



OPEN

# The dynamics of free and phosphopeptide-bound Grb2-SH2 reveals two dynamically independent subdomains and an encounter complex with fuzzy interactions

Karoline Sanches<sup>3</sup>, Icaro P. Caruso<sup>1,2,3</sup>, Fabio C. L. Almeida<sup>1,2</sup>✉ & Fernando A. Melo<sup>3</sup>✉

The growth factor receptor-bound protein 2 (Grb2) is a key factor in the regulation of cell survival, proliferation, differentiation, and metabolism. In its structure, the central Src homology 2 (SH2) domain is flanked by two Src homology 3 (SH3). SH2 is the most important domain in the recognition of phosphotyrosines. Here, we present the first dynamical characterization of Grb2-SH2 domain in the free state and in the presence of phosphopeptide EpYINSQV at multiple timescales, which revealed valuable information to the understanding of phosphotyrosine sensing mechanism. Grb2-SH2 presented two dynamically independent subdomains, subdomain I involved in pY recognition and subdomain II is the pY + 2 specificity pocket. Under semi-saturated concentrations of pY-pep we observed fuzzy interactions, which led to chemical exchange observed by NMR. This information was used to describe the encounter complex. The association with pY-pep is dynamic, involving fuzzy interactions and multiple conformations of pY-pep with negative and hydrophobic residues, creating an electrostatic-potential that drives the binding of pY-pep. The recognition face is wider than the binding site, with many residues beyond the central SH2 binding site participating in the association complex, which contribute to explain previously reported capability of Grb2 to recognize remote pY.

## Abbreviations

Grb2	Growth factor receptor-bound protein 2
SH2	Src homology 2
MAPK	Mitogen activated protein kinase
NMR	Nuclear magnetic resonance
CPMG	Carr–Purcell–Meiboom–Gill pulse train
CSP	Chemical shift perturbation
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation

Cell survival, control, proliferation, differentiation, and metabolism are mediated by a variety of well-orchestrated series of events. In this context, the growth factor receptor-bound protein 2 (Grb2) has demonstrated to be a key factor in regulating many cellular events. Grb2 is not an enzyme, instead, its multiple domains and flexible linkers provide the ability to bind to multiple partners in the cell. Grb2 is a pivotal intermediate in the

<sup>1</sup>Institute of Medical Biochemistry – IBqM, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. <sup>2</sup>National Center for Structural Biology and Bioimaging (CENABIO)/National Center for Nuclear Magnetic Resonance (CNRMN), Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. <sup>3</sup>Multuser Center for Biomolecular Innovation (CMIB), Department of Physics, São Paulo State University (UNESP), São Jose do Rio Preto, São Paulo, Brazil. ✉email: falmeida@bioqmed.ufrj.br; fernando.melo@unesp.br

communication between the cell-surface and downstream signaling<sup>1</sup>. Recently, Grb2 has been reported to have equilibrium between monomeric and dimeric states, which is critical for the activation and regulation of the Ras/mitogen-activated protein kinases (MAPKs) pathway<sup>2</sup>. The plasticity of Grb2 is essential for its biological function. It is widely known that the deregulation of MAPK pathway leads to many diseases, including cancer and developmental defects<sup>3–6</sup>.

Grb2 is composed of a central Src homology 2 (SH2) domain flanked by two Src homology 3 (SH3), a N- and C-terminus. The SH3 domain is well-known to recognize proline-rich sequences with the PxxP motif, while SH2 is the most important domain in the recognition of phosphotyrosines (pY)<sup>7–9</sup>. The SH2 has approximately 100 amino acid residues in globular modules with a well-conserved structure of a central antiparallel  $\beta$ -sheet flanked on each side by an  $\alpha$ -helix. A remarkable point in the SH2 domains is the presence of two distinct binding pockets for the phosphopeptide recognition, where the pocket I (site I) is responsible for the pY binding to two conserved residues, R $\beta$ B5 (R186) and H $\beta$ D4 (H107), and the pocket II (site II) for the pY + 1, pY + 2 or even pY + 3 recognition<sup>8,10</sup>. Site II accounts for the specificity of the interaction. The phosphopeptide binds to SH2 as an extended linear conformation at the domain surface that spans from one  $\alpha$ -helix to the other. An exception for this linear conformation is the interaction with the SH2 domain of Grb2, in which the W121 (site II) is a barrier, hindering the pY + 3 interaction. Grb2-SH2 preferentially selects pY + 2 asparagine sequences<sup>11,12</sup>.

Grb2-SH2 domain is able to form a domain-swapped dimer<sup>13–16</sup>, in which both  $\alpha$ -helix 3 ( $\alpha$ 3) and part of the loop are involved in a major conformational change. The  $\alpha$ 3 opens and associates with the adjacent subunit of the dimer. The domain-swapped dimer is able to bind phosphopeptides with different affinities when compared to the monomeric state, sometimes with higher affinity and others with lower affinity<sup>14,15</sup>, contributing as another degree of freedom for Grb2 plasticity and ability of regulation. The biological role of the swapped domain is not well-known and there is no description of a cellular event with the participation of the swapped Grb2 dimer. The only structure available of the full-length Grb2 is not swapped<sup>17</sup>.

Here we present the first description of the dynamics of Grb2-SH2 domain. We verified that the Grb2-SH2 domain presents two dynamically independent subdomains, one of them in fast exchange regime (subdomain I) and other in intermediate exchange (subdomain II). The thermodynamic profile for a two-state conformational equilibrium of subdomain II in intermediate exchange regime was evaluated. We also evaluated the influence of Grb2-SH2 dynamics in the recognition of endothelial growth factor receptor (EGFR) derived phosphopeptide EpYINSQV (pY-pep) in a saturated and semi-saturated concentration. We characterized a recognition mechanism that involves the formation of an encounter complex as “binding intermediates”. Lastly, we demonstrated that the association of pY-pep to Grb2-SH2 is fuzzy and dynamic, which is essential for the molecular recognition of phosphotyrosines, where many residues beyond the central Grb2-SH2 binding site participates.

## Results

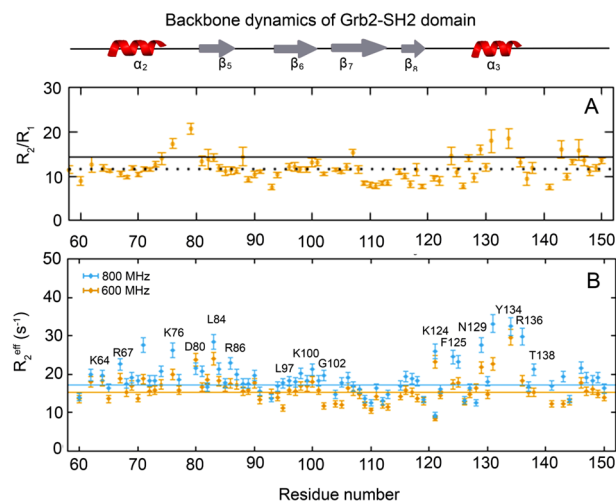
Our main purpose is the measurement of the dynamic properties of Grb2-SH2 domain and its role in the binding of pY-pep. This is the first description of the dynamics of Grb2-SH2 domain in the presence and absence of phosphopeptide. We chose the EGFR derived phosphopeptide EpYINSQV (pY-pep) because it leads to the dissociation of the Grb2, as described by Yuzawa et al. (2001)<sup>18</sup>. We measured the <sup>15</sup>N nuclear spin relaxation parameters  $R_1$ ,  $R_2$  and <sup>1</sup>H-<sup>15</sup>N-heteronuclear NOE (Figure S1), which showed that Grb2-SH2 domain is mostly rigid, displaying only a few residues involved in thermal motion (pico- to nanosecond timescale). They are mainly in the loop between  $\alpha$ 2 and  $\beta$ 5. The most evident feature was the presence of an increased value of the  $R_2/R_1$  ratio (Fig. 1A) in two regions: the first involving residues in  $\alpha$ 2, loop  $\alpha$ 2/ $\beta$ 5 and  $\beta$ 5, and the second in the loop  $\beta$ 8/ $\alpha$ 3 and  $\alpha$ 3. We assigned<sup>19</sup> (BMRB ID 27,781) and calculated the structure of the Grb2-SH2 domain at pH 7.0 (PDB\_id 6VK2) using CS-Rosetta<sup>20</sup> with the ambiguously and unambiguously assigned NOEs derived from Aria/CNS (Figure S2, Table S1). A detailed description of the structure calculation is in the supplementary material.

These regions are in conformational exchange, and to get more insights into the dynamics we measured the <sup>15</sup>N CPMG-RD at four temperatures and two fields (Figures S3, S4, S5, and S6). The CPMG-RD data (Fig. 1A) confirmed the observation of conformational exchange (increased  $R_2/R_1$  ratio, Fig. 1B). The residues in the first region are at the site I, the phosphotyrosine (pY) recognition site and the residues in the second region are at the site II, the pY + 2 recognition site (Fig. 2A).

The analysis of the CPMG-RD curves showed that the residues in conformational exchange at site I are in fast exchange regime, while the residues at site II are in intermediate exchange regime (Figures S3, S4, S5, and S6). Although they belong to the same domain, they are dynamically independent. Grb2-SH2 has two dynamically independent subdomains: subdomain I (Fig. 2A, residues in green) for the residues in fast exchange involved in recognition site I, and subdomain II (Fig. 2A, residues in pink) for the residues in intermediate exchange involved in the recognition site II.

For the dynamic characterization of subdomain I, we looked for residues with no dispersion, but displaying  $R_{2,eff}$  above exchange-free  $R_{2,eff}^\infty$ . The most likely value  $R_{2,eff}^\infty$  is represented by the dotted line in Figs. 1B and 2B, which represents an expected average value of  $R_{2,eff}$  for residue without exchange contribution<sup>21</sup>. The typical plot of <sup>15</sup>N CPMG data observed for fast exchange regimes is where all the points in a flat profile are above the dotted line, meaning that much larger values of  $\nu_{CPMG}$  would be necessary to refocus the exchange contribution to  $R_{2,eff}^\infty$  (Fig. 2B). For details, see Figures S3, S4, S5, and S6.

As subdomain II presents residues in intermediate exchange regime, it was possible to characterize its thermodynamic profile for a two-state conformational equilibrium. We fitted the <sup>15</sup>N CPMG-RD experiments at multiple temperatures using the Bloch-McConnell equation, which describes the evolution of the magnetization in a two-state exchange<sup>22,23</sup>. The Table S1 shows that the two-state exchange regime describes well the dynamics of subdomain II. To evaluate the quality of fitting and validity of the two-state exchange model we compared the statistical parameters ( $\chi^2$  and degrees of freedom, DF) of individual fits for each residue, global fit at each



**Figure 1.** (A) Summary of  $R_2/R_1$  at 18.8 T (800 MHz) and the (B)  $^{15}\text{N}$  CPMG-RD  $R_2^{\text{eff}}$  at  $66.7 \text{ s}^{-1}$  at 14.09 (600 MHz, orange) and 18.8 T (800 MHz, blue) as function of the residue number, acquired at 25 °C. The lines represent the average  $R_2^{\text{eff}}$  for both fields. In A, the dotted line represents the average  $R_2/R_1$  value and the full line the average value plus one standard deviation. The residues above the average lines are in conformational exchange and are labeled in B. There are two main regions in conformational exchange, subdomain I which involves  $\alpha_2$  and  $\beta_5$ , and subdomain II, at  $\alpha_3$ . These data are complemented by  $R_1$ ,  $R_2$  and  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NOE (Figure S2) and by  $^{15}\text{N}$  CPMG-RD experiments at 5, 10, and 17 °C (Figures S3, S4, S5, and S6).

temperature, and the global fit at all temperatures in two situations: (1) without any constraints; and (2) imposing Arrhenius linearity for the transition state and van't Hoff<sup>7</sup> linearity for the equilibrium between major and minor conformational states<sup>24</sup>. Since both global and individual fittings,  $\chi^2 < \text{DF}$  (global) and  $\Sigma\chi^2 < \Sigma\text{DF}$  (individual), the global fitting is consistent with the two-state models. A more detailed description of the analysis is in the supplementary information.

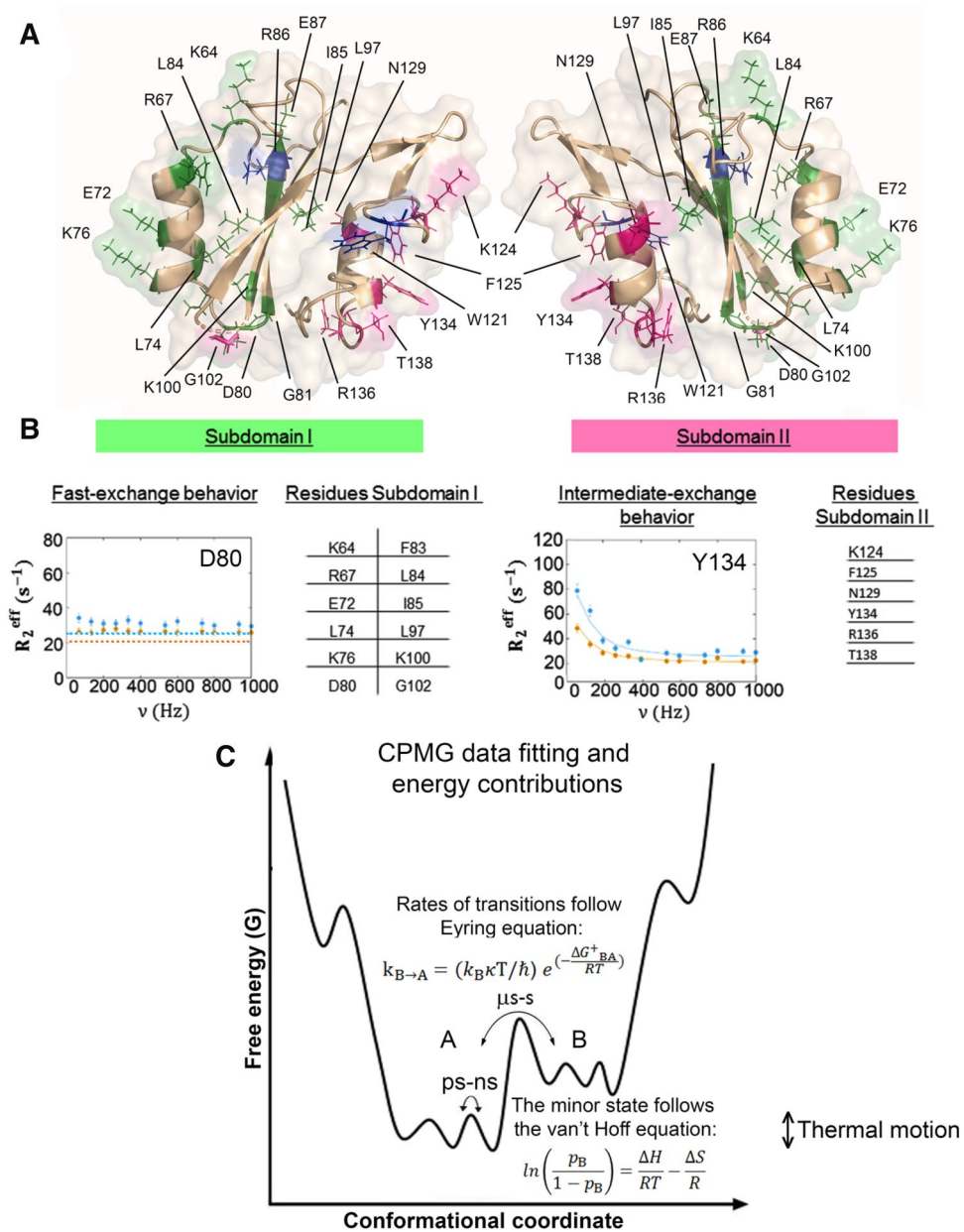
The residues K124, F125, Y134, R136, and T138 presented a good  $^{15}\text{N}$  CPMG-RD profile. Because of the low number of residues in intermediate conformational exchange at subdomain II (Figures S3, S4, S5, and S6), there was a poor convergence of the population of the minor state ( $p_B$ ) for the global fitting at each temperature. Conversely, the global fitting at all temperatures and especially the constrained global fitting enabled a good convergence of  $p_B$ . We obtained a good convergence for exchange constant ( $k_{\text{ex}}$ ) in all situations. Figure 2C summarizes how to obtain the free energy landscape of a protein from the parameters obtained for the RD profiles. We used van't Hoff equation to fit  $p_B$  as a function of temperature and obtain the thermodynamics of equilibrium, and Eyring equation to fit  $k_{\text{AB}}/k_{\text{BA}}$  as a function of temperature to obtain the thermodynamics of the transition. These data enabled to obtain the thermodynamic parameters ( $\Delta G$ ,  $\Delta H$ , and  $\Delta S$ ) for the equilibrium and transition state (Fig. 3).

Figure 3A shows the  $p_B$  and Fig. 3B the  $k_{\text{ex}}$  for global fitting at all temperatures without and with constraints. The increase in temperature led to a small decrease in  $p_B$  and an increase in  $k_{\text{ex}}$ . This behavior reflected in the thermodynamic profile, which is observed in Fig. 3C.

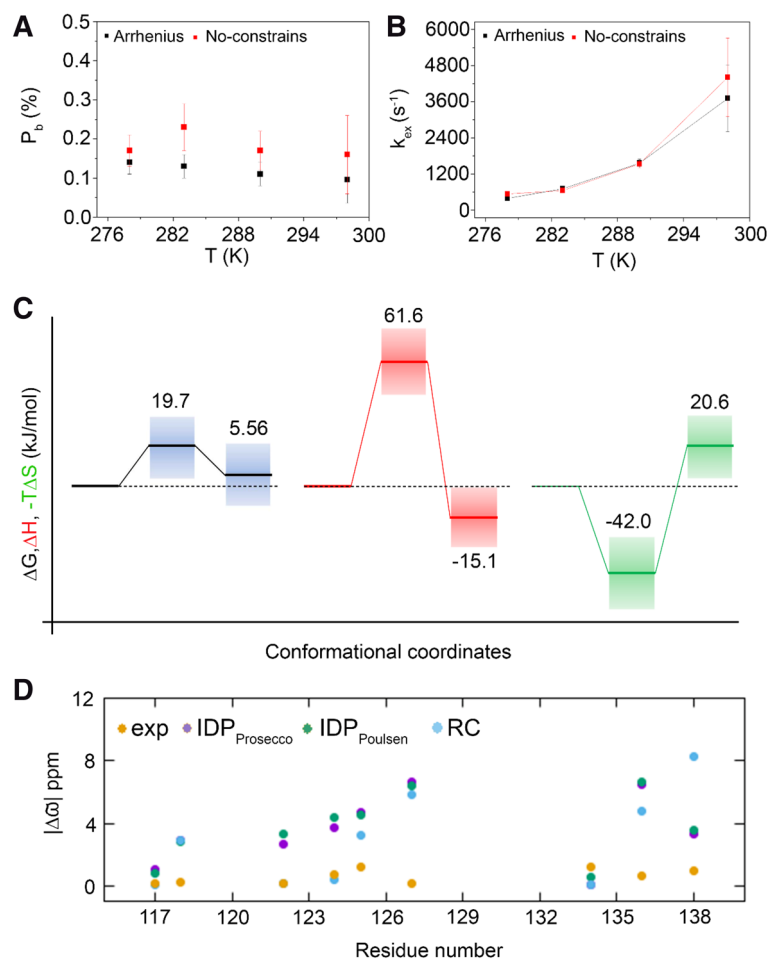
The minor state (Fig. 3C) is enthalpically favorable and entropically unfavorable at 298 K. This is typical behavior of a conformational fluctuation involving the exposure of hydrophobic residues to the solvent<sup>24</sup>. Knowing that the conformational exchange involving loop  $\beta_8/\alpha_3$  and  $\alpha_3$  may be associated with the observation, in several crystal structures, of a domain swap involving  $\alpha_3$ , which opens to interact with the adjacent subunit to form the swapped dimer<sup>13–16</sup>. The question that remains is whether the minor state is dimeric or monomeric. To answer this question, we acquired the  $^{15}\text{N}$ -CPMG-RD profiles at 278 K in a lower concentration (120  $\mu\text{M}$ ,  $\sim 2.5 \times$  diluted). These profiles did not have significant changes (Figure S7), meaning that there was not the concentration-dependence expected for an oligomer. We concluded that the observed conformational exchange does not reflect a dimerization of the minor state nor domain swapping. Instead, the minor state is a result of conformational fluctuations of  $\alpha_3$  in the monomeric state and related to the pY + 2 recognition.

To get insights on the structural changes at the minor state, we analyzed the differences in the chemical shifts ( $\Delta\omega$ ) between the major and minor conformational states and compared with predicted values for the random coil<sup>25</sup> and the intrinsically disordered domain (Prosecco<sup>26</sup> and Poulsen<sup>27</sup>, Fig. 3D). The observed  $\Delta\omega$ 's are small when compared to the expected values for the unfolding/disordering of  $\alpha_3$ . The minor state is probably a result of the repositioning of  $\alpha_3$ , exposing hydrophobic residues to the solvent.

To correlate the dynamics with the phosphotyrosine sensing mechanism, we measured the effect of pY-pep binding on the Grb2-SH2 structure and dynamics. We analyzed the chemical shift perturbation (CSP) upon pY-pep binding (Fig. 4A). There were significant perturbations throughout the protein. When we looked at the most prominent CSPs at Grb2-SH2 structure (Fig. 4B), we verified the correlation with the peptide binding. Based on all the information available on PDB for complexes between Grb2-SH2 and different phosphopeptides, we calculated the atomistic probability density for the phosphopeptides bound to Grb2-SH2. The green mesh at Fig. 4B shows that the highest probability density is at positions pY, pY + 1, and pY + 2, in agreement with the



**Figure 2.** Summary of the regions in conformational exchange of Grb2-SH2 domain obtained from  $^{15}\text{N}$  CPMG-RD experiments at 5, 10, 17 and 25 °C. **(A)** There are two dynamically independent subdomains. The residues comprising the subdomain I (green) are in fast conformational exchange. The residues comprising the subdomain II (pink) are in intermediate conformational exchange. For reference, the R86 and W121 are colored in blue, evidencing the two subsites for molecular recognition of the pY-pep. **(B)** Typical RD profile of the residues in fast exchange regime observed for subdomain I, represented by D80. The table shows all the residues in fast exchange. Typical RD profile of the residues in intermediate exchange regime observed for subdomain II, represented by Y134. The table shows all the residues in intermediate exchange at subdomain II. The dotted lines represent the exchange-free  $R_2^{\text{eff}}$  for each field **(C)** Free energy landscape of a protein undergoing conformational exchange between two states. A is the major conformational state while B is the minor state. For the fitting of the RD profiles of the residues in intermediate exchange using Bloch-McConnell equation, we obtained the population of the minor state ( $p_B$ ), the exchange rate ( $k_{\text{ex}} = k_{AB} + k_{BA}$ ) and the chemical shift difference between major and minor state ( $\Delta\omega$ ). Using van't Hoff equation for fitting  $p_B$  as a function of temperature we obtained the thermodynamic parameter at equilibrium ( $\Delta G$ ,  $\Delta H$ ,  $\Delta S$ , Fig. 3). Using Eyring equation for fitting  $k_{AB}$  or  $k_{BA}$  as a function of temperature we obtained the thermodynamic parameter at the transition state ( $\Delta G^\ddagger$ ,  $\Delta H^\ddagger$ ,  $\Delta S^\ddagger$ , Fig. 3).



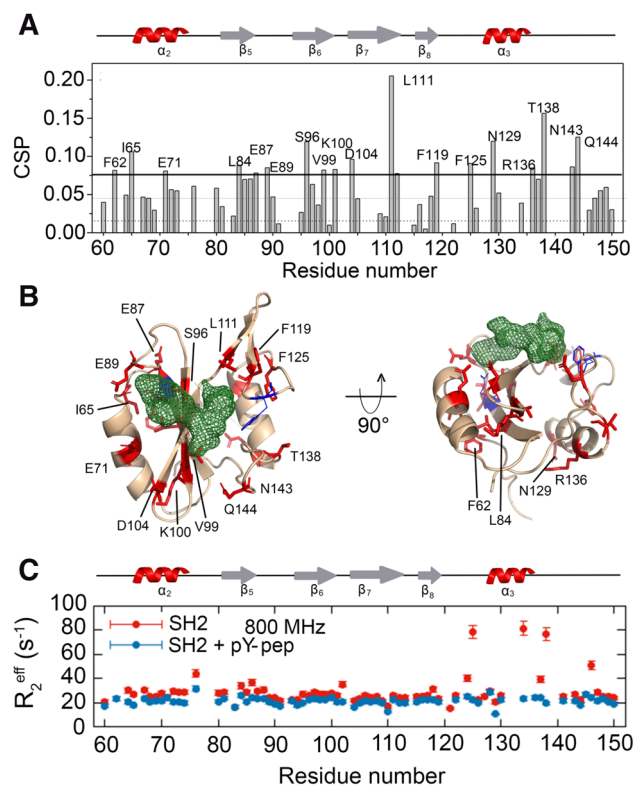
**Figure 3.** Energy landscape of subdomain II of Grb2-SH2 domain. Thermodynamic parameters of equilibrium and transition state from the global fitting using the Bloch-McConnell equation of  $^{15}\text{N}$  CPMG-RD experiments at several temperatures. **(A)** The population of the minor state ( $p_B$ ) and **(B)** the exchange constant ( $k_{\text{ex}}$ ) as function of the temperature. The black and red squares denote the global fitting at all temperatures (no-constrains) and imposing the linearity of Arrhenius and van't Hoff equations, respectively. **(C)** Conformer A is the major state while B is the first thermally accessible conformational state (minor state).  $\Delta G$  is in black,  $\Delta H$  is in red, and  $-\Delta S$  is in green. Note that the minor conformational state is enthalpically favorable. The thermodynamic parameters obtained from  $p_B$  and  $k_{\text{ex}}$  at individual temperatures can be found in Table S1.

CSP. Many of the perturbed residues are directly facing the atomic probability density of the phosphopeptides (green mesh), at subdomain I, E89 in the loop  $\beta 5/\beta 6$ , the same regions observed in the calculated structure as open (Figure S2D), R86 that directly binds to pY, and S96. At subdomain II, L111 of the  $\beta$ -hairpin  $\beta 7/\beta 8$  (the highest CSP) and Q144 also bind directly to the phosphopeptide. In contrast, the probability density map of the backbone of the bound structures in the PDB is closed (Figure S2D). Residue F62, N129, and R136 are on the opposite face. It is noteworthy that in the full-length Grb2 the F62 and R136 are facing the dimerization interface. It was previously reported<sup>18</sup> that the binding of pY-pep leads to the dissociation of the dimeric Grb2. The binding of pY-pep also affected residues at  $\alpha 3$  (N129 and T138). L111 and T138 were the ones with the highest CSP.

Next, we measured the  $^{15}\text{N}$  CPMG-RD for a saturated concentration of pY-pep (6 mM) (Fig. 4C, S8, and S9). At this condition, most of the conformational exchange observed for subdomains I and II is quenched. This is typical of a conformational selection mechanism of binding.

We also measured the  $^{15}\text{N}$  CPMG-RD for a semi-saturated concentration of pY-pep (2.1 mM) (Fig. 5). We took advantage of the chosen experimental conditions at 20 mM inorganic phosphate and 200 mM NaCl. The inorganic phosphate competes with the pY-pep for the binding to R86 while the NaCl shields electrostatic interactions. At this condition the pY-pep binds with millimolar affinity, enabling the observation by NMR of “binding intermediates” in the process of molecular recognition. We observed, in the presence of a semi-saturated concentration of pY-pep, that residues acquired relaxation dispersion (Fig. 5A). Interestingly, these residues are at the same surface of the recognition site as shown in Fig. 5C. It is noteworthy that the residues that took part in exchange processes are uniquely at the pY-pep molecular recognition face. In the context of the full-length Grb2, this surface is exposed to the solvent, not involved in inter-domain or inter-subunit contacts.



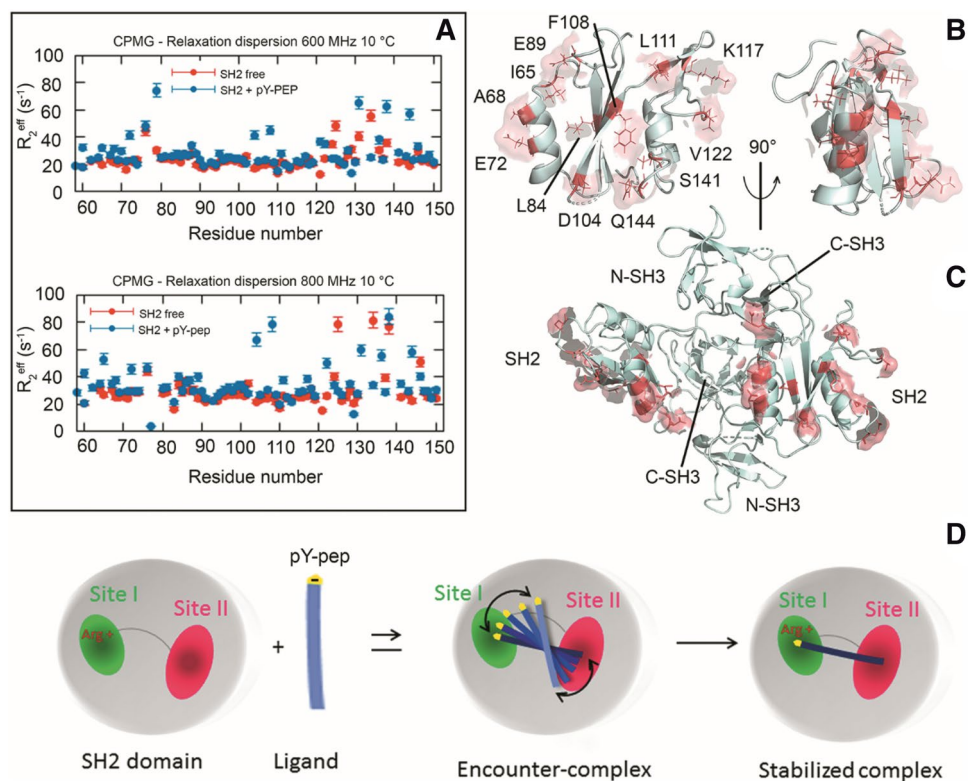


**Figure 4.** Binding of pY-pep to Grb2-SH2. **(A)** Chemical shift perturbation (CSP) as a function of residue number promoted by the presence of 2.1 mM pY-pep to Grb2-SH2 at 298 K. The labeled residues are the ones with CSP larger than the average plus two standard deviations. **(B)** Cartoon representation of the Grb2-SH2 structure (PDB\_id 1BMB) highlighting in red the residues labeled in A. The green mesh is the atomic probability density map of the ensemble of all Grb2-SH2 structures complexed with phosphopeptides available at the PDB website. It presents the highest probability density of finding each atom of the phosphopeptides, which contains the pY, pY + 1, and pY + 2. **(C)**  $R_{2}^{\text{eff}}$  at  $66.7 \text{ s}^{-1}$  as a function of the residue number at 18.8 T (800 MHz) for the Grb2-SH2 with (blue) and without (red) pY-pep at a saturated concentration (6 mM) at 283 K. For the complete results see Figures S4, S8, and S9.

We proposed a mechanism for pY-pep molecular recognition that involves the formation of an encounter complex as “binding intermediates” in the process (Fig. 5D). The recognition surface of the encounter complex is wider than that of the stabilized complex. Fuzzy interactions of pY-pep led to an uncertainty of the position of the pY-pep at the encounter complex that generates a chemical exchange that could be observed by NMR. In this mechanism, the residues that gained exchange are those involved in the tethering of pY-pep (or pY-pep’s) in the encounter complex. It is interesting to observe that some of these residues are negatively charged (E72, E89, D10), which were surrounding the positively charged R86 at site I, the responsible for the direct binding to the phosphate of phosphotyrosine. Hydrophobic residues also were found in the encounter complex (I65, A68, L84, F108, L111, and V122). It is noteworthy that L111 showed the highest CSP, no conformational exchange in the free state and gained relaxation dispersion in the presence of semi-saturated concentrations of pY-pep (Figure S10 and S11). Ross & Subramanian<sup>28</sup> described the importance of hydrophobic surfaces in the formation of encounter complexes, with the first step of mutual penetration of hydration layers causing solvent disorder followed by further short-range interactions.

To better understand the dynamics of Grb2-SH2 domain and the mobility of pY-pep relative to the domain, we run 1  $\mu\text{s}$  molecular dynamics simulations (MD) for the free domain and complexed with the pY-pep. By analyzing the root mean square fluctuation (RMSE, Fig. 6A), we observed four main conformational flexible regions, the N- and C-terminal residues, the loop  $\beta 5/\beta 6$ , and the hairpin  $\beta 7/\beta 8$ , both in the presence and absence of pY-pep (Fig. 6A). The conformational flexibility in these regions could not be observed in the nuclear spin relaxation parameters (Figure S2), which reflect the thermal flexibility in a timescale smaller than the overall rotational correlation time ( $\tau_c$ , sub- $\tau_c$  dynamics). Based on the  $R_2/R_1$  ratio, we estimated  $\tau_c$  to be  $\sim 6.5 \text{ ns}$ . The motions observed in the MD simulations are probably related to supra- $\tau_c$  dynamics (tenths of ns to  $\mu\text{s}$ ), which could not be observed either by  $^{15}\text{N}$  nuclear spin relaxation parameters or  $^{15}\text{N}$  CPMG-RD. Supra- $\tau_c$  dynamics are described to regulate the “on” rates in the molecular recognition process<sup>29</sup>.

In the complex, there was a slight increase in supra- $\tau_c$  flexibility for loop  $\beta 5/\beta 6$  and hairpin  $\beta 7/\beta 8$ , and the C-terminal loop between residues 140 and 150. There was also a decrease in supra- $\tau_c$  flexibility for residues in  $\alpha 2$ , loop  $\alpha 2/\beta 5$ , and loop  $\beta 6/\beta 7$ . This observation corroborates with the CSP results observed upon the addition

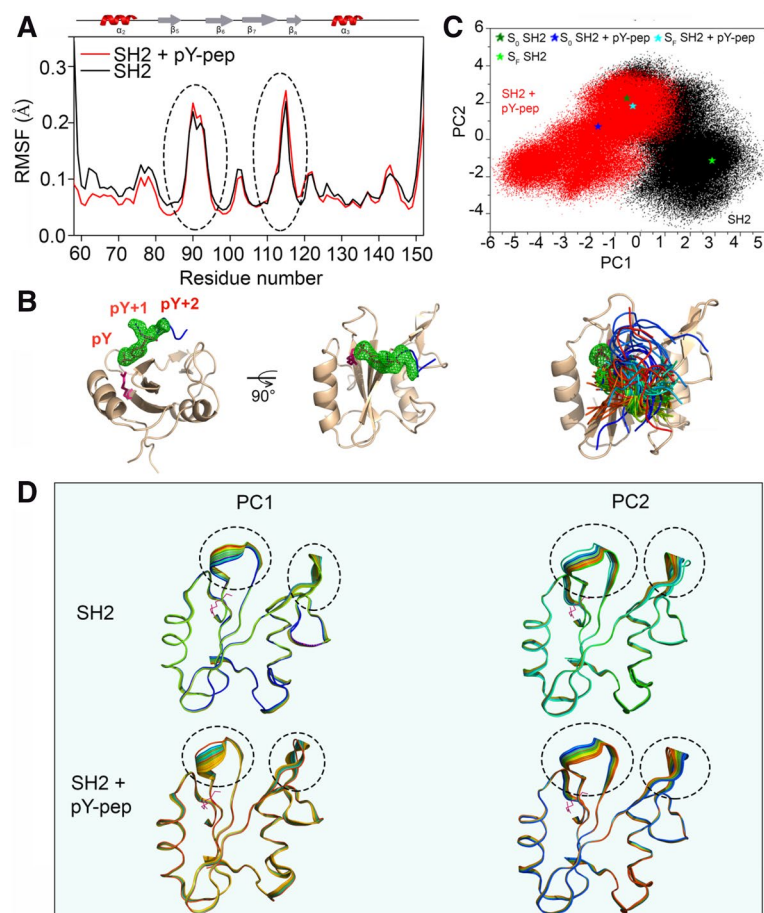


**Figure 5.** Effect of the binding of pY-pep to Grb2-SH2 domain at a semi-saturated concentration. **(A)**  $R_{2eff}$  at  $66.7 \text{ s}^{-1}$  as a function of the residue number in two fields, 14.09 (600 MHz) and 18.8 (800 MHz) T, for the Grb2-SH2 with (blue) and without (red) pY-pep at a semi-saturated concentration (2.1 mM) at 283 K. For the complete  $^{15}\text{N}$  CPMG-RD profiles see Figures S10 and S11. **(B)** Cartoon representation of the Grb2-SH2 structure showing in pink the residues that acquired relaxation dispersion in the presence of a semi-saturated concentration of pY-pep. Note that the residues which gained exchange are exclusively at the pY-pep molecular recognition face. **(C)** Cartoon representation of the full-length Grb2, showing in pink the residues that gained inter-exchange. It is worth to note that they are solvent-exposed on the surface, and not involved in inter-domain or inter-subunit contacts. **(D)** Schematic representation of the proposed molecular recognition mechanism for the phosphopeptides. The pY-pep becomes tethered to the recognition surface, forming an encounter complex and finally finds the recognition site stabilizing the complex. The uncertainty of the position of pY-pep at the encounter complex induces the observed chemical exchange.

of pY-pep (Fig. 4A,B), for which these regions presented significant CSP values. Highlighting that L111 in the hairpin  $\beta 7/\beta 8$ , and N143 and Q144 in the C-terminal loop are very important residues for the binding and L111 is part of the encounter complex (Figure S10 and S11), demonstrating the importance of these regions to the recognition of pY-pep and the complex formation. This behavior of increased flexibility of Grb2-SH2 in the presence of pY-pep was corroborated by the analysis of the root mean square deviation (RMSD) relative to the starting structures (Figure S12A). In the presence of pY-pep there is a wider fluctuation of RMSD around its average value when compared with the free domain. We also observed a stability point for the values of RMSD after  $\sim 500 \text{ ns}$  for the free Grb2-SH2 domain, while for the pY-pep bound domain this stability point was observed since the beginning of the simulation.

The MD simulations helped describe the position and mobility of pY-pep relative to the domain. As expected, positions pY, pY + 1, and pY + 2 were well determined, presenting a high atomistic probability density at these positions (Fig. 6B). In contrast, the positions pY + 3, pY + 4, and pY + 5 showed high mobility and low atomistic probability density. This behavior agrees with the expected specificity mapping for Grb2-SH2<sup>30</sup>, with no participation of pY + 3. It is remarkable that even though R86(+)/pY(-) interaction is present along with all the MD simulation, the backbone position varies significantly for all peptide, even at the positions pY and pY + 1, showing the presence of multiple conformations of the peptide at the recognition face (Fig. 6B).

To describe the supra- $\tau_c$  motions, we analyzed the trajectories of the MD simulations using principal component analysis (PCA). The two main principal-components (PC1 and PC2) showed an oscillation among different conformational sub-states along the simulation time (Figure S12B). For the free domain, we observed a convergence toward a stabilized conformation both in PC1 and PC2 after  $\sim 400 \text{ ns}$  of simulation (2 to 3 for PC1 and 0 to -2 for PC2 in Figure S12B). We must be careful before attributing a convergent structure for the free domain, because, with longer simulation time it could be reverted, representing a lower-frequency oscillation. For the complex, we observed no stability point, both in PC1 and PC2, without a convergence, agreeing with the gained supra- $\tau_c$  flexibility that denotes a higher frequency oscillation of possible conformational sub-states of



**Figure 6.** Molecular dynamic simulation of free and pY-pep bound Grb2-SH2 domain. **(A)** RMSF of Grb2-SH2 in the presence (red) and absence (black) of pY-pep. **(B)** Cartoon representation of the structure of Grb2-SH2 (PDB\_id 1BMB). The green mesh is the atomistic probability density map of the ensemble of pY-pep structures along with the MD simulation. It represents the probability density of finding each atom of the phosphopeptide, which contains the pY, pY + 1, and pY + 2. Note that there is no probability density shown for the residues pY + 3, pY + 4, and pY + 5. The structure in the bottom shows the superposition of the pY-pep structure with the interaction surface of Grb2-SH2 along with the MD simulation (at each 10 ns). Note the high variability in position relative to the domain for pY + 3, pY + 4, and pY + 5. **(C)** PC1 x PC2 observed in the PCA for Grb2-SH2 in the presence (red) and absence of pY-pep (black). Note that both in the presence or in the absence of pY-pep, the MD simulation shows transitions between conformational sub-states. The colored stars show the starting points of the MD simulations (S0 dark green for the free domain and dark blue for the complexed domain) and the ending point of the MD-simulations (SE, dark green for the free domain and dark blue for the complexed domain). **(D)** Flexibility observed in the PCA for the first principal eigenvector (PC1) and second principal eigenvector (PC2) for Grb2-SH2 in the presence and absence of pY-pep. The dotted circles in A and C highlight loop  $\beta 5/\beta 6$  and the hairpin  $\beta 7/\beta 8$ , which are the main flexible regions in the free or complexed Grb2-SH2 domain along with the MD simulation.

the protein. The plot PC2 versus PC1 (Fig. 6C) showed the interconversion among different conformational sub-states. For both free and complexed Grb2-SH2 domain, we observed different interconverting conformational sub-states. Interestingly, the conformational sub-state with PC1 in the range between  $-1.5$  and  $1$  and with PC2 between  $0$  and  $3$  are common between the free and complexed domain. It is important to note that this sub-state is the starting point for the MD simulation of Grb2-SH2 and the ending point for the complexed protein. It is worth mentioning that the starting point was the crystal structure from PDB\_id 1BMB<sup>31</sup> stripped of the complexed peptide. The PCA was done with the two MD simulations concatenated in the same trajectory file, which enabled the direct comparison of the eigenvectors PC1 and PC2 for both free and complexed MD simulations.

The structural extent of supra- $\tau_c$  conformational dynamics is depicted in Fig. 6D. Note that loop  $\beta 5/\beta 6$  and the hairpin  $\beta 7/\beta 8$  presented the main structural changes responsible for PC1 and PC2. Those are the regions with the highest RMSF and higher flexibility in the presence of pY-pep. The extent of structural change in the presence of pY-pep was more pronounced than for the free domain. This could also be observed for the C-terminal loop, containing the residues N143 and Q144. Remarkably, these are the same regions for which we observed the open-to-close conformational transition in the calculated structure (Figure S2C).



## Discussion

Here we present the dynamical behavior of Grb2-SH2 domain and its relationship with the molecular recognition of phosphopeptides. We showed that this domain behaves as two dynamically independent subdomains, suggesting a correlation with previously described properties of Grb2-SH2. There are many reports on the dynamics of SH2 domains. However, a full characterization of an SH2 domain in the milli-to-microsecond timescale was not available. Many of the contributions correlate backbone dynamics in the pico-to-nanosecond timescale with the binding of phosphopeptide recognition<sup>32–37</sup>. In the present manuscript we described the dynamical behavior in multiple timescales, but with an emphasis on the conformational exchange (milli-to-microseconds), which is the timescale related to molecular recognition ( $k_{\text{off}}$ ).

The dynamical data showed that subdomain I is involved in microsecond dynamics while subdomain II in millisecond dynamics, suggesting independence in the association of the pY recognition (subdomain I) and pY + 2 recognition (subdomain II). We proposed that the dynamics at subdomain II involves a minor conformational change that exposes hydrophobic residues to the solvent. We present here the first dynamical characterization of Grb2-SH2, revealing a torsion/reorientation mechanism of  $\alpha 3$ , an important conformational state for the binding specificity of pY + 2 and allosteric events.

Our data showed that pY-pep recognition occurred by conformational selection, in which a conformational state is stabilized. The phosphopeptide association is specific for the exposed-face of the full-length Grb2, the recognition-face (Fig. 5C). Before the formation of a stabilized-complex, an encounter complex was formed involving negatively charged and hydrophobic residues, creating an electrostatic potential that drives the association of the phosphotyrosine (pY) to R86 (Figure S14A). This fact suggests that the association can be driven by Coulombic and hydrophobic interactions. The observation of an encounter complex led to the conclusion that the recognition face is wide, with fuzzy interactions with the phosphopeptide. There are still many unanswered questions regarding the phosphopeptide recognition by SH2 domain. One of the most intriguing is regarding the recognition of remote phosphotyrosines<sup>38</sup>. This is biologically relevant for proteins with multiple phosphorylations. This manuscript did not fully explain this effect, but the wide recognition interface and the fuzzy interaction with the phosphopeptide point toward a possible mechanism of how remote phosphotyrosines are sensed by SH2 domain.

The high-resolution structure of the free-state (PBD\_id 6VK2) enabled the characterization of an open conformation, in which  $\beta 5/\beta 6$  and C-terminal loops undergo an open-to-close conformational transition upon binding of phosphopeptides. These loops, along with the hairpin  $\beta 7/\beta 8$ , are far apart, demonstrating that the recognition surface is wide. The structural changes are in agreement with CSP and dynamic description of the free and bound-states, and the presence of “binding intermediates” in the encounter complex.

The PCA of the MD simulations showed that there are different conformational sub-states for the free and complexed domain, where the pY-pep can assume multiple conformations since the position pY presented the highest atomistic probability density, followed by pY + 1 and pY + 2, and the conformations at positions pY + 3 to pY + 5 were very diverse. Interestingly, the presence of pY-pep increased the extent of conformational transitions among the different sub-states. In agreement with our experimental results, Lindfors et al. (2012)<sup>39</sup> used spin-labeled phosphopeptides to probe the interaction with SH2. The peptide assumes multiple orientations at the interacting surface. The electrostatic interactions involving charged patches of the protein resulted in an ensemble of rapidly exchanging orientations, suggesting a dynamic encounter state. This behavior shows that high-affinity binding can be dynamic and that residues outside the central SH2 binding site are also important for the Src-FAK interaction<sup>39</sup>.

The formation of the dynamic association and encounter complexes is a millisecond timescale event and involved the quenching of the dynamics of subdomain I and subdomain II. The MD simulations pinpointed important motions probably in a faster supra- $\tau_c$  dynamics, which involves mainly three regions of the protein: loop  $\beta 5/\beta 6$  at subdomain I, and hairpin  $\beta 7/\beta 8$  and the C-terminal loop, both at subdomain II. Multiple timescale dynamics are pivotal for recognition events. It is noteworthy that the slow  $\mu\text{s}$ -ms dynamics is related to the “off” rates, while the supra- $\tau_c$  dynamics (tenths of ns to  $\mu\text{s}$ ) to the “on” rates in a molecular recognition process<sup>29</sup>. The wide and dynamic character of the recognition face is also evidenced by the MD simulation. There is a good correspondence between the contact map of Grb2-SH2 (Figure S13) and pY-pep and the CSP. All the regions mapped by CSP and CPMG relaxation dispersion were observed in the MD-simulation.

Our data also showed that the association of pY-pep to the recognition-face led to chemical shift changes in the opposite face, such as F62 and R136 that are involved in dimeric contacts in the full-length Grb2. The F62 is at subdomain I and R136 at subdomain II. This observation is suggestive of an allosteric event triggered by pY-pep association, which is in agreement with the fact that the association of pY-pep leads to the monomerization of full-length Grb2<sup>18</sup>. The monomer-dimer equilibrium of Grb2 is biologically relevant and pivotal for the regulation of signal transduction pathways<sup>3–6</sup>. The Grb2 dimer interacts with the fibroblast growth factor receptor 2 (FGFR2) which in turn activates the Ras/MAPK pathway, while its monomeric form is inhibitory for the signaling. Ras/MAPK is important and widely studied via related to a variety of cancer and development defects<sup>2,3</sup>. Further studies are necessary for a better understanding of the mechanism of monomerization. Plasticity is a major feature of an adaptor protein such as Grb2. The present manuscript detailed many of the residues involved in the dynamic of the free and bound states of the Grb2-SH2 domain, which is the basis for future studies of the plasticity of Grb2.

The dynamics of subdomain II, showed the exposure of hydrophobic surfaces with the repositioning/reorientation of  $\alpha 3$ . The possibility of this dynamics to contribute to the domain swap was considered, once the  $\alpha 3$  is the helix involved in the swapping mechanism<sup>13–16</sup>. However, our results showed that the observed dynamics is not involved in the swapping mechanism. Domain swap for Grb2-SH2 domain occurs in a different timescale (minutes/hours) from the observed  $\alpha 3$  dynamics (ms) and the swapped dimer can be separated in a gel-filtration

column. Moreover, we showed that Grb2-SH2 is monomeric in the major and minor states. The biological role of the domain swap in the Grb2-SH2 is not fully understood and further studies are necessary. Grb2-SH2 domain-swap would add another degree of freedom for Grb2 plasticity and could be important, but it was never observed in the full-length Grb2.

In conclusion, we propose a mechanism of phosphotyrosine recognition by Grb2-SH2 domain, which involves a wide recognition face and a dynamic association of the phosphopeptide. We verified the formation of an encounter complex as “binding intermediates” involving the exposed face of the SH2 domain of the full-length Grb2. This is biologically relevant for an adaptor protein, such as Grb2, since it potentially increases the scope of interaction. In agreement with this conclusion, it has been reported that Grb2 is able to sense remote phosphotyrosines. Huang et al. (2017) showed that the binding affinity for the linker of activation of T-cells (LAT) and Grb2 depends on the phosphorylation of remote tyrosine sites<sup>38</sup>. This capacity could be conferred by the dimeric state of full-length Grb2, but also by the wide and dynamic character of SH2 domain to recognize phosphopeptides.

## Materials and methods

**Protein expression and purification.** The SH2 domain – 6×histidine tagged (Grb2-SH2)—was expressed and purified as previously described by Sanches et al. (2019)<sup>19</sup>. Chemical shifts are deposited in the Biomagnetic Resonance Bank ([www.bmrb.wisc.edu](http://www.bmrb.wisc.edu)) under accession number 27781.

**NMR spectroscopy.** All NMR samples of Grb2-SH2 was in 20 mM sodium phosphate buffer, 200 mM NaCl, pH 7.0, 10% D<sub>2</sub>O. NMR spectra were acquired at 25 °C on Bruker Avance III 600 MHz, Avance III 800 MHz and Avance IIIHD 900 MHz, equipped with <sup>15</sup>N/<sup>13</sup>C/<sup>1</sup>H triple-resonance probes (Bruker TXI). NMR spectra were processed with NMRPipe<sup>40</sup> and analyzed using CCPNmr Analysis software<sup>41</sup>. The detailed experimental setup for all NMR experiments are in Table S3.

**Structure calculation.** Distance restraints were derived from <sup>15</sup>N-NOESY-HSQC and aliphatic and aromatic <sup>13</sup>C-NOESY-HSQC collected at an AVANCE III HD 900 MHz, for aromatics and aliphatics. The 3D-NOESY-HSQC were obtained using non-uniform sampling (NUS, 50%) with multidimensional Poisson Gap scheduling. NMRPipe and iterative soft threshold (hms1ST) fast reconstruction of NMR data were used for processing<sup>42</sup> (Table S3). We run TALOS-N<sup>43</sup> for backbone chemical shift-based dihedral angle prediction. The predicted backbone dihedral angles  $\phi$  and  $\psi$  of the residues involved in secondary structure were used as a restraint for structural calculations.

Structure calculation of the Grb2-SH2 was performed iteratively using ARIA 2.1 program, version 2.3<sup>44,45</sup> combined with CNS version 1.2<sup>46</sup>, using <sup>15</sup>N-NOESY-HSQC and aliphatic and aromatic <sup>13</sup>C-NOESY-HSQC datasets as the source of distance restraints. The CCPN/Aria interface was used<sup>41</sup>.

The converged structure using Aria/CNS generated an ensemble of well-converged structures (Figure S1A) that was generated from 697 unambiguous and 277 ambiguous distance restraints. Next, the structural calculation proceeded using Chemical-Shift Rosetta (CS-Rosetta, ROSETTACOMMON Version 3.7 run at NMRBOX<sup>47</sup>). All the intra-residue NOEs were removed in CS-Rosetta calculation. The sequential, medium, and long-range distance restraints generated using Aria/CNS were then converted to the Rosetta format and used as distance restraints. <sup>13</sup>C, <sup>15</sup>N, and <sup>1</sup>H NMR chemical shifts were used as inputs (BMRB 27781<sup>19</sup> for fragment picking, along with 479 NOE-based distance restraints. 66,317 structures generated by the standard CS-ROSETTA protocol. The 20 lowest energy structures were deposited in the protein data bank (PDB id 6VK2).

The structural ensemble was visualized and analyzed with PyMOL. Quality validation was Protein Structure Validation Software suite (PSVS) (<https://montelionelab.chem.rpi.edu/PSVS/>) and Molprobity (<https://molprobity.biochem.duke.edu/>)<sup>49</sup>.

**Nuclear spin relaxation parameters.** <sup>15</sup>N backbone amide relaxation parameters (<sup>15</sup>N R<sub>1</sub>, <sup>15</sup>N R<sub>2</sub> and <sup>1</sup>H-<sup>15</sup>N heteronuclear Nuclear Overhauser effect—NOE) were measured for a <sup>15</sup>N labeled Grb2-SH2 (300 μM, in 20 mM sodium phosphate buffer, 200 mM NaCl, 10% D<sub>2</sub>O, pH 7.0) using Avance III HD 800 (18.8 T, operating at 800.4 MHz) at 25 °C. R<sub>1</sub> was measured with delays ranging from 0.05 to 1 s. R<sub>2</sub> was measured with delays varying from 17 to 170 ms. The experimental error was evaluated from the signal-to-noise ratio of the spectra. The <sup>1</sup>H-<sup>15</sup>N NOE was acquired with or without proton saturation for 6.0 s. Details of the NMR experiments are in Table S3. The R<sub>1</sub> and R<sub>2</sub> values were obtained using the relaxation module of CcpNmr Analysis<sup>41</sup>. The <sup>1</sup>H-<sup>15</sup>N heteronuclear NOE values were determined using the intensity saturation spectra/intensity without the saturation spectra ratio.

**<sup>15</sup>N Relaxation dispersion measurements.** <sup>15</sup>N Carr–Purcell–Meiboom–Gill relaxation dispersion (CPMG-RD) profiles for a <sup>15</sup>N labeled Grb2-SH2 (300 μM, in 20 mM sodium phosphate buffer, 200 mM NaCl, 10% D<sub>2</sub>O, pH 7.0) were recorded in Bruker Avance III 600 (14.09 T, operating at 600.03 MHz) and Avance III 800 (18.8 T, operating at 800.4 MHz) at four temperatures (278, 283, 290 and 298 K) using the constant relaxation time of T<sub>relax</sub> = 30 ms (<sup>15</sup>N CPMG relaxation compensation<sup>49</sup>). R<sub>2</sub><sup>eff</sup>(ν<sub>CPMG</sub>) were calculated from peak intensities (I) in a series of two-dimensional (2D) <sup>1</sup>H-<sup>15</sup>N correlation spectra recorded in an interleaved way at different CPMG frequencies ν<sub>CPMG</sub>, (ranging from 66.7 to 1,000 s<sup>-1</sup>) using the following equation: R<sub>2</sub><sup>eff</sup>(ν<sub>CPMG</sub>) = -1/T<sub>relax</sub> ln(I/I<sub>0</sub>), where I is the signal intensity in the spectra collected at T<sub>relax</sub> = 30 ms and I<sub>0</sub> is the signal intensity in the reference spectrum recorded at T<sub>relax</sub> = 0. The experimental error in R<sub>2</sub><sup>eff</sup> rates were estimated signal to noise ratio for each resonance ΔR<sub>2</sub><sup>eff</sup>(ν<sub>CPMG</sub>) = 1/T<sub>relax</sub> (1/(SIGNAL/NOISE)).

**Amide chemical shift temperature coefficient.** Amide  $^1\text{H}$ N chemical shift temperature coefficient of Grb2-SH2 (300  $\mu\text{M}$ , in 20 mM sodium phosphate buffer, 200 mM NaCl, 10%  $\text{D}_2\text{O}$ , pH 7.0) was measured in the presence and absence of the pY-pep. A series of two-dimensional  $^{15}\text{N}/^1\text{H}$  HSQC spectra at 278, 283, 290 and 298 K was acquired (Bruker Avance III HD 800 18.8 T, operating at 800.4 MHz). The water chemical shift was referenced using 3-(trimethylsilyl)propane-1-sulfonic acid (DSS). Each spectrum was referenced to the water signal at each temperature. They were processed using NMRPipe and analyzed using CcpNmr Analysis. The amide chemical shift values ( $\delta_{\text{HN}}$ ) of all residues at different temperatures were plotted as function of the temperature, and the slope ( $d\delta_{\text{HN}}/dT$ ) of every curve was plotted for each Grb2-SH2 residue.

**Phosphopeptide binding.** The binding of the phosphopeptide EpYINSQV (sequence from the Epidermal Growth Factor receptor, pY-pep<sup>18</sup> to Grb2-SH2 (300  $\mu\text{M}$ , in 20 mM sodium phosphate buffer, 200 mM NaCl, 10%  $\text{D}_2\text{O}$ , pH 7.0, 298 K) was evaluated by chemical shift perturbation (CSP) using the  $^{15}\text{N}$ -HSQC spectra with and without pY-pep. CSP values were calculated using the equation  $\text{CSP} = \sqrt{0.5(\Delta\delta_H)^2 + 0.5(\Delta\delta_N/10)^2}$ .

**Construction of the probability map.** We created a map of the weighted atomic density using the function VOLMAP in Visual Molecular Dynamics (VMD)<sup>50</sup> software. The atomistic probability map was calculated (i) over an ensemble containing all Grb2-SH2 complexed with phosphopeptides available at the protein data bank (RSCB PDB), except the domain-swapped Grb2-SH2 dimers and (ii) an ensemble of 100 structures collected at each 10.000 frames of the molecular dynamics simulation of Grb2-SH2 and of Grb2-SH2 complexed with the docked pY-pep.

**Computational simulations of the free and pY-pep bound Grb2-SH2 domain.** The structure of the Grb2-SH2 domain used for computer simulations was downloaded from PDB website under access code 1BMB<sup>31</sup>. The phosphopeptide and water molecules in the crystal structure were removed. The domain and pY-pep were prepared using AutoDockTools program<sup>51</sup> for molecular docking simulations, merging non-polar hydrogen atoms and adding atom types. The pY-pep rigid root was generated automatically, setting all possible rotatable bonds defined as active by torsions. The protonation states of ionizable residues of the protein were defined according to PROPKA software<sup>52</sup> from PDB2PQR webserver<sup>53</sup>, considering pH 7.0. The molecular docking calculations were performed by using AutoDock Vina<sup>54</sup>, applying a total of 16 exhaustiveness. The coordinates of the center of the conformational search box at the protein binding site were defined according to the CSP results and the atomic probability density map generated from Grb2-SH2 crystal structures complexed with phosphopeptides available in PDB website. The box dimensions were  $30 \times 26 \times 26$  Å on the three coordinate axes.

Molecular dynamics (MD) simulations were performed in GROMACS version 5.0.1<sup>55</sup> using the AMBER99SB-IDLN force field<sup>56</sup> for modeling the Grb2-SH2 domain and pY-pep. TIP3P<sup>57</sup> was used as water model. The starting position of the pY-pep for the MD simulations was obtained from the molecular docking calculations. The structures of the free protein (PDB-ID 1BMB) and complexed with pY-pep (from AutoDock Vina) were each placed in the center of a 66 Å cubic box solvated by a solution of 200 mM NaCl in water, and the protonation state of ionizable residues was determined by PROPKA results, considering a pH 7.0. The periodic boundary conditions and NPT ensemble were used in all simulations, keeping the systems at 298 K and 1.0 bar using Nose–Hoover thermostat ( $\tau_T = 2.0$  ps) and Parrinello–Rahman barostat ( $\tau_P = 2.0$  ps and compressibility =  $4.5 \times 10^{-5}$  bar $^{-1}$ ). A 12 Å cut-off point for the Lennard–Jones and Coulomb potentials was used. The long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) algorithm. The simulations were performed using a time step of 2.0 fs and all covalent bonds involving hydrogen atoms were constrained to their equilibrium distance. A conjugate gradient minimization algorithm was utilized to relax the superposition of atoms generated in the box construction process. Energy minimizations were carried out with steepest descent integrator and conjugate gradient algorithm, using 500 kJ mol $^{-1}$  nm $^{-1}$  as maximum force criterion. At the end of preparation, 1.0  $\mu\text{s}$  MD simulation of the free and pY-pep bound Grb2-SH2 were performed for data acquisition. Following dynamics, the trajectories of free and pY-pep bound domain were firstly concatenated individually and analyzed according to root mean square deviation (RMSD) of backbone atoms and root mean square fluctuation (RMSF) with respect of the C $\alpha$  atoms of the protein. A total of 100 equally spaced frames over MD simulation time containing only pY-pep structural information were extracted for the evaluation of mobility of the phosphopeptide and for the also construction of the atomic probability density map as mentioned in the previous section. After individual analysis of each simulation, the trajectories for the free and pY-pep bound protein excluding the initial 1.0 ns were concatenated in a single file, and this new trajectory was used to the Principal Component Analysis (PCA). This aspect is important because ensures the same eigenvectors (PC1 and PC2) for both MD simulations, enabling us to directly compare the occurrence of conformational sub-states. The average structure used as a reference in the PCA calculations was obtained from the first 9.0 ns. The values of PC1 and PC2 were calculated considering the respective portions of simulation time for the free and complexed protein. For a visual inspection of motions related to the principal components, 30 frames containing Grb2-SH2 structural information were extracted from PC1 and PC2 in the simulation time ranges correspondent to the free and pY-pep bound protein. The structural representations were prepared using PyMol<sup>58</sup>.

### Accession numbers

The atomic coordinates, experimental restraints and chemical shift assignments are available at the Protein Data Bank (PDBID 6VK2) and Biomagnetic Resonance Data Bank (BMRB ID 27781).

Received: 11 April 2020; Accepted: 20 July 2020

Published online: 03 August 2020

## References

- Tari, A. M. & Lopez-Berestein, G. GRB2: a pivotal protein in signal transduction. *Semin. Oncol.* **28**, 142–147 (2001).
- Ahmed, Z. *et al.* Grb2 monomer–dimer equilibrium determines normal versus oncogenic function. *Nat. Commun.* **6**, 7354 (2015).
- Lin, C.-C. *et al.* Inhibition of basal FGF receptor signaling by dimeric Grb2. *Cell* **149**, 1514–1524 (2012).
- Ornitz, D. M. & Itoh, N. The fibroblast growth factor signaling pathway. *Wiley Interdiscip. Rev. Dev. Biol.* **4**, 215–266 (2015).
- Dhillon, A. S., Hagan, S., Rath, O. & Kolch, W. MAP kinase signalling pathways in cancer. *Oncogene* **26**, 3279–3290 (2007).
- Molina, J. R. & Adjei, A. A. The Ras/Raf/MAPK pathway. *J. Thorac. Oncol.* **1**, 7–9 (2006).
- Rozakis-Adcock, M., Fernley, R., Wade, J., Pawson, T. & Bowtell, D. The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1. *Nature* **363**, 83–85 (1993).
- Machida, K. & Mayer, B. J. The SH2 domain: versatile signaling module and pharmaceutical target. *Biochim. Biophys. Acta Proteins Proteomics* **1747**, 1–25 (2005).
- Lowenstein, E. J. *et al.* The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell* **70**, 431–442 (1992).
- Vidal, M., Gigoux, V. & Garbay, C. SH2 and SH3 domains as targets for anti-proliferative agents. *Crit. Rev. Oncol. Hematol.* **40**, 175–186 (2001).
- Ijaz, M. *et al.* The role of Grb2 in cancer and peptides as Grb2 antagonists. *Protein Pept. Lett.* **24**, 1084–1095 (2018).
- Marengere, L. E. M. *et al.* SH2 domain specificity and activity modified by a single residue. *Nature* **369**, 502–505 (1994).
- Papaioannou, D., Geibel, S., Kunze, M. B. A., Kay, C. W. M. & Waksman, G. Structural and biophysical investigation of the interaction of a mutant Grb2 SH2 domain (W121G) with its cognate phosphopeptide. *Protein Sci.* **25**, 627–637 (2016).
- Benfield, A. P., Whiddon, B. B., Clements, J. H. & Martin, S. F. Structural and energetic aspects of Grb2-SH2 domain-swapping. *Arch. Biochem. Biophys.* **462**, 47–53 (2007).
- Hosoe, Y. *et al.* Structural and functional properties of Grb2 SH2 dimer in CD28 binding. *Biophys. Physicobiol.* **16**, 80–88 (2019).
- Nioche, P. *et al.* Crystal structures of the SH2 domain of Grb2: highlight on the binding of a new high-affinity inhibitor. *J. Mol. Biol.* **315**, 1167–1177 (2002).
- Senior, M. M. *et al.* The three-dimensional solution structure of the Src homology domain-2 of the growth factor receptor-bound protein-2. *J. Biomol. NMR* **11**, 153–164 (1998).
- Yuzawa, S. *et al.* Solution structure of Grb2 reveals extensive flexibility necessary for target recognition. *J. Mol. Biol.* **306**, 527–537 (2001).
- Sanches, K., Caruso, ÍP., Almeida, F. C. L. & Melo, F. A. NMR assignment of free <sup>1</sup>H, <sup>15</sup>N and <sup>13</sup>C-Grb2-SH2 domain. *Biomol. NMR Assign.* <https://doi.org/10.1007/s12104-019-09894-x> (2019).
- Shen, Y. *et al.* Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 4685–4690 (2008).
- Morando, M. A., Barbosa, G. M., Cruz-Oliveira, C., Da Poian, A. T. & Almeida, F. C. L. Dynamics of zika virus capsid protein in solution: the properties and exposure of the hydrophobic cleft are controlled by the  $\alpha$ -helix 1 sequence. *Biochemistry* **58**, 2488–2498 (2019).
- McConnell, H. M. Reaction rates by nuclear magnetic resonance. *J. Chem. Phys.* **28**, 430–431 (1958).
- Carver, J. & Richards, R. A general two-site solution for the chemical exchange produced dependence of T<sub>2</sub> upon the carr-Purcell pulse separation. *J. Magn. Reson.* **6**, 89–105 (1972).
- Machado, L. E. S. F. *et al.* Conformational dynamics of a cysteine-stabilized plant defensin reveals an evolutionary mechanism to expose hydrophobic residues. *Biochemistry* **57**, 5797–5806 (2018).
- Schwarzinger, S. *et al.* Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc.* **123**, 2970–2978 (2001).
- Sanz-Hernández, M. & De Simone, A. The PROSECCO server for chemical shift predictions in ordered and disordered proteins. *J. Biomol. NMR* **69**, 147–156 (2017).
- Kjaergaard, M., Brander, S. & Poulsen, F. M. Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *J. Biomol. NMR* **49**, 139–149 (2011).
- Ross, P. D. & Subramanian, S. Thermodynamics of protein association reactions: forces contributing to stability. *Biochemistry* **20**, 3096–3102 (1981).
- Ban, D. *et al.* Kinetics of conformational sampling in ubiquitin. *Angew. Chem. Int. Ed.* **50**, 11437–11440 (2011).
- Liu Bernard, A., Engelmann Brett, W. & Nash Piers, D. The language of SH2 domain interactions defines phosphotyrosine-mediated signal transduction. *FEBS Lett.* **586**, 2597–2605 (2012).
- Zurini, M. *et al.* Structural and conformational requirements for high-affinity binding to the SH2 domain of Grb2 1. *J. Med. Chem.* **42**, 971–980 (2002).
- Momin, M., Xin, Y. & Hamelberg, D. Allosteric fine-tuning of the binding pocket dynamics in the ITK SH2 domain by a distal molecular switch: an atomistic perspective. *J. Phys. Chem. B* **121**, 6131–6138 (2017).
- Farrow, N. A. *et al.* Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by <sup>15</sup>N NMR relaxation. *Biochemistry* **33**, 5984–6003 (1994).
- Kay, L. E., Muhandiram, D. R., Farrow, N. A., Aubin, Y. & Forman-Kay, J. D. Correlation between dynamics and high affinity binding in an SH2 domain interaction. *Biochemistry* **35**, 361–368 (1996).
- Huculeci, R. *et al.* Dynamically coupled residues within the SH2 domain of FYN are key to unlocking its activity. *Structure* **24**, 1947–1959 (2016).
- Finerty, P. J., Mittermaier, A. K., Muhandiram, R., Kay, L. E. & Forman-Kay, J. D. NMR dynamics-derived insights into the binding properties of a peptide interacting with an SH2 domain. *Biochemistry* **44**, 694–703 (2005).
- Pascal, S. M., Singer, A. U., Yamazaki, T., Kay, L. E. & Forman-Kay, J. D. Structural and dynamic characterization of an SH2 domain-phosphopeptide complex by NMR approaches. *Biochem. Soc. Trans.* **23**, 729–733 (1995).
- Huang, W. Y. C., Ditlev, J. A., Chiang, H.-K., Rosen, M. K. & Groves, J. T. Allosteric modulation of Grb2 recruitment to the intrinsically disordered scaffold protein, LAT, by remote site phosphorylation. *J. Am. Chem. Soc.* **139**, 18009–18015 (2017).
- Lindfors, H. E., Drijfhout, J. W. & Ubbink, M. The Src SH2 domain interacts dynamically with the focal adhesion kinase binding site as demonstrated by paramagnetic NMR spectroscopy. *IUBMB Life* **64**, 538–544 (2012).
- Delaglio, F. *et al.* Nmrpipe—a multidimensional spectral processing system based on unix pipes. *J. Biomol. Nmr* **6**, 277–293 (1995).
- Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
- Hyberts, S. G., Milbradt, A. G., Wagner, A. B., Arthanari, H. & Wagner, G. Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson gap scheduling. *J. Biomol. NMR* **52**, 315–327 (2012).
- Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).



44. Rieping, W. *et al.* ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23**, 381–382 (2007).
45. Linge, J. P., O'Donoghue, S. I. & Nilges, M. Automated assignment of ambiguous nuclear overhauser effects with ARIA. *Methods Enzymol.* **339**, 71–90 (2001).
46. Brunger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **54**, 905–921 (1998).
47. Maciejewski, M. W. *et al.* NMRbox: a resource for biomolecular NMR computation. *Biophys. J.* **112**, 1529–1534 (2017).
48. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
49. Loria, J. P., Rance, M. & Iii, A. G. P. A Relaxation-Compensated Carr–Purcell–Meiboom–Gill sequence for characterizing chemical exchange by NMR spectroscopy. *Analysis* 2331–2332 (1999).
50. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**(33–8), 27–28 (1996).
51. Sanner, M. F. Python: a programming language for software integration and development. *J. Mol. Graph. Model.* **17**, 57–61 (1999).
52. Olsson, M. H. M., Sondergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: consistent treatment of internal and surface residues in empirical  $pK_a$  predictions. *J. Chem. Theory Comput.* **7**, 525–537 (2011).
53. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: An automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665–W667 (2004).
54. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2009).
55. Pronk, S. *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).
56. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinforma.* **78**, 1950–1958 (2010).
57. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
58. DeLano, W. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr.* (2002).

## Acknowledgments

We thank the National Center of Nuclear Magnetic Resonance (CNRMN, UFRJ, Rio de Janeiro-Brazil) and the Mutiuser Center of Molecular Inovation (CMIB, UNESP, São José do Rio Preto, Brazil) for the NMR spectra acquisition, and NMRbox (National Center for Biomolecular NMR Data Processing and Analysis, a Biomedical Technology Research Resource—BTRR). We thank Professor John Ladbury for kindly providing the plasmid of Grb2-SH2. We also thank Prof. Dmitry Korzhnev for discussion of the results and Prof. Ana Paula Valente for the paper review and suggestions. The work was supported by FAPERJ grants 215141, 210361, 239229, and 204432, CNPq Grants 309564/2017-4 and 457773/2014-6 awarded to FCLA. FAPESP Grants (2014/17630-0, 2017/20642-8), CNPq Grant 442951/2014-0 and UNESP-PROPG grants (09/2017, 12/2017) awarded to FAM. We also thank INBEB-INCT for funding. KS is funded by CNPq scholarship and IPC is funded by a postdoctoral fellowship by FAPERJ. The authors declare no conflict of interest.

## Author contributions

K.S. contributed with the sample preparation, experimental design, resonance assignments, N.M.R. data collection, data analysis and structure calculation. I.P.C. contributed with experimental design, N.M.R. data collection and molecular dynamics simulation. F.A.M. is a biophysist that contributed with data analysis and funding. F.C.L.A. is a structural biologist and N.M.R. spectroscopist, which contributed with the structure and dynamics, data analysis, experimental design and funding. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-70034-w>.

**Correspondence** and requests for materials should be addressed to F.C.L.A. or F.A.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020