Christmann, N. Bachmann, D. Mitter, F. W. Cremer, K. Mayer, C. Daumer-Haas, C. Nevinny-Stickel-Hinzpeter, F. Oeffner, G. Schlüter, M. Gencik, B. Überlacker, C. Lissewski, I. Schanze, M. H. Greene, C. Spix, M. Zenker, Cancer spectrum and frequency among children with Noonan, Costello, and cardio-facio-cutaneous syndromes. *Br J Cancer* **112**, 1392–1397 (2015).

39. M. Tartaglia, B. D. Gelb, M. Zenker, Noonan syndrome and clinically related disorders. *Best Practice & Research Clinical Endocrinology & Metabolism* **25**, 161–179 (2011).

40. A. E. Roberts, J. E. Allanson, M. Tartaglia, B. D. Gelb, Noonan syndrome. *The Lancet* **381**, 333–342 (2013).

41. T. Niihori, Y. Aoki, H. Ohashi, K. Kurosawa, T. Kondoh, S. Ishikiriyama, H. Kawame, H. Kamasaki, T. Yamanaka, F. Takada, K. Nishio, M. Sakurai, H. Tamai, T. Nagashima, Y. Suzuki, S. Kure, K. Fujii, M. Imaizumi, Y. Matsubara, Functional analysis of PTPN11/SHP-2 mutants identified in Noonan syndrome and childhood leukemia. *J Hum Genet* **50**, 192–202 (2005).

42. H. Keilhack, F. S. David, M. McGregor, L. C. Cantley, B. G. Neel, Diverse Biochemical Properties of Shp2 Mutants: IMPLICATIONS FOR DISEASE PHENOTYPES *. *Journal of Biological Chemistry* **280**, 30984–30993 (2005).

43. M. I. Kontaridis, K. D. Swanson, F. S. David, D. Barford, B. G. Neel, PTPN11 (Shp2) Mutations in LEOPARD Syndrome Have Dominant Negative, Not Activating, Effects. *Journal of Biological Chemistry* **281**, 6785–6792 (2006).

44. Z.-H. Yu, R.-Y. Zhang, C. D. Walls, L. Chen, S. Zhang, L. Wu, S. Liu, Z.-Y. Zhang, Molecular Basis of Gain-of-Function LEOPARD Syndrome-Associated SHP2 Mutations. *Biochemistry* **53**, 4136–4151 (2014).

45. Z. Jiang, A. E. van Vlimmeren, D. Karandur, A. Semmelman, N. H. Shah, Deep mutational scanning of the multi-domain phosphatase SHP2 reveals mechanisms of regulation and pathogenicity. *Nat Commun* **16**, 5464 (2025).

46. T. Niihori, Y. Aoki, H. Ohashi, K. Kurosawa, T. Kondoh, S. Ishikiriyama, H. Kawame, H. Kamasaki, T. Yamanaka, F. Takada, K. Nishio, M. Sakurai, H. Tamai, T. Nagashima, Y. Suzuki, S. Kure, K. Fujii, M. Imaizumi, Y. Matsubara, Functional analysis of PTPN11/SHP-2 mutants identified in Noonan syndrome and childhood leukemia. *J Hum Genet* **50**, 192–202 (2005).

47. Q. Riller, F. Rieux-Laucat, RASopathies: From germline mutations to somatic and multigenic diseases. *Biomedical Journal* **44**, 422–432 (2021).

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

49. C. J. Markin, D. A. Mokhtari, F. Sunden, M. J. Appel, E. Akiva, S. A. Longwell, C. Sabatti, D. Herschlag, P. M. Fordyce, Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics. *Science* **373**, eabf8761 (2021).

50. C. J. Markin, D. A. Mokhtari, S. Du, T. Doukov, F. Sunden, J. A. Cook, P. M. Fordyce, D. Herschlag, Decoupling of catalysis and transition state analog binding from mutations throughout a phosphatase revealed by high-throughput enzymology. *Proceedings of the National Academy of Sciences* **120**, e2219074120 (2023).

51. J. R. LaRochelle, M. Fodor, X. Xu, I. Durzynska, L. Fan, T. Stams, H. M. Chan, M. J. LaMarche, R. Chopra, P. Wang, P. D. Fortin, M. G. Acker, S. C. Blacklow, Structural and Functional Consequences of Three Cancer-Associated Mutations of the Oncogenic Phosphatase SHP2. *Biochemistry* **55**, 2269–2277 (2016).

52. A. Zeke, T. Takács, P. Sok, K. Németh, K. Kirsch, P. Egri, Á. L. Póti, I. Bento, G. E. Tusnády, A. Reményi, Structural insights into the pSer/pThr dependent regulation of the SHP2 tyrosine phosphatase in insulin and CD28 signaling. *Nat Commun* **13**, 5439 (2022).

53. D. A. Keedy, Z. B. Hill, J. T. Biel, E. Kang, T. J. Rettenmaier, J. Brandão-Neto, N. M. Pearce, F. von Delft, J. A. Wells, J. S. Fraser, An expanded allosteric network in PTP1B by multitemperature crystallography, fragment screening, and covalent tethering. *eLife* **7**, e36307 (2018).

54. D. S. Cui, V. Beaumont, P. S. Ginther, J. M. Lipchock, J. P. Loria, Leveraging Reciprocity to Identify and Characterize Unknown Allosteric Sites in Protein Tyrosine Phosphatases. *Journal of Molecular Biology* **429**, 2360–2372 (2017).

55. R. L. Redler, J. Das, J. R. Diaz, N. V. Dokholyan, Protein Destabilization as a Common Factor in Diverse Inherited Disorders. *J Mol Evol* **82**, 11–16 (2016).

56. P. Yue, Z. Li, J. Moult, Loss of Protein Structure Stability as a Major Causative Factor in Monogenic Disease. *Journal of Molecular Biology* **353**, 459–473 (2005).

57. A. Stein, D. M. Fowler, R. Hartmann-Petersen, K. Lindorff-Larsen, Biophysical and Mechanistic models for disease-causing protein variants. *Trends Biochem Sci* **44**, 575–588 (2019).

58. L. Pagano, V. Pennacchietti, F. Malagrinò, M. Di Felice, J. Toso, E. Puglisi, S. Gianni, A. Toto, Folding and Binding Kinetics of the Tandem of SH2 Domains from SHP2. *International Journal of Molecular Sciences* **25**, 6566 (2024).

59. Y. M. Agazie, M. J. Hayman, Molecular Mechanism for a Role of SHP2 in Epidermal Growth Factor Receptor Signaling. *Molecular and Cellular Biology* **23**, 7875–7886 (2003).

60. M. Brune, J. L. Hunter, J. E. T. Corrie, M. R. Webb, Direct, Real-Time Measurement of Rapid Inorganic Phosphate Release Using a Novel Fluorescent Probe and Its Application to Actomyosin Subfragment 1 ATPase. *Biochemistry* **33**, 8262–8271 (1994).

61. R.-Y. Zhang, Z.-H. Yu, L. Chen, C. D. Walls, S. Zhang, L. Wu, Z.-Y. Zhang, Mechanistic insights explain the transforming potential of the T507K substitution in the protein-tyrosine phosphatase SHP2. *Journal of Biological Chemistry* **295**, 6187–6201 (2020).

62. A. E. van Vlimmeren, R. Voleti, C. A. Chartier, Z. Jiang, D. Karandur, P. A. Humphries, W.-L. Lo, N. H. Shah, The pathogenic T42A mutation in SHP2 rewires the interaction specificity of its N-terminal regulatory domain. *Proceedings of the National Academy of Sciences* **121**, e2407159121 (2024).

63. K. K. Osbak, K. Colclough, C. Saint-Martin, N. L. Beer, C. Bellanné-Chantelot, S. Ellard, A. L. Gloyn, Update on mutations in glucokinase (GCK), which cause maturity-onset diabetes of the young, permanent neonatal diabetes, and hyperinsulinemic hypoglycemia. *Human Mutation* **30**, 1512–1526 (2009).

64. S. Richards, N. Aziz, S. Bale, D. Bick, S. Das, J. Gastier-Foster, W. W. Grody, M. Hegde, E. Lyon, E. Spector, K. Voelkerding, H. L. Rehm, Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* **17**, 405–423 (2015).

65. A. E. McEwen, M. Tejura, S. Fayer, L. M. Starita, D. M. Fowler, Multiplexed assays of variant effect for clinical variant interpretation. *Nat Rev Genet*, 1–18 (2025).

66. A. M. Bennett, T. L. Tang, S. Sugimoto, C. T. Walsh, B. G. Neel, Protein-tyrosine-phosphatase SHPTP2 couples platelet-derived growth factor receptor beta to Ras. *Proceedings of the National Academy of Sciences* **91**, 7335–7339 (1994).

67. R. A. Stewart, T. Sanda, H. R. Widlund, S. Zhu, K. D. Swanson, A. D. Hurley, M. Bentires-Alj, D. E. Fisher, M. I. Kontaridis, A. T. Look, B. G. Neel, Phosphatase-Dependent and -Independent Functions of Shp2 in Neural Crest Cells Underlie LEOPARD Syndrome Pathogenesis. *Developmental Cell* **18**, 750–762 (2010).

68. V. Vemulapalli, L. A. Chylek, A. Erickson, A. Pfeiffer, K.-H. Gabriel, J. LaRochelle, K. Subramanian, R. Cao, K. Stegmaier, M. Mohseni, M. J. LaMarche, M. G. Acker, P. K. Sorger, S. P. Gygi, S. C. Blacklow, Time-resolved phosphoproteomics reveals scaffolding and catalysis-responsive patterns of SHP2-dependent signaling. *eLife* **10**, e64251 (2021).

69. G. Zhu, J. Xie, W. Kong, J. Xie, Y. Li, L. Du, Q. Zheng, L. Sun, M. Guan, H. Li, T. Zhu, H. He, Z. Liu, X. Xia, C. Kan, Y. Tao, H. C. Shen, D. Li, S. Wang, Y. Yu, Z.-H. Yu, Z.-Y. Zhang, C. Liu, J. Zhu, Phase Separation of Disease-Associated SHP2 Mutants Underlies MAPK Hyperactivation. *Cell* **183**, 490-502.e18 (2020).

70. J. Cheng, G. Novati, J. Pan, C. Bycroft, A. Žemgulytė, T. Applebaum, A. Pritzel, L. H. Wong, M. Zielinski, T. Sargeant, R. G. Schneider, A. W. Senior, J. Jumper, D. Hassabis, P. Kohli, Ž. Avsec, Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* **381**, eadg7492 (2023).

71. E. F. McDonald, K. E. Oliver, J. P. Schlebach, J. Meiler, L. Plate, Benchmarking AlphaMissense pathogenicity predictions against cystic fibrosis variants. *PLOS ONE* **19**, e0297560 (2024).

72. H. Murali, P. Wang, E. C. Liao, K. Wang, Genetic variant classification by predicted protein structure: A case study on IRF6. *Computational and Structural Biotechnology Journal* **23**, 892–904 (2024).

73. Y.-H. Wang, E. Masson, Z. Liao, C. Férec, W.-B. Zou, J.-M. Chen, Strengths and limitations of AlphaMissense in CPA1 missense variant classification. doi: 10.1136/gutjnl-2024-332120 (2024).

74. Z. Sondka, N. B. Dhir, D. Carvalho-Silva, S. Jupe, Madhumita, K. McLaren, M. Starkey, S. Ward, J. Wilding, M. Ahmed, J. Argasinska, D. Beare, M. S. Chawla, S. Duke, I. Fasanella, A. G. Neogi, S. Haller, B. Hetenyi, L. Hodges, A. Holmes, R. Lyne, T. Maurel, S. Nair, H. Pedro, A. Sangrador-Vegas, H. Schuilenburg, Z. Sheard, S. Y. Yong, J. Teague, COSMIC: a curated database of somatic variants and clinical data for cancer. *Nucleic Acids Res* **52**, D1210–D1217 (2024).

75. M. Tartaglia, S. Martinelli, L. Stella, G. Bocchinfuso, E. Flex, V. Cordeddu, G. Zampino, I. van der Burgt, A. Palleschi, T. C. Petrucci, M. Sorcini, C. Schoch, R. Foà, P. D. Emanuel, B. D. Gelb, Diversity and Functional Consequences of Germline and Somatic PTPN11 Mutations in Human Disease. *The American Journal of Human Genetics* **78**, 279–290 (2006).

76. D. Xu, X. Liu, W.-M. Yu, H. J. Meyerson, C. Guo, S. L. Gerson, C.-K. Qu, Non–lineage/stage-restricted effects of a gain-of-function mutation in tyrosine phosphatase Ptpn11 (Shp2) on malignant transformation of hematopoietic cells. *J Exp Med* **208**, 1977–1988 (2011).

77. M. Bentires-Alj, S. G. Gil, R. Chan, Z. C. Wang, Y. Wang, N. Imanaka, L. N. Harris, A. Richardson, B. G. Neel, H. Gu, A role for the scaffolding adapter GAB2 in breast cancer. *Nat Med* **12**, 114–121 (2006).

78. M. Hatakeyama, Oncogenic mechanisms of the Helicobacter pylori CagA protein. *Nat Rev Cancer* **4**, 688–694 (2004).

79. N. M. Sodir, G. Pathria, J. I. Adamkewicz, E. H. Kelley, J. Sudhamsu, M. Merchant, R. Chiarle, D. Maddalo, SHP2: A Pleiotropic Target at the Interface of Cancer and Its Microenvironment. *Cancer Discov* **13**, 2339–2355 (2023).

80. Y. Tao, J. Xie, Q. Zhong, Y. Wang, S. Zhang, F. Luo, F. Wen, J. Xie, J. Zhao, X. Sun, H. Long, J. Ma, Q. Zhang, J. Long, X. Fang, Y. Lu, D. Li, M. Li, J. Zhu, B. Sun, G. Li, J. Diao, C. Liu, A novel partially open state of SHP2 points to a “multiple gear” regulation mechanism. *Journal of Biological Chemistry* **296** (2021).

81. M. Anselmi, J. S. Hub, Atomistic ensemble of active SHP2 phosphatase. *Commun Biol* **6**, 1289 (2023).

82. R. J. Lechleider, S. Sugimoto, A. M. Bennett, A. S. Kashishian, J. A. Cooper, S. E. Shoelson, C. T. Walsh, B. G. Neel, Activation of the SH2-containing phosphotyrosine phosphatase SH-PTP2 by its binding site, phosphotyrosine 1009, on the human platelet-derived growth factor receptor. *Journal of Biological Chemistry* **268**, 21478–21481 (1993).

83. M. Lopes-Pacheco, CFTR Modulators: The Changing Face of Cystic Fibrosis in the Era of Precision Medicine. *Front. Pharmacol.* **10** (2020).

84. K. Clift, S. Macklin, C. Halverson, J. B. McCormick, A. M. Abu Dabrh, S. Hines, Patients’ views on variants of uncertain significance across indications. *J Community Genet* **11**, 139–145 (2020).

85. A. Beltran, X. Jiang, Y. Shen, B. Lehner, Site-saturation mutagenesis of 500 human protein domains. *Nature* **637**, 885–894 (2025).

86. G. M. Findlay, R. M. Daza, B. Martin, M. D. Zhang, A. P. Leith, M. Gasperini, J. D. Janizek, X. Huang, L. M. Starita, J. Shendure, Accurate classification of BRCA1 variants with saturation genome editing. *Nature* **562**, 217–222 (2018).

87. M. D. Tyka, D. A. Keedy, I. André, F. DiMaio, Y. Song, D. C. Richardson, J. S. Richardson, D. Baker, Alternate States of Proteins Revealed by Detailed Energy Landscape Mapping. *Journal of Molecular Biology* **405**, 607–618 (2011).

88. J. S. Fraser, H. van den Bedem, A. J. Samelson, P. T. Lang, J. M. Holton, N. Echols, T. Alber, Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proceedings of the National Academy of Sciences* **108**, 16247–16252 (2011).

89. I.-J. L. Byeon, X. Meng, J. Jung, G. Zhao, R. Yang, J. Ahn, J. Shi, J. Concel, C. Aiken, P. Zhang, A. M. Gronenborn, Structural Convergence between Cryo-EM and NMR Reveals Intersubunit Interactions Critical for HIV-1 Capsid Function. *Cell* **139**, 780–790 (2009).

90. R. Nussinov, C.-J. Tsai, Allostery in Disease and in Drug Discovery. *Cell* **153**, 293–305 (2013).

91. I. R. Outhwaite, I. Kwan, A. Leyte-Vidal, N. P. Shah, I. Bahar, M. A. Seeliger, Resistance to Allosteric Inhibitors. *Journal of Molecular Biology* **437**, 169133 (2025).

92. H. N. Motlagh, J. O. Wrabl, J. Li, V. J. Hilser, The ensemble nature of allostery. *Nature* **508**, 331–339 (2014).

93. M. Azam, R. R. Latek, G. Q. Daley, Mechanisms of Autoinhibition and STI-571/Imatinib Resistance Revealed by Mutagenesis of BCR-ABL. *Cell* **112**, 831–843 (2003).

94. P. Lito, M. Solomon, L.-S. Li, R. Hansen, N. Rosen, Allele-specific inhibitors inactivate mutant KRAS G12C by a trapping mechanism. *Science* **351**, 604–608 (2016).

95. A. S. Little, K. Balmanno, M. J. Sale, S. Newman, J. R. Dry, M. Hampson, P. A. W. Edwards, P. D. Smith, S. J. Cook, Amplification of the Driving Oncogene, KRAS or BRAF, Underpins Acquired Resistance to MEK1/2 Inhibitors in Colorectal Cancer Cells. *Science Signaling* **4**, ra17–ra17 (2011).

96. C. Wiesmann, K. J. Barr, J. Kung, J. Zhu, D. A. Erlanson, W. Shen, B. J. Fahr, M. Zhong, L. Taylor, M. Randal, R. S. McDowell, S. K. Hansen, Allosteric inhibition of protein tyrosine phosphatase 1B. *Nat Struct Mol Biol* **11**, 730–737 (2004).

97. B. Marsh-Armstrong, J. M. Fajnzylber, S. Korntner, B. A. Plaman, A. C. Bishop, The Allosteric Site on SHP2's Protein Tyrosine Phosphatase Domain is Targetable with Druglike Small Molecules. *ACS Omega* **3**, 15763–15770 (2018).

Acknowledgements

We thank members of the Herschlag and Fordyce laboratories for discussions and review of the manuscript, and Dr. David A Stevenson and Dr. Bruce D Gelb for helpful discussions. **Funding**: This work was supported by NIH grant R01 (GM064798) awarded to P.M.F. and D.H., an Ono Pharma Foundation Breakthrough Innovation Prize, and the Gordon and Betty Moore Foundation (Grant 8415). This work was also supported by NIH grant 1R01CA272484 awarded to S.C.B.. P.M.F. is a Chan Zuckerberg Biohub San Francisco Investigator. A.A.L. acknowledges the support from the Stanford School of Medicine Dean's Postdoctoral Fellowship. D.A.M. acknowledges support from the Stanford Medical Scientist Training Program and a Stanford Interdisciplinary Graduate Fellowship (Anonymous Donor) affiliated with Stanford ChEM-H. **Author contributions**: A.A.L., D.A.M., D.H. and P.M.F. conceived the project. A.A.L. and D.A.M. designed experiments and collected and analyzed data. E.D.E. collected and analyzed data. A.A.L. performed modeling and simulations. S.C.B., D.H., and P.M.F. supervised the project. A.A.L., D.H., and P.M.F. wrote the manuscript, with input from all authors. **Competing interests**: D.A.M. is a co-founder and employee of Velocity Bio, Inc. P.M.F. is a co-founder of Velocity Bio, Inc. S.C.B. has equity in and is chair of the SAB for Odyssey Therapeutics, is on the SAB for ERASCA Inc., is a consultant for Antares Therapeutics, and is on the boards of the non-profit Revson Foundation, Institute for Protein Innovation, and Harvard Medical School chapter of the Kalaniyot Foundation. All other authors declare no competing interests. **Data and materials availability**: Summary tables of all kinetic and thermodynamic parameters measured for each variant are included in the supplementary materials and available on the Fordyce Lab website (www.fordycelab.com/data). All data acquired in this study are available in a registered Open Science Foundation Repository (https://doi.org/10.17605/OSF.IO/YGHQV).

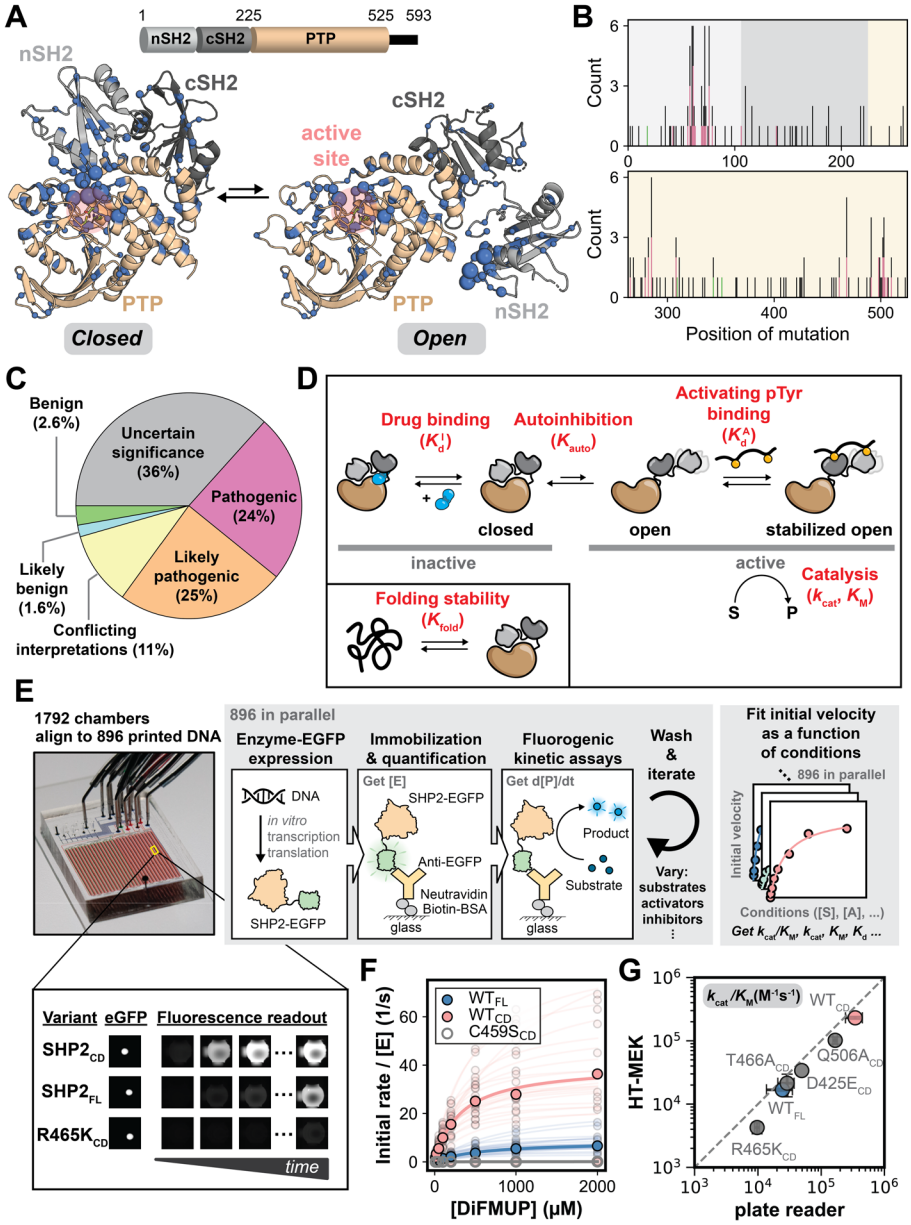

Figure 1. Overview of SHP2 and assaying variants in high throughput. (A) Crystal structure of SHP2 (PDB ID: 4DGP, left & 6CRF, right) highlighting positions mutated in the library in blue spheres; sphere size corresponds to the number of mutations at that position. **(B)** Distribution of unique single amino acid variants by position. Variants are color-coded by clinical classification: pathogenic (magenta), benign (green), and others (gray). **(C)** ClinVar classification of SHP2 variants in this study. **(D)** Schematic representation of key biochemical processes constituting SHP2 function. **(E)** Overview of HT-MEK microfluidic device and SHP2 on-chip expression, purification, and assay procedure. **(F)** Example Michaelis-Menten fits to initial rates. Opaque lines and points indicate the median (from 5–23 replicates), while semi-transparent lines and points represent individual replicates. **(G)** Comparison of median k_{cat}/K_M for DiFMUP hydrolysis measured with HT-MEK and plate reader.

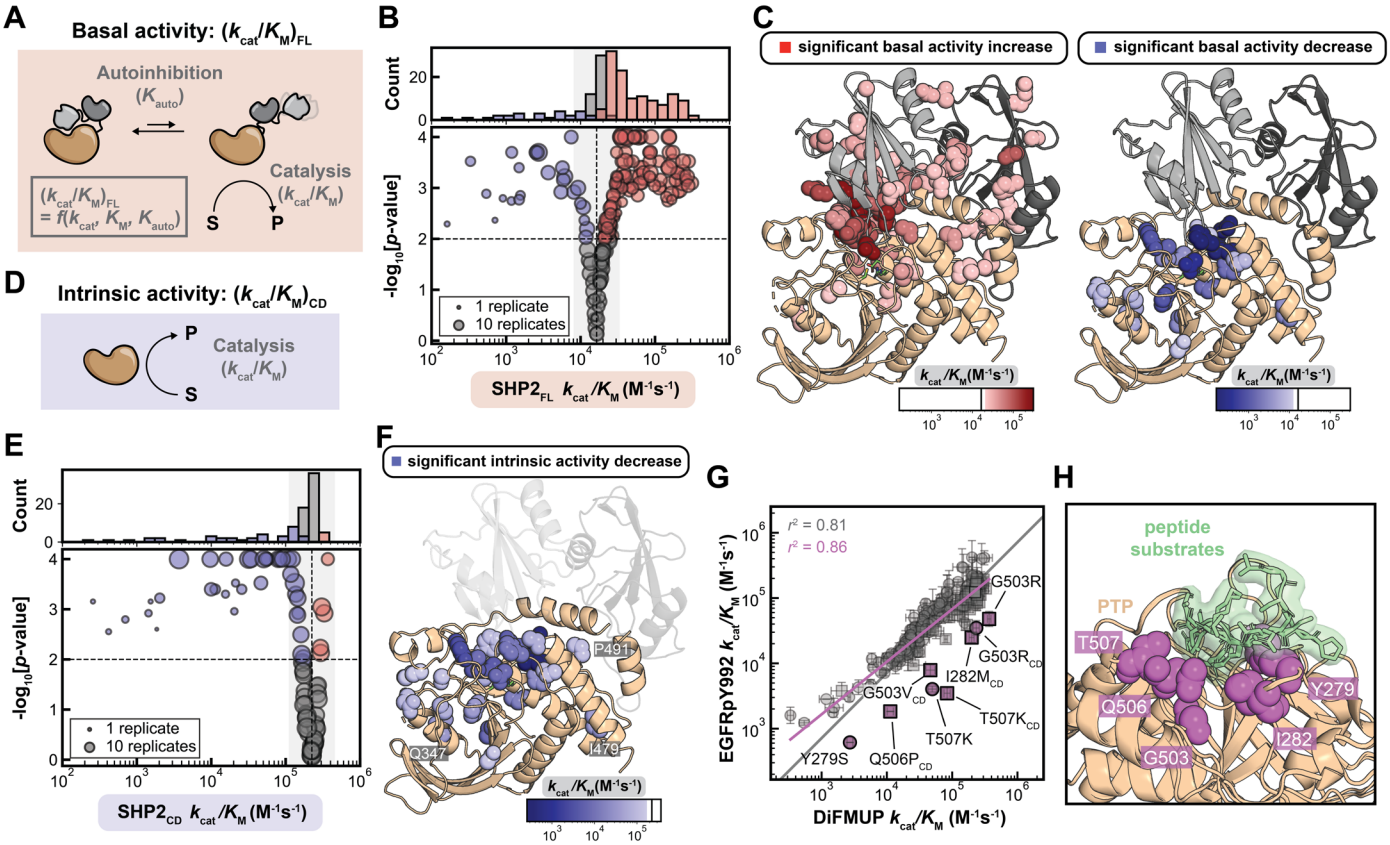

Figure 2. Widespread SHP2 variant activity reveals distinct structural hotspots for activity modulation. **(A)** Basal activity is defined as the k_{cat}/K_M of $SHP2_{FL}$, which is determined by the autoinhibition equilibrium and the catalytic domain activity. **(B)** Median k_{cat}/K_M for DiFMUP hydrolysis across 190 $SHP2_{FL}$ variants. Red: 120 significantly activating variants ($p < 0.01$). Blue: 26 significantly deactivating variants ($p < 0.01$). Gray: not significantly different from WT ($p \geq 0.01$). Gray shaded area is within 2-fold from WT. **(C)** Structural mapping of mutations with significantly increased (left) or decreased (right) basal SHP2 activity. Color intensity of spheres reflects average variant activity at each position. The WT activity is marked by a black reference line on scale bar. **(D)** Intrinsic activity is defined as the k_{cat}/K_M of $SHP2_{CD}$. **(E)** Median k_{cat}/K_M for DiFMUP hydrolysis across 95 $SHP2_{CD}$ variants. Red: 5 significantly activating variants ($p < 0.01$). Blue: 36 significantly deactivating variants ($p < 0.01$). Gray: not significantly different from WT ($p \geq 0.01$). **(F)** Structural mapping of mutations with significantly decreased intrinsic activity. Blue spheres indicate side chain positions and color intensity reflects average variant activity at each position. **(G)** Comparing mutational effects on hydrolysis of DiFMUP and EGFRpY992 peptide substrate. Grey line: log-log fit constrained to slope = 1 (r^2 = 0.81); purple line: unconstrained log-log linear regression (r^2 = 0.86, slope = 0.81). Variants deviating from the regression (residual |*z*-score| > 3) are colored in purple. **(H)** Structural location of the mutants (panel **G**, purple) with differentially altered activity against DiFMUP versus EGFRpY992 (PDB ID: 7PPL, 7PPM, 7PPN; overlayed).

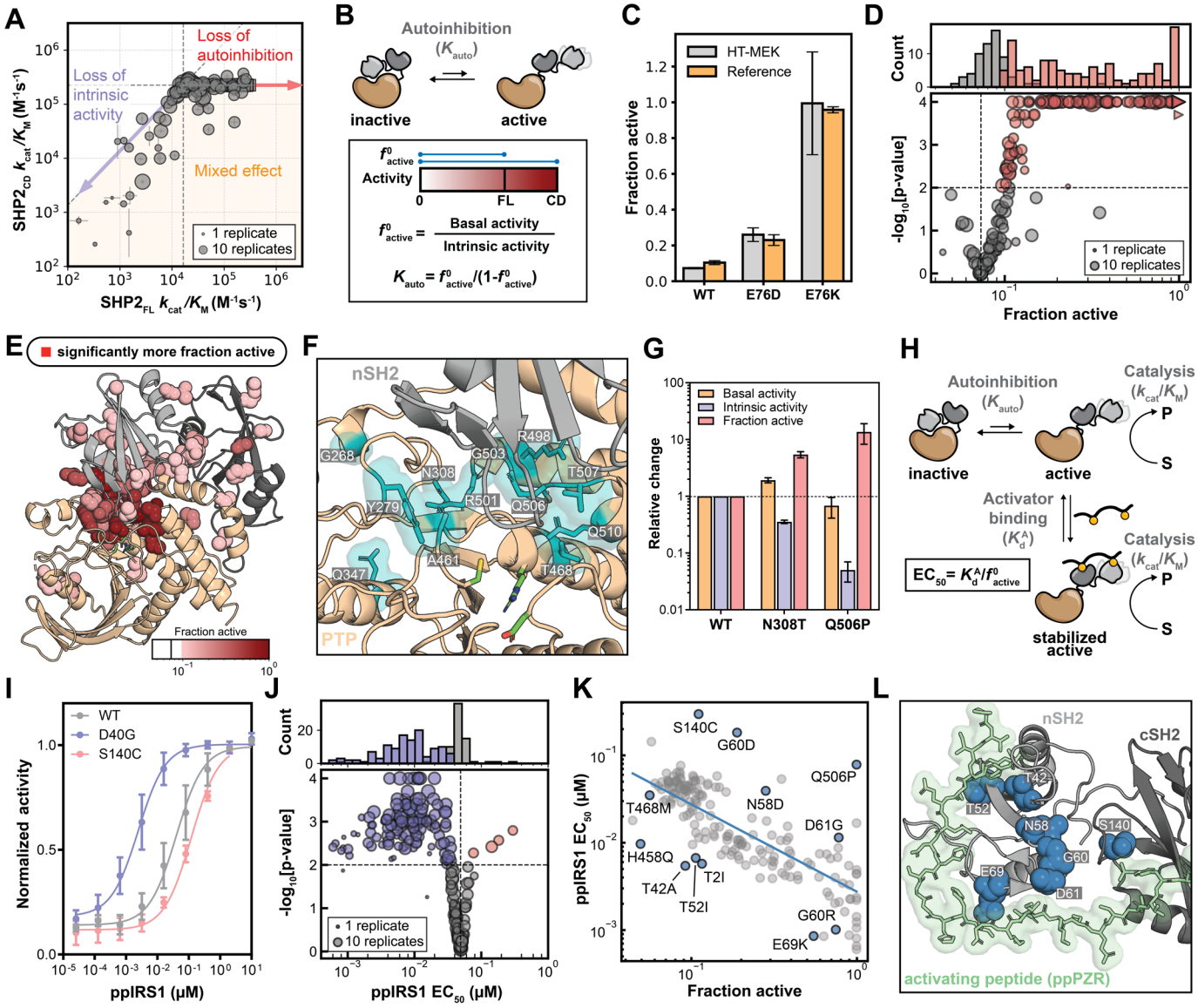

Figure 3. Autoinhibition equilibrium regulates SHP2 variant basal activity and ligand-dependent activation. (A) k_{cat}/K_M of DiFMUP hydrolysis in $SHP2_{CD}$ versus $SHP2_{FL}$ construct. WT values are marked by the intersecting dash lines. Error bar: bootstrap standard error. Square markers indicate mutations in the SH2 domains. **(B)** Calculation of fraction active. **(C)** HT-MEK measured fraction active compared to literature values (*48*). **(D)** Median fraction active of 185 SHP2 variants. Red: 121 significantly activating variants ($p < 0.01$). Gray: not significantly different from WT ($p \geq 0.01$). **(E)** Positions with mutations causing significantly more fraction active shown as red spheres, with intensity reflecting the average effect at each position. **(F)** nSH2–PTP autoinhibitory interface. Positions with mutations that significantly reduce intrinsic activity and increase fraction active are highlighted in cyan. **(G)** Variants with significantly reduced intrinsic activity and increased fraction active, but <2-fold change in basal activity relative to WT. Data are presented as mean ± SEM. **(H)** Regulation of SHP2 by activators and parameters influencing EC_{50}. **(I)** Example HT-MEK-measured dose-response curves for activating peptide ppIRS1. **(J)** Median EC_{50} of 188 SHP2 variants. Red: significantly activating variants ($p < 0.01$). Blue: significantly deactivating variants ($p < 0.01$). **(K)** ppIRS1 EC_{50} versus fraction active. Blue line indicates fitted model (**Eqn. 1**) assuming a constant affinity to activator (K_d^A). Highlighted variants (blue) significantly deviate by >3-fold from the global fit ($p < 0.01$) and have at least 3 replicates. **(L)** Mutations within the SH2 domains with significant deviation from the global EC_{50} versus fraction active fit from panel **K** plotted onto the structure (PDB ID: 9MQ5).

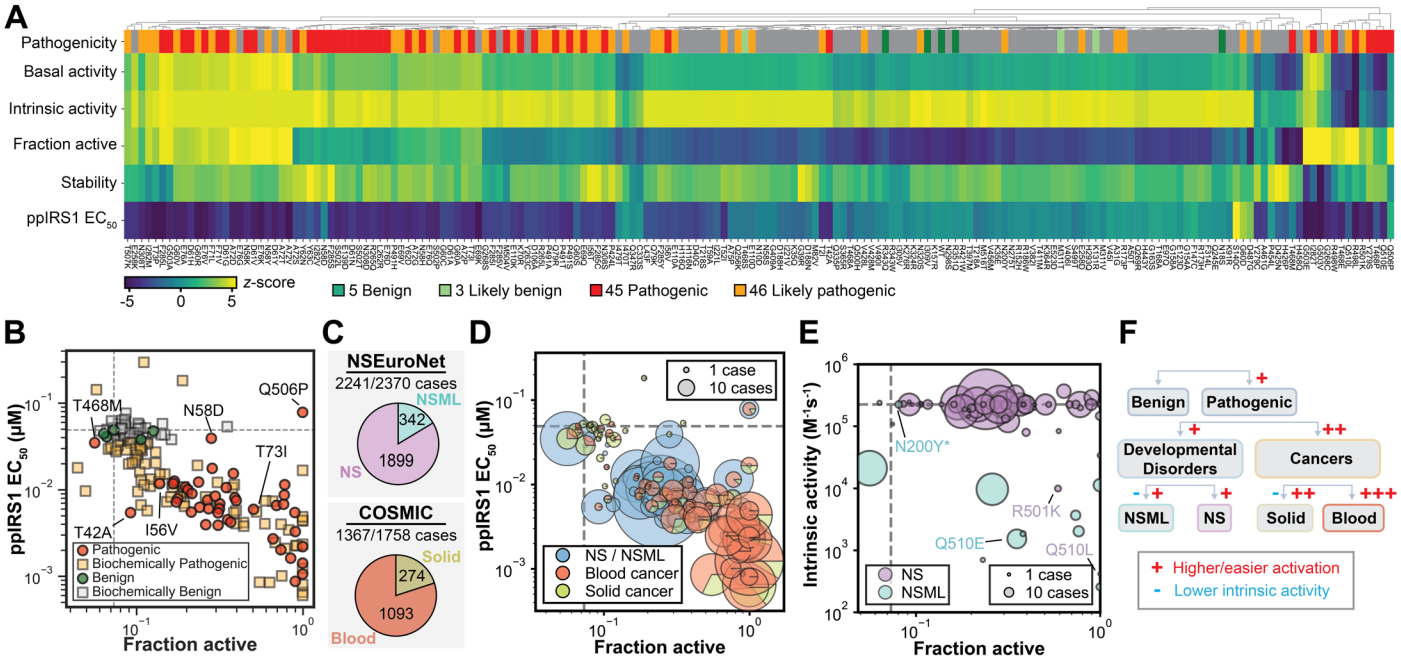

Figure 4. Biochemical signatures segregate SHP2 variants by disease class. (A) Hierarchical clustering of 179 SHP2 variants based on five measured biochemical parameters. **(B)** ppIRS1 EC_{50} versus fraction active for 179 SHP2 variants. Variants with classification other than pathogenic or benign in ClinVar (squares) were classified based on *k*-nearest neighbor model. **(C)** NSEuroNet: 2241 (1899 NS; 342 NSML) of 2370 cases harbor variants in our library. COSMIC: 1367 (1093 blood; 274 solid) of 1758 cancer cases harbor variants in our library. **(D)** Comparing measured ppIRS1 EC_{50} versus fraction active values for variants reported in developmental disorder versus cancer cases; marker area is proportional to the number of case reports for each variant. **(E)** Intrinsic activity versus fraction active for variants reported as NS or NSML in the NSEuroNet database. Marker area is proportional to the number of case reports for each variant. *Confirmed misannotated variant. **(F)** Hierarchical classification of SHP2-related diseases and the biochemical distinction between subclasses.

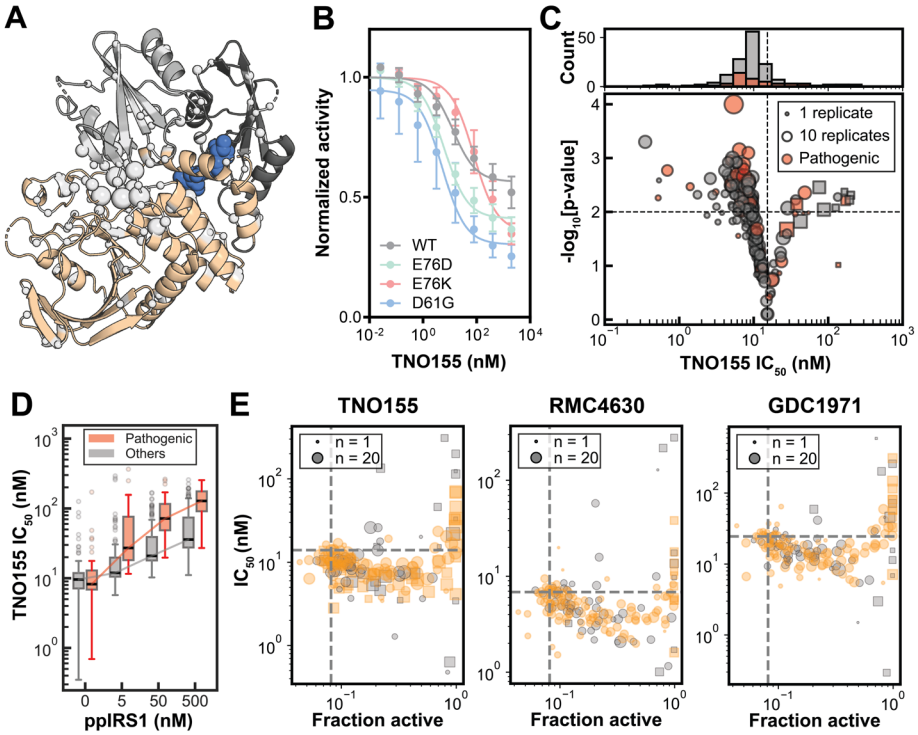

Figure 5. Profiling SHP2 variants reveals variable allosteric inhibitor sensitivity. (A) Crystal structure of TNO155-bound SHP2 (PDB ID:7JVM); the drug (blue) is located between the PTP and SH2 domains. White spheres indicate positions mutated in the library; sphere size reflects mutation count. **(B)** HT-MEK measured inhibition dose-response curves for select variants. Data shown in mean ± SEM. **(C)** Volcano plot showing significance of difference from WT (*p*-values) versus median TNO155 IC_{50} for 190 SHP2 variants. Red: Pathogenic variants in ClinVar. Variants with IC_{50} measurements outside the range of inhibitor concentrations used are plotted in squares. **(D)** TNO155 IC_{50} across different ppIRS1 concentrations. Pathogenic variants are marked in red. **(E)** TNO155, RMC-4630 and GDC-1971 IC_{50} versus f^0_{active}. Gold: variants with mutation > 15 Å away from the bound drug. Dash lines show WT values.

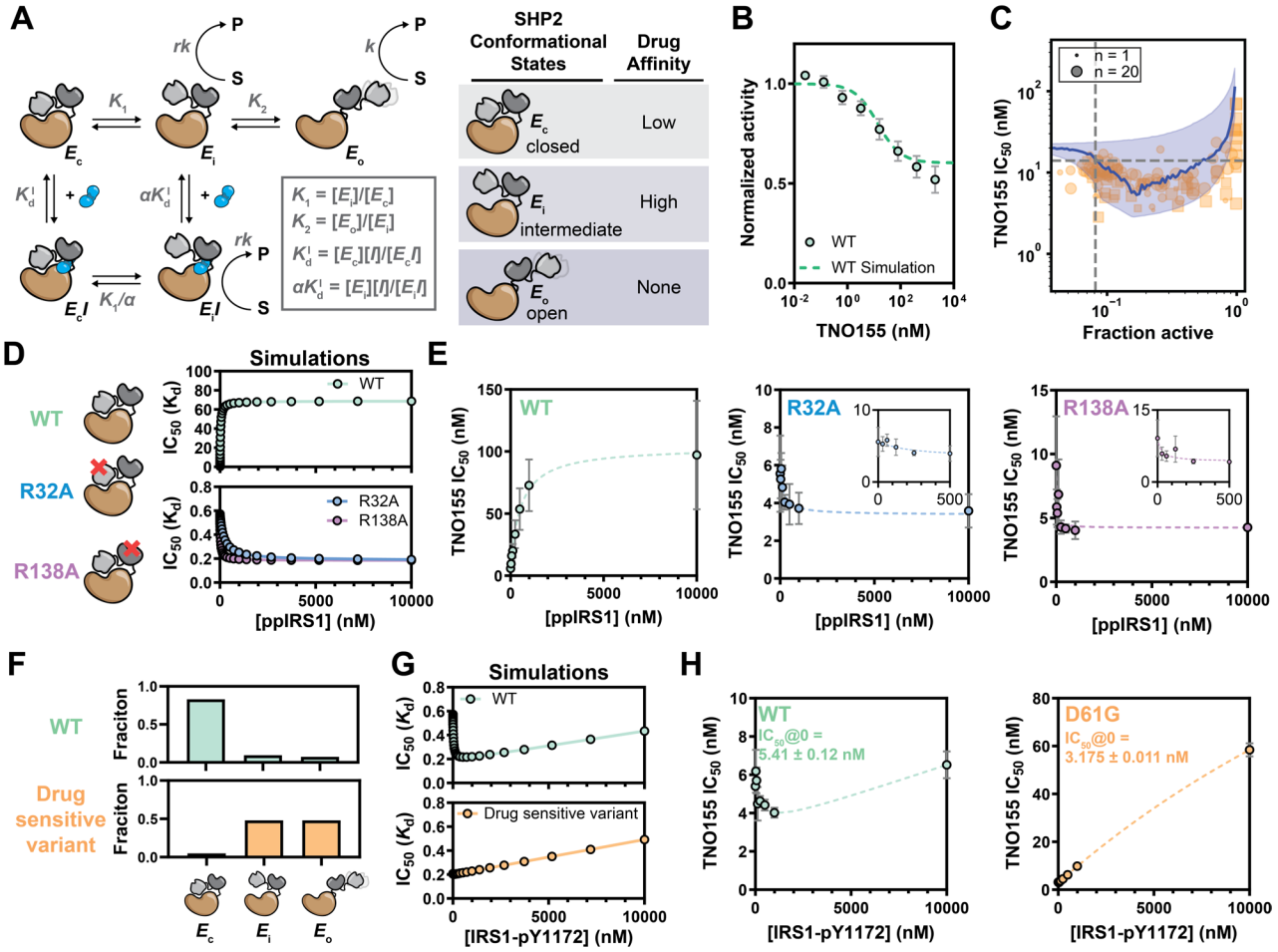

Figure 6. A partially active intermediate conformation is the primary target for clinical allosteric inhibitors. (A) Schematic of the three-conformational-state model of SHP2 conformational states and inhibition. **(B)** Simulated dose-response curve using the three-conformational-state model overlaid with experimental data for WT (Mean ± SEM). **(C)** Simulated IC_{50} versus f^{0}_{active} data from the three-conformational-state model (100,000 simulations). Blue line and shading: binned median and 95th percentile. Gold: TNO155 data for variants with mutations >15 Å from drug. **(D)** Simulated inhibitor IC_{50} dependence on ppIRS1 for WT, R32A, and R138A. **(E)** Experimental data of TNO155 IC_{50} dependence on ppIRS1 concentration for WT (left), R32A (center), and R138A (right). Data shown in mean ± SEM and fit to a general binding coupling model. **(F)** Simulated conformational state distribution of WT and a hypothetical drug sensitive variant. **(G)** Simulated inhibitor IC_{50} dependence on IRS1-pY1172 for WT (top) and a drug sensitive variant (bottom). **(H)** Experimental data of TNO155 IC_{50} dependence on IRS1-pY1172 concentration for WT (left) and D61G (right). Data shown in mean ± SEM and fit with a dual coupling model. See **Materials and Methods** for details of all simulations.