QuantumSi

Control More of Your Protein Research

Introducing Platinum™ *Next-Generation* Protein Sequencer

The power of protein sequencing is in your hands! Now, you can sequence proteins right in your lab with Platinum™, the NEW benchtop solution from Quantum-Si.

Our platform offers single-molecule resolution, a simple workflow and automated, cloud-based data analysis software that delivers proteoform information without added turnaround time or extensive infrastructure.

Platinum provides deeper insights into protein biology to complement your existing "omics" approaches by enabling you to:

- Conduct proteomics experiments in your lab *at your bench*
- Interrogate protein variants and PTMs, and correlate with biological function
- Achieve deeper proteomics insights faster
- Perform analytics with no bioinformatics expertise required

Introducing Platinum™

The Protein Sequencing Company™

Received: 20 April 2023 | Revised: 26 June 2023 | Accepted: 12 July 2023

DOI: 10.1002/pro.4729

RESEARCH ARTICLE

THE PROTEIN SOCIETY WILEY

An intramolecular energetic network regulates ligand recognition in a SH2 domain

Caterina Nardella | Livia Pagano | Valeria Pennacchietti | Mariana Di Felice | Sara Di Matteo | Awa Diop | Francesca Malagrinò | Lucia Marcocci | Paola Pietrangeli | Stefano Gianni | Angelo Toto

Dipartimento di Scienze Biochimiche "A. Rossi Fanelli", Laboratory affiliated to Istituto Pasteur Italia—Fondazione Cenci Bolognetti, Sapienza Università di Roma, Rome, Italy

Correspondence
Stefano Gianni and Angelo Toto, Dipartimento di Scienze Biochimiche "A. Rossi Fanelli", Laboratory affiliated to Istituto Pasteur Italia—Fondazione Cenci Bolognetti, Sapienza Università di Roma, P.le Aldo Moro 5, 00185, Rome, Italy.
Email: stefano.gianni@uniroma1.it and angelo.toto@uniroma1.it

Present address
Francesca Malagrinò, Dipartimento di Farmacia, Università degli Studi di Napoli Federico II, Naples, Italy.

Funding information
European Union's Horizon 2020, Grant/Award Number: 860517; Institut Pasteur Paris, ACIP, Grant/Award Number: 485–21; Istituto Pasteur Italia; Ministero dell'Istruzione, dell'Università e della Ricerca; Regione Lazio, Grant/Award Number: A0375–2020–36559; Sapienza Università di Roma, Grant/Award Numbers: RP11715C34AEAC9B, RM1181641C2C24B9, RM11916B414C897E, RG12017297FA7223, AR22117A3CED340A; Associazione Italiana per la Ricerca sul Cancro, Grant/Award Number: IG 24551

Review Editor: Aitziber L. Cortajarena

Abstract

In an effort to investigate the molecular determinants of ligand recognition of the C-terminal SH2 domain of the SHP2 protein, we conducted extensive site-directed mutagenesis and kinetic binding experiments with a peptide mimicking a specific portion of a physiological ligand (the scaffold protein Gab2). Obtained data provided an in-depth characterization of the binding reaction, allowing us to pinpoint residues topologically far from the binding pocket of the SH2 domain to have a role in the recognition and binding of the peptide. The presence of a sparse energetic network regulating the interaction with Gab2 was identified and characterized through double mutant cycle analysis, performed by challenging all the designed site-directed variants of C-SH2 with a Gab2 peptide mutated at +3 position relative to its phosphorylated tyrosine, a key residue for C-SH2 binding specificity. Results highlighted non-optimized residues involved in the energetic network regulating the binding with Gab2, which may be at the basis of the ability of this SH2 domain to interact with different partners in the intracellular environment. Moreover, a detailed analysis of kinetic and thermodynamic parameters revealed the role of the residue at +3 position on Gab2 in the early and late events of the binding reaction with the C-SH2 domain.

KEYWORDS
binding, double mutant cycle, Gab2, kinetics, protein–protein interaction, SHP2

1 | INTRODUCTION

SH2 domains are among the most abundant protein–protein interaction modules, mediating several key cellular pathways. SH2 domain can recognize short linear interaction motif typically characterized by the presence of a phosphorylated tyrosine (Liu et al., 2011). Tyrosine phosphorylation represents one of the most important post-translational modifications that regulates a vast number of molecular and physiological pathways in the eukaryotic cell. As a consequence, the recognition of such sequences is a key event ensuring correct signal transduction and cell viability, and their dysregulation is at the basis of several pathological states (Diop et al., 2022; Lappalainen et al., 2008; Mayer, 2017; Morlacchi et al., 2014; Zhang et al., 2017).

Despite their structure appears to be highly conserved in the proteome (Waksman et al., 1992; Waksman et al., 1993), a single SH2 domain usually displays a well-defined specificity (Huang et al., 2008; Songyang et al., 1993). In another case, it has been reported that a single SH2 domain may interact with different ligands (Zhang et al., 2011). SH2 domains are ~100 residue globular domains, comprising two α-helices that flank a central β-sheet composed of three to five anti-parallel β strands. The binding pocket with ligands is characterized by a positively charged groove accommodating the phosphotyrosine, and a hydrophobic pocket that interacts with residues located at the C-terminal of the phosphotyrosine, generally referred to as +1, +2, +3, or +4 in the sequence of the ligand (Marasco & Carlomagno, 2020). This allows SH2 domain to bind different partners with distinct affinities and specificities, avoiding non-specific interactions that may lead to harmful reactions and cellular signaling dysregulation.

The molecular determinants that drive the mechanism of interaction of SH2 domain with their ligands are still not completely understood, although a large amount of structural and thermodynamic data is available about SH2 domains in complex with their ligands (Diop et al., 2022; Kuriyan & Cowburn, 1997; Pawson & Gish, 1992; Waksman et al., 2004). In this work, we provide a detailed mutational analysis of the binding reaction occurring between the C-terminal SH2 domain of SHP2 (Asmamaw et al., 2022; Zhang et al., 2015) and a peptide mimicking a portion of the adapter protein Gab2 (Adams et al., 2012) specifically recognized by the domain (Nardella et al., 2021). By employing a double mutant cycle analysis (Horovitz, 1996; Pagano et al., 2021), we characterize an intramolecular energetic network within the SH2 domain regulating the recognition of Gab2 that may be at the basis of its ability to modulate affinity and specificity for a different physiological partner in the cell. Provided data are compared and discussed under the light of previous work on SH2 domains.

2 | RESULTS AND DISCUSSION

2.1 | Mutational analysis of the binding reaction between the C-SH2 domain of SHP2 and $Gab2_{637–649}$

A feasible methodology to infer the details of a binding reaction lies in perturbing a protein system by mutagenesis while monitoring the effect of substitutions on the binding kinetic parameters. Accordingly, we designed, purified, and characterized 33 site-directed variants of the C-SH2 domain by performing a time-resolved kinetic binding experiment. Substitutions were designed to be conservative, by following the generally accepted rules of phi-value analysis (Fersht & Sato, 2004.) By following an analogous approach used in a previous work (Nardella et al., 2021), we analyzed the binding reaction occurring between the C-SH2 variants and a peptide mimicking the portion of Gab2 ranging from residue 637 to 649 ($Gab2_{637–649}$), chemically modified with a dansyl group covalently linked at the N-terminus. This allowed us to follow the binding reaction spectroscopically, exploiting the tryptophan residue in position 112 of the C-SH2 domain and the dansyl group for FRET (Forster Resonance Energy Transfer) experiments.

Kinetic binding experiments were conducted by using a stopped-flow apparatus, by rapidly mixing a fixed concentration of dansylated $Gab2_{637–649}$ (2 μM) with increasing concentrations of the C-SH2 domain variants (from 2 to 10 μM). Buffer used was 50 mM sodium-acetate pH 5.5 (to improve the affinity of the C-SH2 domain for $Gab2_{637–649}$[19]), containing 40% w/v sucrose (corresponding to 1.17 M) and 2 mM 1,4-dithiothreitol (DTT) and the temperature was set to 283.15 K. The addition of sucrose to the buffer increased the viscosity of the solution, reduced the rate of collision of the molecules slowing down the binding reaction and allowing an improved resolution of kinetic traces at the stopped flow. All the binding traces obtained by the time-resolved fluorescence monitoring were satisfactorily fitted with a single-exponential equation (see Materials and Methods section) to calculate binding observed rate constants (k_{obs}). Calculated k_{obs} were plotted versus the different concentrations of the site-directed mutants of the C-SH2 domain and compared with the C-SH2 domain wt (Figure 1). Data were fitted by a linear equation (see Materials and Methods section), with the slope and the y-axis intercept of the line representing the microscopic association (k_{on}) and dissociation rate constants (k_{off}), respectively. Given

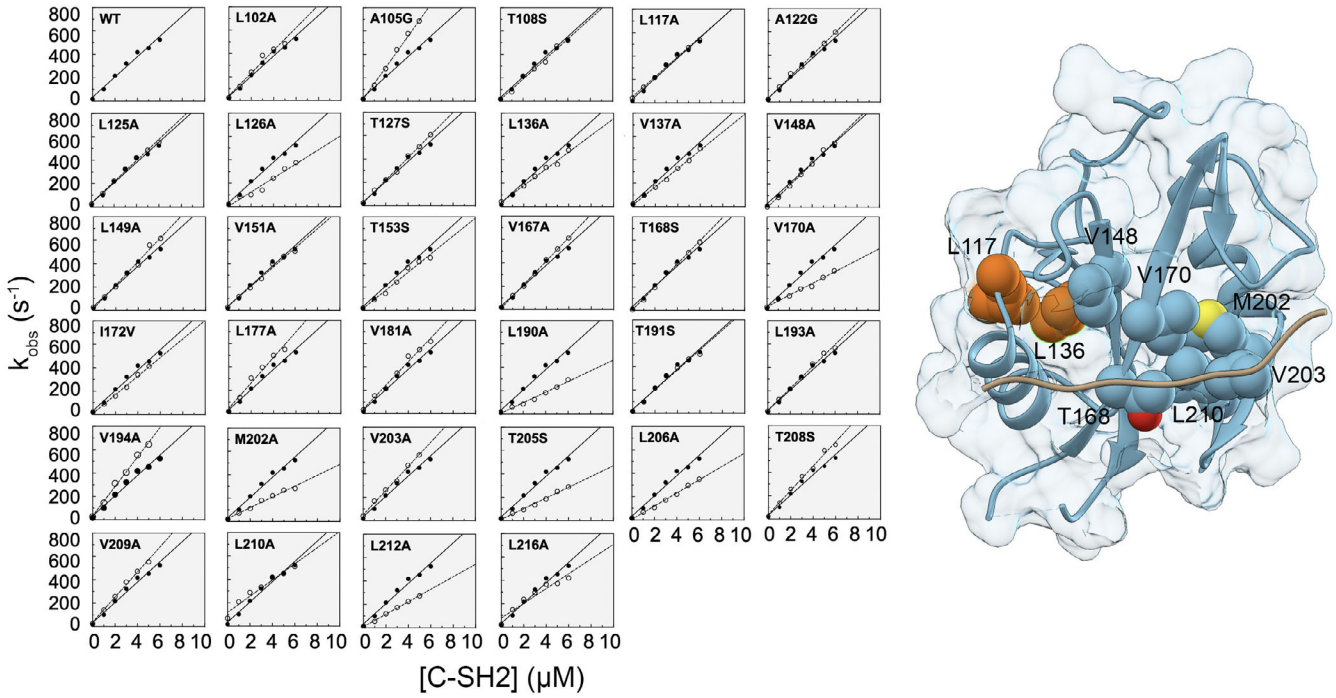

FIGURE 1 Dependences of the binding observed rate constants recorded by rapidly mixing wild-type $Gab2_{637–649}$ with different concentrations of C-SH2 domain wt (black-filled circles) and its site-directed variants (empty circles). Lines represent the best fit to a linear equation. On the right, a graphical representation of the positions targeted for mutagenesis (represented as spheres) on the three-dimensional structure of the C-SH2 domain (obtained from a structural alignment between the PDB 4qsy and PDB 4jeg. Structural alignment was performed by using UCSF Chimera software) reporting a major effect on the affinity for Gab2. In orange, residues L117 and L136, are highlighted as physically far from the binding pocket of the domain.

the high experimental error associated with the extrapolation of k_{off}, this parameter was directly measured by independent displacement experiments, performed by rapidly mixing a pre-formed C-SH2 domain mutant (2 μM) + dansylated $Gab2_{637–649}$ (10 μM) complex with a high excess of non-dansylated $Gab2_{637–649}$ peptide (100 μM) (Antonini & Brunori, 1971). Dependences of k_{obs} as a function of [C-SH2] for all C-SH2 variants are reported in Figure 1 and calculated kinetic and thermodynamic parameters are listed in Table S1. Affinity was calculated as $K_D = k_{off}/k_{on}$. Notably, while most of the substitutions affecting K_D are physically located in the binding site of the SH2 domain (V148A, T168S, V170A, M202A, V203A, and L210A), the positions L117A and L136A appear to be far from the interface between the SH2 domain and Gab2 (highlighted in orange in Figure 1). It is of interest to compare these data with kinetic binding data obtained for the other SH2 domain of the SHP2 protein, that is, the N-SH2 domain, with a different portion of Gab2 (ranging from residue 608 to 620) (Bonetti et al., 2018; Visconti, Malagrinò, et al., 2020). In the case of N-SH2, an extensive mutational analysis of binding kinetic parameters reported an effect on the affinity for $Gab2_{608–620}$ only for residues directly interacting with the ligand. Since the N-SH2 and C-SH2 domains can bind a different portion of the same interactor, our results suggest that the two domains may engage in binding with a different mechanism of recognition, the C-SH2 characterized by the presence of an intramolecular energetic network regulating its binding properties. However, such a feature cannot be depicted from a simple mutational analysis and demands a deeper thermodynamic investigation.

2.2 | Double mutant cycle analysis

Structural biology techniques represent a fundamental tool to analyze the overall complex topology between a protein and its ligand and to fully understand the contacts involved in such events. However, the function of protein–protein interaction domains can be finely regulated by sparce energetic networks (Gianni et al., 2011; Lockless & Ranganathan, 1999; Malagrinò et al., 2019) that may escape a structural characterization, as they may not be associated with detectable changes in conformation or dynamics. A powerful methodology to assess the strength of interactions and energetic coupling

occurring between a ligand and specific residues of the protein that are spatially distant from the binding site is represented by the double mutant cycle analysis. This approach combines extensive site-directed mutagenesis with quantitative measurements of the biophysical properties of a protein system and allows to pinpoint the energetic interactions between discrete residues, even spatially far from the binding pocket, thereby obtaining a map of the energetic contributions of different structural elements of the protein, energetically communicating in exerting protein function. The rationale and details of this methodology have been extensively described before (Cockroft & Hunter, 2007; Horovitz, 1996; Horovitz et al., 2019; Pagano et al., 2021).

SH2 domains are among the most diffuse protein–protein interaction domains able to bind motifs characterized by the presence of a phosphorylated tyrosine, being involved in a large number of molecular pathways. Given their fundamental role and their abundance in the intracellular milieu, the affinity and specificity of SH2 domains for different ligands must be tightly regulated in order to ensure proper signaling. Such regulation is mainly achieved by the recognition of additional residue(s) at the C-terminal of the phosphorylated tyrosine in the ligand (Huang et al., 2008), which in the case of the C-SH2 domain of SHP2, is the residue at +3 position relative to the phosphorylated tyrosine. Thus, to investigate the energetic network regulating the function of C-SH2 domain of SHP2 we resorted to perform a double mutant cycle analysis by performing kinetic binding experiments between all the site-directed variants of C-SH2 and a modified version of the $Gab2_{637–649}$ peptide, that is, V to A in position 646 which corresponds to position +3 relative to the phosphotyrosine. Experimental conditions were the same used for wt $Gab2_{637–649}$. The dependences of k_{obs} as a function of the concentration of all the site-directed variants of C-SH2 are reported in Figure 2 and all the kinetic and thermodynamic parameters calculated are listed in Table S2.

2.3 | Mapping the energetic network regulating the binding of C-SH2 domain with Gab2

The thermodynamic and kinetic parameters obtained from pseudo-first-order binding experiments between the

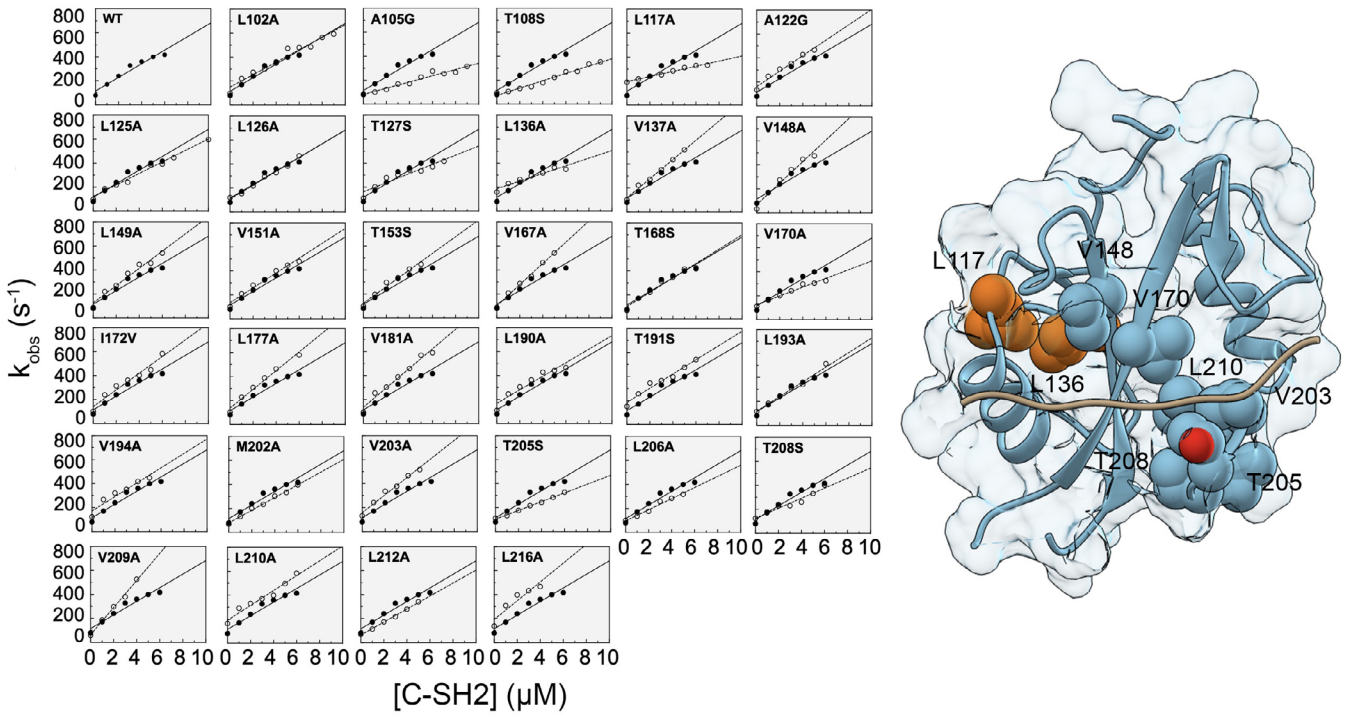

FIGURE 2 Dependences of the binding observed rate constants recorded by rapidly mixing $Gab2_{637–649}$ V646A with different concentrations of C-SH2 domain wt (black-filled circles) and its site-directed variants (empty circles). Lines represent the best fit to a linear equation. On the right, a graphical representation of the positions targeted for mutagenesis (represented as spheres) on the three-dimensional structure of the C-SH2 domain (obtained from a structural alignment between the PDB 4qsy and PDB 4jeg. Structural alignment was performed by using UCSF Chimera software) reporting a major effect on the affinity for Gab2. In orange, residues L117 and L136, are highlighted as physically far from the binding pocket of the domain.

C-SH2 domain and $Gab2_{637–649}$ wt and $Gab2_{637–649}$ V646A, listed in Table S1 and Table S2, were exploited to calculate coupling free energies (ΔΔΔG) that are reported in Table 1. ΔΔΔG values were calculated as follows:

$$\Delta\Delta\Delta G = \Delta\Delta G_{eq}^{Gab2wt} - \Delta\Delta G_{eq}^{Gab2\ V646A}.$$

Notably, we found that four positions, A105G, T108S, L117A, and T168S, reported a ΔΔΔG < −0.4 kcal mol^{-1}, while one position, V209A, reported a positive ΔΔΔG > 0.4 kcal mol^{-1}.

It is of interest to analyze the structural distribution of the residues that reported a detectable ΔΔΔG value, by mapping thermodynamic data on the three-dimensional structure of the C-SH2 domain (Figure 3). Since there is no available structure of the C-SH2 domain in complex with $Gab2_{637–649}$ we performed a structural alignment between the PDB 4qsy, which corresponds to the structure of the N-SH2 domain in complex with Gab1, and PDB 4jeg, which corresponds to the C-SH2

TABLE 1 Thermodynamic parameters and coupling free energies (ΔΔΔG) of the binding reaction of the C-SH2 domain with $Gab2_{637–649}$ wild-type and V646A.

	$\Delta\Delta G_{eq}$ Gab2 wt (kcal Mol^{-1})	$\Delta\Delta G_{eq}$ Gab2 V646A (kcal Mol^{-1})	ΔΔΔG (kcal Mol^{-1})
L102A	0.08 ± 0.06	0.16 ± 0.07	−0.08 ± 0.09
A105G	−0.29 ± 0.04	0.47 ± 0.07	−0.76 ± 0.08
T108S	−0.05 ± 0.05	0.45 ± 0.07	−0.50 ± 0.08
L117A	0.46 ± 0.04	1.04 ± 0.07	−0.57 ± 0.08
A122G	0.04 ± 0.04	0.22 ± 0.08	−0.17 ± 0.09
L125A	0.12 ± 0.04	0.21 ± 0.07	−0.09 ± 0.08
L126A	0.28 ± 0.06	0.22 ± 0.07	0.06 ± 0.09
T127S	0.12 ± 0.04	0.39 ± 0.10	−0.28 ± 0.11
L136A	0.57 ± 0.05	0.71 ± 0.09	−0.14 ± 0.10
V137A	0.18 ± 0.04	−0.06 ± 0.07	0.23 ± 0,08
V148A	−0.81 ± 0.04	−0.77 ± 0.08	−0.04 ± 0.09
L149A	0.11 ± 0.04	−0.08 ± 0.08	0.19 ± 0.09
V151A	0.24 ± 0.04	0.09 ± 0.07	0.16 ± 0.08
T153S	0.10 ± 0.05	0.01 ± 0.08	0.10 ± 0.09
V167A	−0.08 ± 0.04	−0.19 ± 0.06	0.11 ± 0.07
T168S	−0.66 ± 0.04	−0.10 ± 0.08	−0.55 ± 0.09
V170A	0.54 ± 0.05	0.44 ± 0.08	0.10 ± 0.09
I172V	0.10 ± 0.04	0.03 ± 0.08	0.07 ± 0.09
L177A	−0.12 ± 0.06	−0.04 ± 0.07	−0.08 ± 0.09
V181A	0.26 ± 0.05	−0.11 ± 0.07	0.38 ± 0.09
L190A	0.27 ± 0.05	0.18 ± 0.10	0.08 ± 0.11
T191S	−0.04 ± 0.05	0.33 ± 0.08	−0.37 ± 0.09
L193A	0.14 ± 0.04	0.17 ± 0.07	−0.03 ± 0.08
V194A	0.01 ± 0.04	0.21 ± 0.11	−0.20 ± 0.12
V203A	0.33 ± 0.04	0.12 ± 0.07	0.21 ± 0.08
T205S	0.34 ± 0.04	0.43 ± 0.06	−0.09 ± 0.08
L206A	0.18 ± 0.04	0.24 ± 0.07	−0.06 ± 0.09
T208S	0.03 ± 0.05	0.42 ± 0.08	−0.39 ± 0.10
V209A	0.04 ± 0.04	−0.56 ± 0.07	0.60 ± 0.08
L210A	0.86 ± 0.06	0.36 ± 0.13	0.05 ± 0.14
L212A	−0.02 ± 0.04	−0.06 ± 0.08	0.04 ± 0.09
L216A	0.29 ± 0.08	0.12 ± 0.14	0.17 ± 0.16
M202A	0.37 ± 0.05	−0.04 ± 0.08	0.41 ± 0.10

domain in complex with monobody inhibitor CS1. Structural alignment was performed by using UCSF Chimera software (Pettersen et al., 2004). This allowed us to obtain a graphical representation of the C-SH2 domain and a ligand, in order to spatially highlight the position of the binding pocket, which is reported in Figure 3 (as well as in Figure 1 and Figure 2). Interestingly, T168S is physically located in the binding cleft of the C-SH2 domain, suggesting that this residue may directly interact with Gab2 in the binding reaction, while A105G, T108S, L117A, and V209A positions appear to be far from the binding pocket. This aspect let us conclude that although T108, L117, and V209 do not engage a direct contact with the ligand during binding, they are part of a sparse energetic network within the C-SH2 domain that finely regulate its function and specificity, playing a key role in the recognition of $Gab2_{637–649}$.

It is also important to notice that the ΔΔΔG values of the residues reported in Figure 3 are prevalently negative. This result implies that the overall effect of substitutions on the binding appears to be higher when the reaction is monitored versus the mutated ligand as compared to the wild-type peptide. Thus, it may be concluded that the sequence of C-SH2 appears to be not optimized for binding with Gab2. Our data are also in accordance with what has been previously observed for the C-terminal SH3 domain of Grb2, analyzing its binding reaction with a different portion of Gab2 (Malagrinò et al., 2019). Both SHP2 and Grb2 mediate the interactions with several partners in the intracellular environment, being part of and regulating several complex molecular pathways for the cell (Asmamaw et al., 2022; Belov & Mohammadi, 2012; Giubellino et al., 2008; Jang et al., 2009; Zhang et al., 2015). Thus, our results suggest that the presence of a sub-optimal sequence might be a general feature that may allow protein–protein interaction domains to accommodate different interactors in the binding pocket and mediate the recognition and binding with several distinct partners.

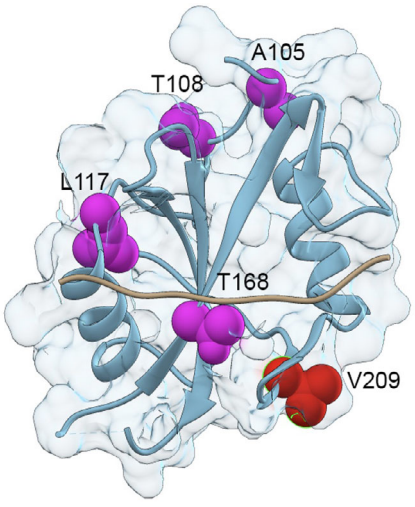

FIGURE 3 Energetic network regulating the binding of the C-SH2 domain of SHP2 with Gab2. Residues reported as magenta spheres (A105, T108, L117, and T168) are the ones reporting a $\Delta\Delta\Delta G < -0.4$ kcal mol^{-1}, in red the residue (V209) with a $\Delta\Delta\Delta G > 0.4$ kcal mol^{-1}.

Complex binding mechanisms implying energetic networks underlying the function of other SH2 domains have been established and characterized in the past (Huculeci et al., 1993; Lubman & Waksman, 2002). In particular, the SH2 domain of Fyn displayed a thermodynamic link between residues located in the binding site to residues located throughout the whole structure. Interestingly, a few residues in the linker between the SH2 and SH3 domain of Fyn appeared to be perturbed upon binding, suggesting an interdomain cross-talk (Huculeci et al., 1993), suggesting that the presence of complex energetic networks regulating the function of the SH2 domain may represent a general feature of this protein–protein interaction module. Furthermore, our data reinforce the previously proposed concept that SH2 domains tend, in general, to display suboptimal affinities for their ligands because a too-strong interaction might lead to aberrant signaling in the cell (Kaneko et al., 2012). This would imply an evolutive pressure on SH2 domains not to improve their affinities to the point of disrupting fundamental molecular pathways. It has also to be considered that protein–protein interactions strongly depend on the intracellular microenvironment in which they occur (Ladbury & Arold, 2000; Zarrinpar et al., 2003), as well as on the presence of contiguous multiple domains (Ottinger et al., 1998; Stiegler et al., 2022), that can possibly modulate the affinity for the ligand and increase the complexity of the regulation of physiological pathway mediated by those interactions.

2.4 | Deciphering the role of Gab2 V646 residue in the early and late events of binding through LFER analysis

Linear Free Energy Relationship (LFER) analysis correlates the change in activation free energy ($\Delta G^{\#}$) to the change in equilibrium free energy (ΔG_{eq}) and is usually exploited to determine the position of the transition state of a given reaction along the reaction coordinate in organic chemistry, enzymology, protein folding, and protein–protein interaction studies (Eaton et al., 1991; Edelstein & Changeux, 2010; Giri et al., 2013; Haq et al., 2012; Leffler, 1953; Matouschek & Fersht, 1993; Toto et al., 2016; Visconti et al., 2019). In the latter case,

LFER analysis can be particularly useful to understand the effect of given perturbations on the kinetic parameters, allowing us to pinpoint the contribution of single residues on the early and late events of the binding reaction.

In Figure 4a,b, we reported the dependence of the logarithm of k_{on} and k_{off} as a function of the logarithm of K_D measured for each C-SH2 variant. Data clearly show that substitutions have minor effects on the early event of the recognition occurring between C-SH2 and Gab2 (the slope of log k_{on} vs. log K_D being -0.25 ± 0.08, with a $R^2 = 0.23$). On the other hand, a greater contribution of the microscopic dissociation rate constant on the stability of the complex is evident, with k_{off} increasing at higher K_D values (the slope of log k_{off} vs. log K_D being 0.76 ± 0.08, with a $R^2 = 0.75$). Given the nature of designed site-directed variations, occurring only on hydrophobic side-chains and without perturbing polar and charged residues (Fersht & Sato, 2004), this result well correlates with a scenario in which the early recognition events are driven by electrostatic contributions, while non-polar interactions lock the complex in place in the late events of binding, in analogy to what has been previously described for other SH2 domains (Bonetti et al., 2018; Nardella et al., 2021; Nardella et al., 2022; Visconti, Malagrinò, et al., 2020; Visconti, Toto, et al., 2020).

Interestingly, LFER analysis of $Gab2_{637–649}$ V646A reports a linear correlation of both log k_{on} (the slope of log k_{on} vs. log K_D being -0.54 ± 0.05, with a $R^2 = 0.73$) and log k_{off} (the slope of log k_{off} vs. log K_D being 0.46 ± 0.06, with a $R^2 = 0.65$) as a function of the logarithm of the equilibrium dissociation constant. As we briefly reported in the Introduction, the presence of a hydrophobic pocket in the binding cleft of SH2 domains allows them to interact with residues located at the C-terminal of the phosphotyrosine. LFER analysis and binding kinetics unequivocally demonstrate a direct involvement of V646 residue of Gab2 in the binding with C-SH2. Altogether, these results suggest a synergistic contribution of electrostatic interactions, driven by the phosphorylated tyrosine, and hydrophobic interactions, mediated by V646, to the recognition of Gab2 by the C-SH2 domain of SHP2, indicating V646 as a key mediator of specificity in both the early and late events of binding.

3 | CONCLUSIONS

SH2 domains play a fundamental role in several molecular and physiological pathways, acting downstream of tyrosine kinases. The interactions SH2 domains mediate must be highly specific, in spite of sharing a highly

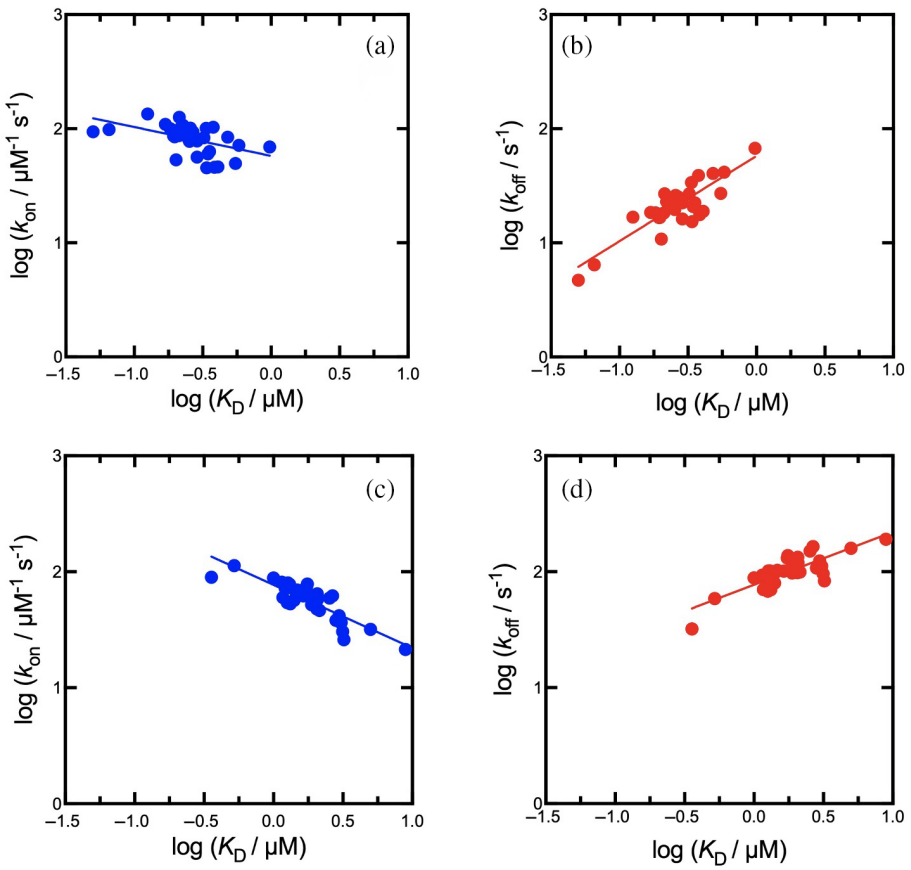

FIGURE 4 Linear free energy relationships for the binding of $Gab2_{637–649}$ wild-type (panels a and b) and V646A (panels c and d) peptides with C-SH2. (a) and (c) Dependences of the microscopic association rate constant k_{on} on the equilibrium dissociation constant. (b) and (d) Dependences of the microscopic dissociation rate constant k_{off} on the equilibrium dissociation constant.

conserved three-dimensional structure in the proteome. Thus, together with a structural characterization, a quantitative analysis of its thermodynamic components is essential to comprehensively understand these protein systems. The double mutant cycle analysis represents a powerful approach for the determination of sparse energetic networks underlying protein functions and for measuring the energetic coupling occurring between residues of a protein that are located far from the binding site and a ligand. This methodology combines extensive site-directed mutagenesis with quantitative measurements of the biophysical properties of a protein system. The employment of double mutant cycle methodology allowed us to map an intramolecular energetic network that takes place in the binding reaction of C-SH2 with Gab2, regulating its binding specificity, that would possibly escape a structural characterization, since there may not be any detectable structural change, upon binding, for residues located far from the binding pocket. Importantly, through a rigorous analysis of thermodynamic data, we could highlight a scenario in which residues involved in this energetic network appear to be not optimized for binding with Gab2. This aspect suggests that such residues, which are not directly involved in the function of the protein, can still modulate binding through the intramolecular energetic network that may be at the basis of the ability of C-SH2 to accommodate different partners to its binding pocket in the intracellular environment. However, although these 'allosteric' sites may represent possible targets for pharmacological strategies aimed to modulate the binding properties of protein–protein interaction domains, the functional relevance of the energetic networks in which they are involved is currently the subject of debate (Gianni & Jemth, 2023). Moreover, the analysis of thermodynamic parameters allowed us to highlight the role of V646 residue of Gab2 in the binding with C-SH2, pinpointing a dual contribution from electrostatics and hydrophobic interactions in the early and late events of binding with the ligand. Future structural characterization on this and other SH2 domains will further clarify the possible generality of this binding mechanism.

4 | MATERIALS AND METHODS

4.1 | Protein expression and purification

The C-SH2 domain (residues 97–220) of the SHP2 protein (Uniprot Q06124) was expressed and purified as previously reported (Nardella et al., 2021). Site-directed variants of the C-SH2 domain were obtained using QuikChange® Lightning kit from Agilent Technologies, according to the instructions of the manufacturer. Peptides mimicking the region 637–649 of Gab2 (Uniprot Q9UQC2), sequences SDEKVDpYVQVDKE for wt and SDEKVDpYVQADKE for V646A variant, with phosphorylated tyrosine in position 643, with and without the dansyl N-terminal modification, were purchased from Genscript Biotech Corp (purity >90%).

4.2 | Kinetic binding and displacement experiments

Kinetic binding experiments were performed using a SX-18 stopped-flow apparatus (Applied Photophysics), by challenging a constant concentration of dansylated Gab2 peptide (2 μM) with increasing concentrations of the C-SH2 domain variants (from 2 to 10 μM). Buffer used was 50 mM sodium-acetate pH 5.5, containing 40% w/v sucrose and 2 mM 1,4-dithiothreitol (DTT). The temperature was set to 283.15 K. Samples were excited at 280 nm and FRET signal was collected through a 475 nm cut-off filter. For each experiment conducted, an average calculated from 3 to 5 independently acquired kinetic traces was satisfactorily fitted with a single exponential equation to obtain the observed rate constant k_{obs}.

$$y = a * exp(-k_{obs} * t) + c.$$

Dependences of k_{obs} as a function of the concentration of C-SH2 domain were fitted with the following linear equation:

$$k_{obs} = k_{on}[CSH2] + k_{off}.$$

Displacement experiments were employed to directly calculate the microscopic dissociation, and performed by rapidly mixing a pre-formed C-SH2 domain mutant (2 μM) + dansylated $Gab2_{637–649}$ (10 μM) complex versus a high excess of non-dansylated $Gab2_{637–649}$ peptide (100 μM).

AUTHOR CONTRIBUTIONS

Caterina Nardella: investigation (lead), methodology (equal), formal analysis (equal), writing–review and editing (equal); **Livia Pagano**: investigation (equal), writing–review and editing (equal); **Valeria Pennacchietti**: investigation (equal), writing–review and editing (equal); **Mariana Di Felice**: investigation (equal), writing–review and editing (equal); **Sara Di Matteo**: investigation (equal), writing–review and editing (equal); **Awa Diop**: investigation (equal), writing–review and editing (equal); **Francesca Malagrinò**: investigation

(equal), writing–review and editing (equal); **Lucia Marcocci**: investigation (equal), writing–review and editing (equal); **Paola Pietrangeli**: investigation (equal), writing–review and editing (equal); **Stefano Gianni**: conceptualization (equal), formal analysis (equal), funding acquisition (equal), methodology (equal), writing–review and editing (lead); **Angelo Toto**: conceptualization (equal), formal analysis (lead), funding acquisition (equal), methodology (equal), writing–original draft preparation (lead), writing–review and editing (equal).

ACKNOWLEDGMENTS

This work was partly supported by grants from the Italian Ministero dell'Istruzione dell'Università e della Ricerca (Progetto di Interesse "Invecchiamento" to Stefano Gianni), Sapienza University of Rome (RP11715 C34AEAC9B and RM1181641C2C24B9, RM11916B414-C897E, RG12017297FA7223 to Stefano Gianni, AR22117-A3CED340A to Caterina Nardella), by an ACIP grant (ACIP_485–21) from Institut Pasteur Paris to Stefano Gianni, the Associazione Italiana per la Ricerca sul Cancro (Individual Grant—IG 24551 to Stefano Gianni), the Regione Lazio (Progetti Gruppi di Ricerca LazioInnova A0375–2020–36,559 to Stefano Gianni), European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement UBIMOTIF No 860517 (to Stefano Gianni) the Istituto Pasteur Italia, "Teresa Ariaudo Research Project" 2018, and "Research Program 2022 to 2023 Under 45 Call 2020" (to Angelo Toto). We ackowledge co-funding from Next Generation EU, in the context of the National Recovery and Resilience Plan, Investment PE8-Project Age-It: "Ageing Well in an Ageing Society." This resource was co-financed by the Next Generation EU [DM 1557 11 October 2022]. The views and opinions expressed are only those of the authors and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

CONFLICT OF INTEREST STATEMENT

All authors declare no conflict of interest.

ORCID

Stefano Gianni https://orcid.org/0000-0003-1653-1925
Angelo Toto https://orcid.org/0000-0002-4620-3677

REFERENCES

Adams SJ, Aydin IT, Celebi JT. GAB2—a scaffolding protein in cancer. Mol Cancer Res MCR. 2012;10:1265–70.

Antonini E, Brunori M. Hemoglobin and myoglobin in their reactions with ligands. North Holl: Amst; 1971.

Asmamaw MD, Shi X-J, Zhang L-R, Liu H-M. A comprehensive review of SHP2 and its role in cancer. Cell Oncol Dordr. 2022; 45:729–53.

Belov AA, Mohammadi M. Grb2, a double-edged sword of receptor tyrosine kinase signaling. Sci Signal. 2012;5:pe49. https://doi.org/10.1126/scisignal.2003576

Bonetti D, Troilo F, Toto A, Travaglini-Allocatelli C, Brunori M, Gianni S. Mechanism of folding and binding of the N-terminal SH2 domain from SHP2. J Phys Chem B. 2018; 122:11108–14.

Cockroft SL, Hunter CA. Chemical double-mutant cycles: dissecting non-covalent interactions. Chem Soc Rev. 2007;36:172–88.

Diop A, Santorelli D, Malagrinò F, Nardella C, Pennacchietti V, Pagano L, et al. SH2 domains: folding, binding and therapeutical approaches. Int J Mol Sci. 2022;23:15944.

Eaton WA, Henry ER, Hofrichter J. Application of linear free energy relations to protein conformational changes: the quaternary structural change of hemoglobin. Proc Natl Acad Sci. 1991;88:4472–5.

Edelstein SJ, Changeux J-P. Relationships between structural dynamics and functional kinetics in oligomeric membrane receptors. Biophys J. 2010;98:2045–52.

Fersht AR, Sato S. Phi-value analysis and the nature of protein-folding transition states. Proc Natl Acad Sci U S A. 2004;101: 7976–81.

Gianni S, Haq SR, Montemiglio LC, Jürgens MC, Engström Å, Chi CN, et al. Sequence-specific long range networks in PSD-95/discs large/ZO-1 (PDZ) domains tune their binding selectivity. J Biol Chem. 2011;286:27167–75.

Gianni S, Jemth P. Allostery frustrates the experimentalist. J Mol Biol. 2023;435:167934.

Giri R, Morrone A, Toto A, Brunori M, Gianni S. Structure of the transition state for the binding of c-Myb and KIX highlights an unexpected order for a disordered system. Proc Natl Acad Sci. 2013;110:14942–7.

Giubellino A, Burke TR, Bottaro DP. Grb2 signaling in cell motility and cancer. Expert Opin Ther Targets. 2008;12:1021–33.

Haq SR, Chi CN, Bach A, Dogan J, Engström Å, Hultqvist G, et al. Side-chain interactions form late and cooperatively in the binding reaction between disordered peptides and PDZ domains. J Am Chem Soc. 2012;134:599–605.

Horovitz A. Double-mutant cycles: a powerful tool for analyzing protein structure and function. Fold Des. 1996;1:R121–6.

Horovitz A, Fleisher RC, Mondal T. Double-mutant cycles: new directions and applications. Curr Opin Struct Biol. 2019;58: 10–7.

Huang H, Li L, Wu C, Schibli D, Colwill K, Ma S, et al. Defining the specificity space of the human SRC homology 2 domain. Mol Cell Proteomics MCP. 2008;7:768–84.

Huculeci R, Cilia E, Lyczek A, Buts L, Houben K, Seeliger MA, et al. Lenaerts T (2016) dynamically coupled residues within the SH2 domain of FYN are key to unlocking its activity. Struct Lond Engl. 1993;24:1947–59.

Jang IK, Zhang J, Gu H. Grb2, a simple adapter with complex roles in lymphocyte development, function, and signaling. Immunol Rev. 2009;232:150–9.

Kaneko T, Huang H, Cao X, Li X, Li C, Voss C, et al. Superbinder SH2 domains act as antagonists of cell signaling. Sci Signal. 2012;5:ra68.

Kuriyan J, Cowburn D. Modular peptide recognition domains in eukaryotic signaling. Annu Rev Biophys Biomol Struct. 1997; 26:259–88.

Ladbury JE, Arold S. Searching for specificity in SH domains. Chem Biol. 2000;7:R3–8.

Lappalainen I, Thusberg J, Shen B, Vihinen M. Genome wide analysis of pathogenic SH2 domain mutations. Proteins. 2008;72: 779–92.

Leffler JE. Parameters for the description of transition states. Science. 1953;117:340–1.

Liu BA, Shah E, Jablonowski K, Stergachis A, Engelmann B, Nash PD. The SH2 domain-containing proteins in 21 species establish the provenance and scope of phosphotyrosine signaling in eukaryotes. Sci Signal. 2011;4:ra83.

Lockless SW, Ranganathan R. Evolutionarily conserved pathways of energetic connectivity in protein families. Science. 1999;286: 295–9.

Lubman OY, Waksman G. Dissection of the energetic coupling across the Src SH2 domain-tyrosyl phosphopeptide interface. J Mol Biol. 2002;316:291–304.

Malagrinò F, Troilo F, Bonetti D, Toto A, Gianni S. Mapping the allosteric network within a SH3 domain. Sci Rep. 2019;9:8279.

Marasco M, Carlomagno T. Specificity and regulation of phosphotyrosine signaling through SH2 domains. J Struct Biol X. 2020; 4:100026.

Matouschek A, Fersht AR. Application of physical organic chemistry to engineered mutants of proteins: Hammond postulate behavior in the transition state of protein folding. Proc Natl Acad Sci. 1993;90:7814–8.

Mayer BJ. What have we learned from SH2 domains? Methods Mol Biol Clifton NJ. 2017;1555:37–43.

Morlacchi P, Robertson FM, Klostergaard J, McMurray JS. Targeting SH2 domains in breast cancer. Future Med Chem. 2014;6: 1909–26.

Nardella C, Malagrinò F, Pagano L, Rinaldo S, Gianni S, Toto A. Determining folding and binding properties of the C-terminal SH2 domain of SHP2. Protein Sci Publ Protein Soc. 2021;30: 2385–95.

Nardella C, Toto A, Santorelli D, Pagano L, Diop A, Pennacchietti V, et al. Folding and binding mechanisms of the SH2 domain from Crkl. Biomolecules. 2022;12:1014.

Ottinger EA, Botfield MC, Shoelson SE. Tandem SH2 domains confer high specificity in tyrosine kinase signaling. J Biol Chem. 1998;273:729–35.

Pagano L, Toto A, Malagrinò F, Visconti L, Jemth P, Gianni S. Double mutant cycles as a tool to address folding, binding, and allostery. Int J Mol Sci. 2021;22:828.

Pawson T, Gish GD. SH2 and SH3 domains: from structure to function. Cell. 1992;71:359–62.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera?A visualization system for exploratory research and analysis. J Comput Chem. 2004;25: 1605–12.

Songyang Z, Shoelson SE, Chaudhuri M, Gish G, Pawson T, Haser WG, et al. SH2 domains recognize specific phosphopeptide sequences. Cell. 1993;72:767–78.

Stiegler AL, Vish KJ, Boggon TJ. Tandem engagement of phosphotyrosines by the dual SH2 domains of p120RasGAP. Structure. 2022;30:1603–1614.e5.

Toto A, Camilloni C, Giri R, Brunori M, Vendruscolo M, Gianni S. Molecular recognition by templated folding of an intrinsically disordered protein. Sci Rep. 2016;6:21994.

Visconti L, Malagrinò F, Gianni S, Toto A. Structural characterization of an on-pathway intermediate and transition state in the folding of the N-terminal SH2 domain from SHP2. FEBS J. 2019;286:4769–77.

Visconti L, Malagrinò F, Pagano L, Toto A. Understanding the mechanism of recognition of Gab2 by the N-SH2 domain of SHP2. Life Basel Switz. 2020;10:85.

Visconti L, Toto A, Jarvis JA, Troilo F, Malagrinò F, De Simone A, et al. Demonstration of binding induced structural plasticity in a SH2 domain. Front Mol Biosci. 2020;7:89.

Waksman G, Kominos D, Robertson SC, Pant N, Baltimore D, Birge RB, et al. Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. Nature. 1992;358:646–53.

Waksman G, Kumaran S, Lubman O. SH2 domains: role, structure and implications for molecular medicine. Expert Rev Mol Med. 2004;6:1–18.

Waksman G, Shoelson SE, Pant N, Cowburn D, Kuriyan J. Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. Cell. 1993;72:779–90.

Zarrinpar A, Park S-H, Lim WA. Optimization of specificity in a cellular protein interaction network by negative selection. Nature. 2003;426:676–80.

Zhang J, Zhang F, Niu R. Functions of Shp2 in cancer. J Cell Mol Med. 2015;19:2075–83.

Zhang M, Jang H, Gaponenko V, Nussinov R. Phosphorylated calmodulin promotes PI3K activation by binding to the SH2 domains. Biophys J. 2017;113:1956–67.

Zhang Y, Zhang J, Yuan C, Hard RL, Park I-H, Li C, et al. Simultaneous binding of two peptidyl ligands by a SRC homology 2 domain. Biochemistry. 2011;50:7637–46.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Nardella C, Pagano L, Pennacchietti V, Felice MD, Matteo SD, Diop A, et al. An intramolecular energetic network regulates ligand recognition in a SH2 domain. Protein Science. 2023;32(8):e4729. https://doi.org/10.1002/pro.4729