Journal of Theoretical and Computational Chemistry
Vol. 13, No. 3 (2014) 1440003 (15 pages)

DOI: 10.1142/S0219633614400033

Distinct mechanisms of a phosphotyrosyl peptide binding to two SH2 domains

Xiaodong Pang and Huan-Xiang Zhou*

Department of Physics and Institute of Molecular Biophysics
Florida State University, Tallahassee, FL 32306
**hzhou4@fsu.edu*

Received 14 October 2013
Accepted 11 December 2013
Published 26 March 2014

Protein phosphorylation is a very common post-translational modification, catalyzed by kinases, for signaling and regulation. Phosphotyrosines frequently target SH2 domains. The spleen tyrosine kinase (Syk) is critical for tyrosine phosphorylation of multiple proteins and for regulation of important pathways. Phosphorylation of both Y342 and Y346 in Syk linker B is required for optimal signaling. The SH2 domains of Vav1 and PLC-γ both bind this doubly phosphorylated motif. Here we used a recently developed method to calculate the effects of Y342 and Y346 phosphorylation on the rate constants of a peptide from Syk linker B binding to the SH2 domains of Vav1 and PLC-γ. The predicted effects agree well with experimental observations. Moreover, we found that the same doubly phosphorylated peptide binds the two SH2 domains via distinct mechanism, with apparent rigid docking for Vav1 SH2 and dock-and-coalesce for PLC-γ SH2.

Keywords: Transient complex; association rate constant; association mechanism; dock-and-coalesce; electrostatic interactions; phosphorylation.

1. Introduction

Protein phosphorylation is a common post-translational modification, catalyzed by kinases, for signaling and regulation.[1] As a result of phosphorylation or dephosphorylation of hydroxyl-containing residues (serine, threonine and tyrosine), enzymes can become activated or inactivated.[2] The phosphorylated residues may create docking sites for signal-transduction proteins, exemplified by those containing the Src homology 2 (SH2) domain.[3]

SH2 domains specifically recognize phosphotyrosine (pY). They were first identified in the oncoproteins Src and Fps.[4] They contain approximately 100 amino acids, comprised of two α-helices and seven β-strands.[5,6] As shown in Fig. 1, the two helices (αA and αB) sandwich a central β-sheet formed by strands βB, βC and βD.

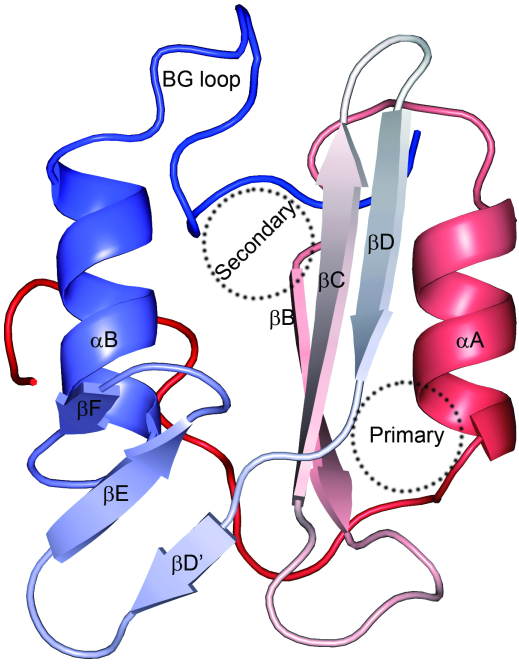


Fig. 1. Structure of an SH2 domain (as found in PDB entry 2FCI; peptide not displayed), displayed as ribbon with color varying from red at the N-terminus to blue at the C-terminus. The primary and secondary sites for phosphotyrosine binding are indicated by dash circles.

An SH2 domain harbors two pY binding sites, referred to as primary and secondary and located on opposite sides of the central β-sheet. The primary site, positioned between the central β-sheet and αA, typically binds peptides with a single pY, which can interact with two highly conserved residues, an arginine and a histidine, at the βB5 and βD4 positions, respectively. Tyrosine phosphorylation on the peptides could increase the binding affinities by several orders of magnitude.[7] For peptides with two pY residues, the secondary site, situated between the central β-sheet and αB, is also engaged, resulting in enhanced binding affinity and higher sequence specificity.[8]

The spleen tyrosine kinase (Syk) has multiple proteins as substrates and is essential in B cell signaling.[9] Murine Syk contains three conserved tyrosine residues, Y317, Y342 and Y346, in linker B, which connects the second SH2 domain and the catalytic domain. Phosphorylation of Y317 negatively regulates the function of Syk,[10] whereas phosphorylation of both Y342 and Y346 is required for optimal signaling.[11] The SH2 domains of Vav1 and PLC-γ can both bind to the doubly phosphorylated Syk linker B (referred to as pYpY hereafter), with pY342 in the primary site and pY346 in the secondary site.[8,12] The dissociation constants of pYpY are 7.4 μM and 0.07 μM, respectively, for Vav1 SH2 and PLC-γ SH2.[8,12] For Vav1 SH2, the binding affinity is reduced by 10-fold when Y342 is unphosphorylated (abbreviated as YpY)

and by 2-fold when Y346 is unphosphorylated (abbreviated as pYY).[12] The binding affinity of pYY for PLC-γ SH2 is 6-fold lower.[8] For both systems, the experimental studies[8,12] showed that the decreases in affinity come mostly from the decreases of association rate constant (k_a), while the dissociation rate constants (k_d) are only modestly affected.

What is the physical basis for the enhancements in association rate constants (and binding affinities) by tyrosine phosphorylation? A direct consequence of phosphorylation is the addition of one to two negative charges (depending on the pH) to each tyrosine, which may introduce favorable electrostatic interactions with the SH2 domains. Electrostatic interactions have indeed been suggested to make significant contributions to the binding affinities of pY-containing peptides for SH2 domains.[13] In particular the addition of a phosphate to the YEEI peptide results in a $\sim 10^4$-fold increase in binding affinity for Src SH2.[7]

We and others have found that, for stereospecific protein association, electrostatic attraction can significantly enhance association rate constants, by as much as three or more orders of magnitude.[14–21] To calculate the association rate constant, we envisioned that association proceeds by first reaching via translational and rotational diffusion a transient complex, in which the two subunits have near-native separation and relative orientation but have yet to form most of the native interactions. Thereafter the two subunits undergo conformational rearrangement and further tightening to form the native complex. Assuming that the second substep is fast (see Ref. 22), the overall association rate constant can be predicted as[15]

$$k_a = k_{a0} \exp(-\Delta G^*_{\rm el}/k_B T), \tag{1}$$

where k_{a0} is the "basal" rate constant for reaching to transient complex by unbiased diffusion, $\Delta G^*_{\rm el}$ is the electrostatic interaction energy in the transient complex and $k_B T$ is the thermal energy. Our method for k_a calculations has been automated[18] and is accessible as the TransComp web server at http://pipe.sc.fsu.edu/transcomp/. The basal rate constant is mainly determined by the shape of the interface between the two subunits in the native complex, and $\Delta G^*_{\rm el}$ is dictated by the degree of charge complementarity across the interface. By separating the two contributing factors to k_a, we gain better understanding on how the magnitudes of k_a are modulated in different protein complexes.

In this study, we used TransComp to calculate the effects of Y342 and Y346 phosphorylation on the rate constants of a peptide (DTEVY$_{342}$ESPY$_{346}$ADPE) from Syk linker B binding to the SH2 domains of Vav1 and PLC-γ. The predicted effects agree well with the experimental observations.[8,12] Moreover, we found that the same doubly phosphorylated peptide binds the two SH2 domains via distinct mechanisms. The binding with Vav1 SH2 can be characterized as rigid docking, whereby contacts with both the primary and secondary sites are formed nearly all at once to produce the native complex. In contrast, binding to PLC-γ SH2 proceeds via a dock-and-coalesce mechanism, whereby docking of pY342 to the primary site is followed by coalescing of pY346 and the flanking residues

around the secondary site, leading to eventual insertion of pY346 into the secondary site.

2. Methods

2.1. *TransComp calculations*

The procedure for TransComp calculations has been described previously.[15,18] Here we briefly summarize the three constituent steps. The transient complex is identified through mapping the interaction energy landscape in and around the native-complex energy well. Starting from the structure of the native complex, the two subunits are translated and rotated, and configurations that are clash-free are saved. For each clash-free configuration, the number of inter-subunit contacts (N_c) is calculated. As the two subunits move away from the native complex, N_c decreases but the rotational freedom, as measured by the standard deviation (σ_χ) of the relative rotation angle (χ) among the clash-free configurations at a given N_c, increases (Fig. 2 and S3). The dependence of σ_χ on N_c is fitted to a function used for modeling reversible two-state protein denaturation. The midpoint of this fit, where N_c is denoted as N_c^*, defines the transient complex. That is, configurations with $N_c = N_c^*$, make up the transient-complex ensemble.

The basal rate constant k_{a0} is calculated from force-free Brownian dynamics simulations according to an algorithm developed previously.[31] Each Brownian trajectory starts from a configuration inside the native-complex energy well (i.e. $N_c > N_c^*$) and is propagated in the translational and rotational space. At each step during the simulation, if $N_c > N_c^*$ then the subunits are allowed to form the native complex; if successful the trajectory is then terminated. The survival fraction of the Brownian trajectories as a function of time is used to calculate k_{a0}.

To calculate the electrostatic interaction energy$\Delta G_{\rm el}^*$, 100 configurations are randomly chosen from the transient-complex ensemble, using the APBS program[32] and the results averaged.

All our calculations were done through the TransComp web server at http://pipe.sc.fsu.edu/transcomp/. The inputs, including the structure of the native complex and the ionic strength, are described below.

2.2. *pYpY binding to Vav1 SH2*

Here we used the first model in the NMR structure (PDB entry 2LCT) of the Syk-derived doubly phosphorylated peptide pYpY (DTEVY$_{342}$ESPY$_{346}$ADPE) bound to Vav1 SH2.[12] We acetylated the N-terminus and amidated the C-terminus of the peptide and optimized the positions of hydrogen atoms by energy minimization (Amber with Amber99SB force field[33]). The resulting structure in PQR format (with Amber charges and Bondi radii) was input to the TransComp server, along an ionic strength of 120 mM (experimental condition of Ref. 12). The partial charges for phosphotyrosine were from Homeyer *et al.*[34]

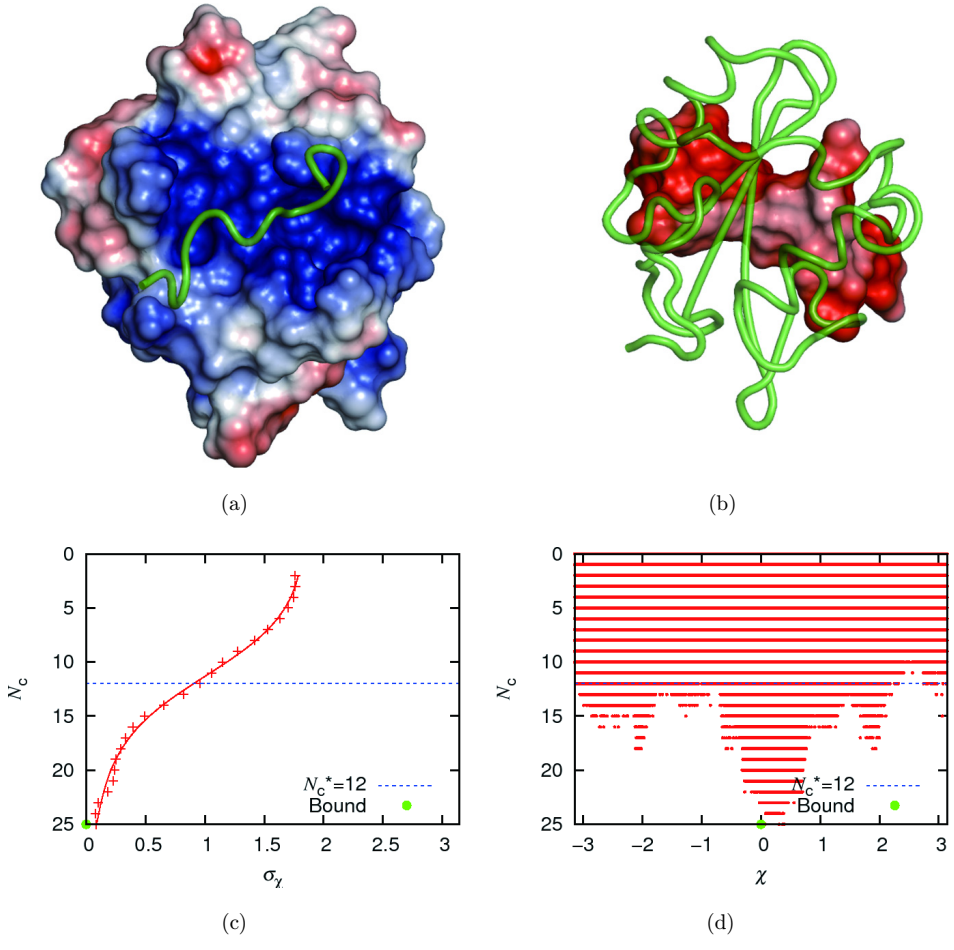


Fig. 2. Output of the TransComp run for pYpY binding to Vav1 SH2. (a) Vav1 SH2 is shown as electrostatic surface accompanied by a ribbon representation of pYpY to indicate the binding site. Blue and red indicate positive and negative electric potentials, respectively. (b) The representations of Vav1 SH2 and pYpY are reversed and the view is rotated around vertical axis by 180°. (c) The N_c versus σ_χ curve. (d) Scatter plot of N_c versus χ for clash-free configurations. The native complex and the transient complex are indicated by a green circle and blue line, respectively.

2.3. *pYpY binding to PLC-γ SH2*

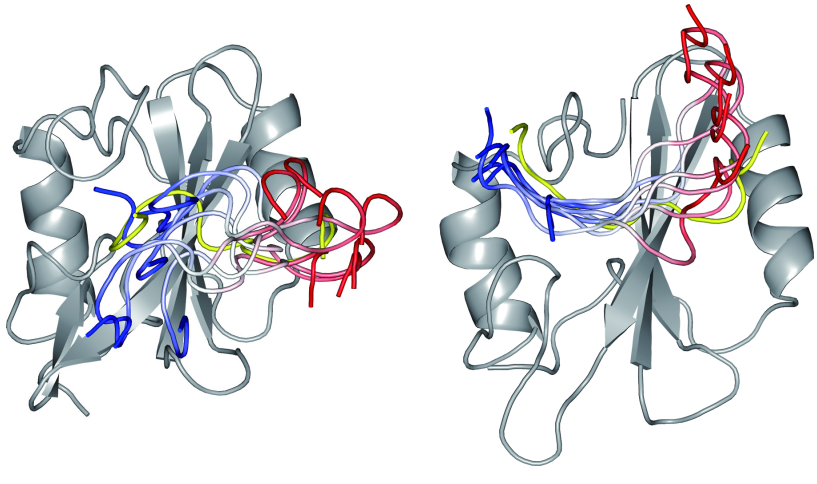
An initial, aborted TransComp run used the first model in the NMR structure (PDB entry 2FCI[8]) for the PLC-γ SH2-pYpY native complex as input. In subsequent calculations, we used a docked intermediate modeled on the PLC-γ SH2-pY1021 complex (PDB 2PLD[27]) (Fig. S2). After aligning Y342 of pYpY to the only tyrosine of pY1021 (Fig. S2A), a homology model for the PLC-γ SH2-pYpY docked intermediate was generated by the Modeller program.[35] Subsequent steps were similar to

those described in the last subsection, except that the ionic strength was 155 mM (experimental condition of Ref. 8).

3. Results

3.1. *Syk linker B peptide binding to Vav1 SH2*

TransComp calculations for association rate constants use the structures of the native complexes as input.[18] Each k_a calculation consists of three steps: generation of the transient-complex ensemble; determination of the basal rate constant k_{a0} for reaching the transient complex; and evaluation of the electrostatic interaction free energy $\Delta G^*_{\rm el}$ in the transient complex. For pYpY binding to Vav1 SH2, we used the structure in Protein Data Bank (PDB) entry 2LCT[12] as input. The output from the server is shown in Fig. 2. The transient complex was identified by mapping the interaction energy surface in the six-dimensional space of relative translation and relative rotation in and around the native-complex energy well. As illustrated in Fig. 2(d), in the native-complex energy well, the subunits form a large number of contacts (N_c) but have little freedom in the relative rotation angle (χ). Outside the native-complex energy well, the two subunits lose most of the contacts but gain nearly complete rotational freedom. The transient complex is located at the rim of the native-complex energy well (Fig. 2(d)), identified by the mid-point of the transition from restricted rotation to complete rotational freedom (Fig. 2(c)). Representative configurations of the transient complex are shown in Fig. 3(a).



(a) (b)

Fig. 3. Representative transient-complex configurations for the (a) Vav1 SH2-pYpY and (b) PLC-γ SH2-pYpY systems. Seven poses of pYpY with color varying from red at the N-terminus to blue at the C-terminus are shown. pYpY in the native complex is also shown in yellow as reference.

Table 1. Calculated and experimental results for the binding kinetics of Vav1 SH2 with the Syk linker B peptide in different phosphorylation forms (ionic strength = 120 mM). Phosphotyrosines are in the singly charged state.

Peptides/SH2	$k_{a0}(10^6\,\mathrm{M^{-1}s^{-1}})$	ΔG^*_{el} (kcal/mol)	$k_a(10^6\,\mathrm{M^{-1}s^{-1}})$	Fold decrease	Experimental fold decrease[12]
pYpY	0.49	−3.34	136	—	—
YpY	0.49	−1.64	7.78	17.5	17.3
pYY	0.49	−2.65	42.5	3.2	2.1
YY	0.49	−0.93	2.37	57.4	—
βD3 K→Q[a]	0.49	−2.57	37.6	3.6	—

[a]Binding of pYpY to Vav1 SH2 βD3 K→Q mutant.

Successful completion of a TransComp run is a sign that the complex is formed via rigid docking, whereby the inter-subunit contacts in the native complex are formed nearly all at once.[18,23,24] This binding mechanism for forming the Vav1 SH2-pYpY complex is further described below.

The rate constants (k_a) and their contributing factors (k_{a0} and ΔG^*_{el}) for the binding of Vav1 SH2 with the Syk linker B peptide in different phosphorylation forms are listed in Table 1 (with net charge of phosphotyrosine at −1). The basal rate constant k_{a0} is largely dictated by the shape and size of binding interface, with more convoluted and larger interfaces leading to lower k_{a0} whereas flatter and smaller interface leading to higher k_{a0}.[18,21,24] For the doubly phosphorylated peptide, a basal rate constant of $0.49 \times 10^6\,\mathrm{M^{-1}s^{-1}}$ and an electrostatic interaction energy of −3.34 kcal/mol in the transient complex were found, resulting in an association rate constant of $1.36 \times 10^8\,\mathrm{M^{-1}s^{-1}}$. The significant attractive electrostatic interaction between the subunits comes from strong complementarity between the cationic binding site on Vav1 SH2 and the anionic pYpY peptide, as shown in Figs. 2(a) and 2(b).

We investigated the contributions of Y342 and Y346 phosphorylation to the Vav1 SH2-pYpY association rate constant by calculating the k_a values for the singly phosphorylated (YpY and pYY and unphosphorylated (YY) forms. Following our previous studies,[18,19] we assumed that a change in the phosphorylation state of a tyrosine residue, just like a mutation, does not affect the location of the transient complex. Hence the basal rate constant remained the same and only ΔG^*_{el} was re-calculated for YpY, pYY and YY, using the transient complex determined for pYpY binding. With Y342 unphosphorylated, ΔG^*_{el} changed from −3.34 kcal/mol to −1.64 kcal/mol, resulting in a 17.5-fold decrease in k_a. In comparison, with Y346 unphosphorylated, ΔG^*_{el} changed to −2.65 kcal/mol, resulting in a 3.2-fold decrease in k_a.

Experimentally, the decreases in k_a for binding YpY and pYY were 17.3- and 2.1-fold, respectively.[12] These results agree well with our calculations, suggesting that the contributions of Y342 and Y346 phosphorylation to the Vav1 SH2-pYpY association rate constant can be largely accounted for by the electrostatic attraction between the added phosphate groups on the peptide and the cationic residues on the

binding site of Vav1 SH2. We note that the magnitudes of the experimental associate rate constants, determined by surface plasmon resonance (SPR), were much lower than those of our calculations. As noted previously,[16,25] compared to methods such as stopped-flow spectroscopy that operate in solution, SPR has confounding effects such as mass transport and surface immobilization, which can be especially severe for fast association, a scenario implicated here by our calculation for the Vav1 SH2-pYpY system and recognized for other SH2-phosphotyrosyl systems.[25] Nevertheless SPR can still be useful for measuring changes in k_a for related systems.

The above calculations were performed with phosphotyrosine in the singly charged state. For comparison, we performed calculations with phosphotyrosine in the doubly charged state (Table S1). For pYpY, the electrostatic interaction energy changed to -4.61 kcal/mol, and the association rate constant became $9.58 \times 10^8\,\mathrm{M^{-1}s^{-1}}$. With Y342 unphosphorylated (YpY), ΔG^*_{el} weakened to -1.97 kcal/mol, corresponding to a 71.0-fold decrease in k_a. On the other hand, with Y346 unphosphorylated (pYY), ΔG^*_{el} changed to -3.70 kcal/mol, corresponding to a 3.9-fold decrease in k_a. This 71.0-fold decrease in k_a from dephosphorylation of Y342 is significantly higher than the experimental observation (17.3-fold decrease). We used the singly charged state of phosphotyrosine for the rest of the paper.

Our calculations showed that, when both Y342 and Y346 are unphosphorylated, ΔG^*_{el} further decreased in magnitude to -0.93 kcal/mol (corresponding to a 60-fold decrease in k_a). This ΔG^*_{el} is nearly equal to what is expected when the effects of separately dephosphorylating the two tyrosines are added. This additive effect is consistent with the wide separations between the two phosphate groups (distance between the two phosphorus atoms $\sim$18 Å) and between the primary and secondary sites in the native complex. It is unknown whether the YY peptide still has measurable binding affinity for Vav1 SH2. If so our calculated k_a stands to be tested. We also calculated the effect of mutating a conserved lysine (at βD3) around the secondary site, and found that binding of pYpY to this K$\rightarrow$Q mutant has $\Delta G^*_{\mathrm{el}} = -2.57$ kcal/mol, resulting in a 3.6-fold decrease in k_a. This result also remains to be tested.

The transient-complex ensemble used in the above calculations was generated using the first model in the NMR structure of the Vav1 SH2-pYpY complex. To test the sensitivity of the calculation results to the structure used, we repeated the calculations on all the 20 models in the NMR structure. Of the 20 models, calculations using 10 models completed successfully, while TransComp runs on the rest failed because of poor fit of the N_c versus σ_χ curve, possibly arising from low quality of the models. While the results, shown in Table S2, exhibited significant variations among the 10 models, the average quantities are very close to those reported above for model 1.

Very recently Chen *et al.*[26] determined the structure of the Vav1 SH2-YpY complex (PDB entry 2MC1). Using this structure we repeated the k_a calculation for the Vav1 SH2-YpY system. The results were very similar to those obtained above by using the transient complex determined for the Vav1 SH2-pYpY system: The basal

rate constant was modest higher (by 1.8-fold) while ΔG^*_{el} was virtually unchanged. Our calculated results thus appear to be robust.

3.2. *Syk linker B peptide binding to PLC-γ SH2*

Using the structure for the PLC-γ SH2-pYpY complex (PDB entry 2FCI[8]) as input, we carried out a TransComp run for this system. However, this TransComp run was aborted, due to a large gap of 14 in the sampled N_c values. Such a large N_c gap showed that the formation of the PLC-γ SH2-pYpY native complex could not be modeled as rigid docking.[18] In support of this contention, comparison of chemical shift perturbations of PLC-γ SH2 upon pYpY and pYY binding indicated that the pY342 residues of the two peptides formed similar interactions at the primary site, but pY346, relative to its unphosphorylated counterpart, induced significant conformational changes around the secondary site.[8] The latter observation is reinforced when one compares the structures of PLC-γ SH2 bound to either the doubly phosphorylated pYpY peptide or the singly phosphorylated peptide from the platelet-derived growth factor receptor (referred to as pY1021; PDB entry 2PLD).[27] These structures show a degree of similarity at the primary site but extensive differences at the secondary site (Fig. S1A). In the pY1021-bound structure, the BG loop wedges between the αB helix and the βD strand, and pushes away the C-terminal portion of the peptide. In contrast, in the pYpY-bound structure, the BG loop moves outward, allowing the αB helix to closely approach the central β-sheet and bind the secondary site. Moreover, two sidechains, from βD3 K56 and αB9 Y84, extend over the secondary site to cover the bound pY346 (Fig. S1B).

We hypothesized that the binding of pYpY to PLC-γ SH2 involves two steps. The first step is the docking of pY342 to the primary site. The resulting complex will be referred to as the docked intermediate, which should resemble the (yet undetermined) structure formed by binding pYY. The second step is the coalescence of the C-terminal portion of pYpY around the secondary site and the eventual insertion of pY346 into the secondary site. This dock-and-coalesce mechanism was found to be widely followed by the binding of intrinsically disordered proteins (IDPs) to structured targets.[18,28] Accordingly, we built a homology model for the docked intermediate using the structure of the PLC-γ SH2-pY1021 complex[27] as template (Fig. S2). In the model for the docked intermediate, pY342 is inserted into the primary site, but pY346 points away from the SH2 domain, which does not have a well-formed secondary site (Fig. S2C).

Using the model for the docked intermediate as input, a TransComp run completed successfully. The output is shown in Fig. S3. Representative configurations in the transient complex are shown in Fig. 3(b). The calculated k_{a0}, ΔG^*_{el} and k_a results for the docking step are listed in Table 2. Assuming that the coalescing step is fast (relative to the undocking step), these calculations yield the overall rate constant for the full binding of the peptide to the primary and secondary sites. For the binding of pYpY, the basal rate constant was $0.39 \times 10^5\,M^{-1}s^{-1}$ and the electrostatic

Table 2. Calculated and experimental results for the binding kinetics of PLC-γ SH2 with the Syk linker B peptide in different phosphorylation forms (ionic strength = 155 mM). Phosphotyrosines are in the singly charge state.

Peptides/SH2	$k_{a0}(10^5\,\mathrm{M^{-1}s^{-1}})$	ΔG^*_{el} (kcal/mol)	$k_a(10^6\,\mathrm{M^{-1}s^{-1}})$	Fold decrease	Experimental fold decrease[8]
pYpY	0.39	−3.84	23.8	—	—
YpY	0.39	−2.91	5.22	4.6	—
pYY	0.39	−3.26	9.40	2.5	6
YY	0.39	−2.30	1.89	12.6	—
βD3 K→Q[a]	0.39	−2.75	4.01	5.9	0.6

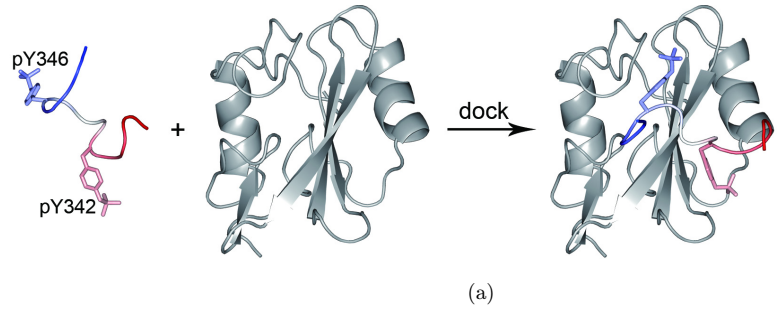
[a]Binding of pYpY to PLC-γ SH2 βD3 K→Q mutant.

interaction energy in the transient complex was −3.84 kcal/mol, resulting in a k_a of $2.38 \times 10^7\,\mathrm{M^{-1}s^{-1}}$. When either Y342 (YpY) or Y346 (pYY) or both (YY) are unphosphorylated, weakened electrostatic attraction resulted in decreases in the association rate constant by 4.6-, 2.5- and 12.6-fold, respectively. Experimentally it was observed that the association rate constant of pYY was smaller than that of pYpY by 6-fold.[8] We also calculated the effect of the PLC-γ SH2 βD3 K→Q mutation on the docking rate constant of pYpY and predicted a 5.9-fold decrease. Experimentally the mutation was found to result in slight increase (1.5-fold) in the overall binding rate constant. A better model for the docked intermediate perhaps can reduce the discrepancy for the K→Q mutant.

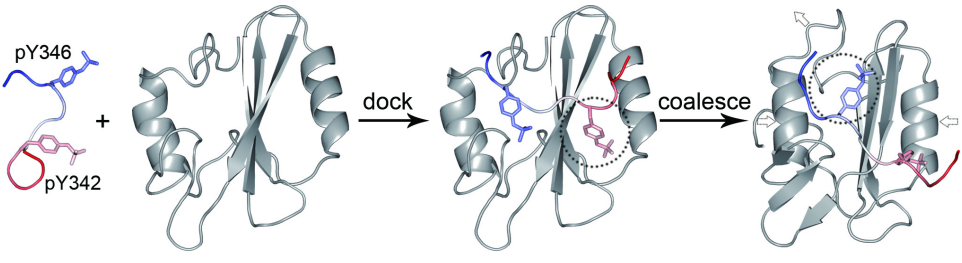
3.3. *Distinct mechanisms for binding to Vav1 and PLC-γ SH2 domains*

Our TransComp calculations suggested that the doubly phosphorylated Syk B linker peptide binds to Vav1 SH2 and PLC-γ SH2 via distinct mechanisms (Fig. 4). Apparently, the pYpY peptide can rapidly sample near-native conformations and, upon approaching Vav1 SH2, forms appropriate inter-subunit contacts at both the primary and secondary sites; additional fast conformational rearrangement of both the peptide and the SH2 domain finally results in the native complex (Fig. 4(a)).

In contrast, when binding to PLC-γ SH2, pYpY forms near-native contacts at the primary and secondary sites not all at once but sequentially (Fig. 4(b)). pY342 first docks to the primary site while the secondary site has yet to be properly formed. Once anchored on PLC-γ SH2, both the C-terminal portion of the peptide and the structural elements around the future secondary site undergo extensive conformational changes. These include outward movement of the BG loop and approach of the αB helix toward the central β-sheet. The result is the proper formation of the secondary site and insertion of pY346 into this site.



(a)



(b)

Fig. 4. Two distinct binding mechanisms. (a) pYpY rapidly samples near-native conformations and then docks to the primary and secondary sites of Vav1 SH2 at once nearly as a rigid body. (b) pYpY binds to the primary and secondary sites of PLC-γ SH2 in a sequential manner. First a docking step anchors pY342 to the primary site. In the subsequent coalescing step, the C-terminal portion of the peptide and the elements around the would-be secondary site undergo conformational rearrangement, leading to the insertion of pY346 into the secondary site.

4. Discussion

4.1. *Roles of electrostatic interactions in SH2 recognition*

For binding to Vav1 SH2 and PLC-γ SH2, phosphorylation of a single tyrosine was found to result in an increase in binding affinity of up to 10-fold.[8,12] For other systems, much greater effects have been reported.[7] An increase in affinity can be achieved by either an increase in the association rate constant or a decrease in the dissociation rate constant or a combination. For binding to both Vav1 SH2 and PLC-γ SH2, the contributions of tyrosine phosphorylation have been linked mostly to increases in association rate constant, with dissociation rate constants only modestly affected.[8,12] Our calculations here have now further established that the increases in association rate constant are due to the additional electrostatic attraction with the SH2 domains afforded by the phosphate groups on the tyrosines. This conclusion is in line with observations on many other protein–protein complexes, where electrostatic attraction has been implicated in elevating the probability of the two subunits being in near-native separation and orientation, leading to significant enhancement of association rates.[14–21,29,30]

At the primary site, the conserved βB5 arginine has been shown to be important for the electrostatic interaction with a pY.[7] In the secondary site of both Vav1 SH2 and PLC-γ SH2, βD3 lysine is in close contact with pY346 of the Syk B linker peptide.[8,12] Mutation of this lysine to a neutral residue glutamine would eliminate the electrostatic interaction with pY346 and possibly neighboring anionic residues on the peptide. In the case of Vav1 SH2, our calculations predicted that the βD3 K→Q mutation would decrease the association rate constant by 3.6-fold. Though this prediction is yet to be tested experimentally, it appears reasonable, since it is similar to the effect of dephosphorylating pY346 (3.2-fold decrease in k_a) and the latter result is consistent with experiment.[12] In the case of PLC-γ SH2, the βD3 K→Q mutation was found experimentally to increase k_a by 1.5-fold[8] whereas our calculations predicted a decrease in k_a by 5.9-fold. These calculations were based on assuming the dock-and-coalesce mechanism and used a modeled structure for the docked intermediate; the latter, as suggested above, could contribute to the discrepancy from experiment. Further discussion is given below.

4.2. *Dock-and-coalesce mechanism*

Our calculations identified two distinct mechanisms for the Syk linker B peptide pYpY's binding to Vav1 SH2 and PLC-γ SH2. In the former case, pYpY rapidly samples near-native conformations and then docks to the primary and secondary sites at once nearly as a rigid body. In the latter case, pYpY binds to the primary and secondary sites in a sequential manner. First pY342 docks to the primary site, and then the C-terminal portion of the peptide and the elements around the would-be secondary site undergo conformational rearrangement, leading to the insertion of pY346 into the secondary site. The former rigid docking mechanism is common for the association between globular proteins that do not undergo significant conformational changes,[18,23] whereas the latter dock-and-coalesce mechanism is common for the binding of IDPs to structured targets.[18,28]

The distinct binding mechanisms have some experimental support. For Vav1 SH2, the chemical shift perturbations of residues around the primary and secondary sites were found to be identical upon binding pYpY, YpY and pYY.[26] This observation can be easily explained if the Vav1 SH2-pYpY complex is stabilized by a cooperative set of interactions that spans both the primary and secondary sites, such that dephosphorylation of either one of the two phosphotyrosines does not sufficiently weaken the cooperative set. This is consistent with the proposed rigid-docking mechanism. In contrast, for pYpY and pYY binding to PLC-γ SH2, chemical shift perturbations showed negligible differences for residues around the primary site but significant differences for residues around the secondary site.[8] This observation is consistent with the proposed dock-and-coalesce mechanism, in which the docking of pY342 into the primary site facilitates subsequent conformational search. When Y346 is also phosphorylated, additional coalesced structure in the SH2-peptide interface is formed.

Our k_a calculations for peptide binding to PLC-γ SH2 were based on the dock-and-coalesce mechanism and used a modeled structure for the docked intermediate. Moreover, we assumed that the docking step is rate-limiting or, equivalently, the coalescing step is much faster than the undocking step. If the docking step is only partially rate-limiting, then we could underestimate the effect of a mutation that slows down the coalescing step and overestimate the effect of a mutation that speeds up the coalescing step. Our calculations apparently underestimated the effect of dephosphorylating pY346 (calculated 2.5-fold decrease in k_a versus an experimental 6-fold decrease[8]). It is indeed reasonable to expect that dephosphorylating pY346 would slow down the coalescing step. On the other hand, we overestimated the adverse effect of the βD3 K→Q mutation (calculated 5.9-fold decrease in k_a versus a slight 1.5-fold increase by experiment[8]). Upon pYpY binding in the wild-type SH2, the sidechain of βD3 K56 (along with that of αB9 Y84) extends over the secondary site to cover the bound pY346 (Fig. S1B). This could be a slow event that is eliminated by the K→Q mutation, a scenario that is consistent with the observation that the mutation increases the dissociation rate constant by 10-fold.[8] It is thus conceivable that the K→Q mutant has a faster coalescing step than the wild-type SH2.

What is the advantage of dock-and-coalesce mechanism? In a case where the proper formation of a secondary site involves significant conformational rearrangement, anchoring of the peptide to the primary site enables the cooperation of the elements around the secondary site and the cognate residues on the peptide in their search for the native structure, leading to acceleration of the overall binding process. Further increase in binding affinity can be produced in the coalescing step by covering up the ligand bound to the secondary site, as in the PLC-γ SH2-pYpY system through the extension of the βD3 K56 and αB9 Y84 sidechains over the bound pY346 (Fig. S1B).

5. Conclusions

We have used TransComp to calculate the rate constants for the binding of Vav1 SH2 and PLC-γ SH2 with a Syk-derived peptide in different phosphorylation forms. The calculations allowed for the isolation of the contributions of electrostatic interactions in SH2 recognition. Moreover, we identified distinct mechanisms for recruiting the doubly phosphorylated peptide to the two SH2 domains: rigid docking for Vav1 SH2 and dock-and-coalesce for PLC-γ SH2. Applications to other binding processes regulated by phosphorylation will further advance our quantitative and mechanistic understanding of this important post-translational modification.

Acknowledgments

This work is supported by National Institutes of Health Grant GM058187.

References

1. Pawson T, Scott JD, Protein phosphorylation in signaling–50 years and counting, *Trends Biochem Sci* **30**:286–290, 2005.
2. Graves JD, Krebs EG, Protein phosphorylation and signal transduction, *Pharmacol Ther* **82**:111–121, 1999.
3. Pawson T, Specificity in signal transduction: From phosphotyrosine-SH2 domain interactions to complex cellular systems, *Cell* **116**:191–203, 2004.
4. Sadowski I, Stone JC, Pawson T, A Noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of fujinami sarcoma-virus P130gag-Fps, *Mol Cell Biol* **6**:4396–4408, 1986.
5. 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* **72**:779–790, 1993.
6. Liu BA, Engelmann BW, Nash PD, The language of SH2 domain interactions defines phosphotyrosine-mediated signal transduction, *FEBS Lett* **586**:2597–2605, 2012.
7. Bradshaw JM, Mitaxov V, Waksman G, Investigation of phosphotyrosine recognition by the SH2 domain of the Src kinase, *J Mol Biol* **293**:971–985, 1999.
8. Groesch TD, Zhou F, Mattila S, Geahlen RL, Post CB, Structural basis for the requirement of two phosphotyrosine residues in signaling mediated by Syk tyrosine kinase, *J Mol Biol* **356**:1222–1236, 2006.
9. Sada K, Takano T, Yanagi S, Yamamura H, Structure and function of Syk protein-tyrosine kinase, *J Biochem* **130**:177–186, 2001.
10. Keshvara LM, Isaacson CC, Yankee TM, Sarac R, Harrison ML, Geahlen RL, Syk- and Lyn-dependent phosphorylation of Syk on multiple tyrosines following B cell activation includes a site that negatively regulates signaling, *J Immunol* **161**:5276–5283, 1998.
11. Hong JJ, Yankee TM, Harrison ML, Geahlen RL, Regulation of signaling in B cells through the phosphorylation of Syk on linker region tyrosines — A mechanism for negative signaling by the Lyn tyrosine kinase, *J Biol Chem* **277**:31703–31714, 2002.
12. Chen CH, Martin VA, Gorenstein NM, Geahlen RL, Post CB, Two closely spaced tyrosines regulate NFAT signaling in B cells via Syk association with Vav, *Mol Cell Biol* **31**:2984–2996, 2011.
13. Grucza RA, Bradshaw JM, Mitaxov V, Waksman G, Role of electrostatic interactions in SH2 domain recognition: Salt-dependence of tyrosyl-phosphorylated peptide binding to the tandem SH2 domain of the Syk kinase and the single SH2 domain of the Src kinase, *Biochemistry* **39**:10072–10081, 2000.
14. Gabdoulline RR, Wade RC, Protein–protein association: Investigation of factors influencing association rates by Brownian dynamics simulations, *J Mol Biol* **306**:1139–1155, 2001.
15. Alsallaq R, Zhou HX, Electrostatic rate enhancement and transient complex of protein–protein association, *Proteins* **71**:320–335, 2008.
16. Schreiber G, Haran G, Zhou H-X, Fundamental aspects of protein–protein association kinetics, *Chem Rev* **109**:839–860, 2009.
17. Zhou HX, Rate theories for biologists, *Q Rev Biophys* **43**:219–293, 2010.
18. Qin S, Pang XD, Zhou HX, Automated prediction of protein association rate constants, *Structure* **19**:1744–1751, 2011.
19. Pang XD, Qin SB, Zhou HX, Rationalizing 5000-fold differences in receptor-binding rate constants of four cytokines, *Biophys J* **101**:1175–1183, 2011.
20. Grant BJ, Gheorghe DM, Zheng W, Alonso M, Huber G, Dlugosz M, McCammon JA, Cross RA, Electrostatically biased binding of kinesin to microtubules, *PLoS Biol* **9**: e1001207, 2011.

21. Pang XD, Zhou KH, Qin SB, Zhou HX, Prediction and dissection of widely-varying association rate constants of actin-binding proteins, *PLoS Comput Biol* **8**:e1002696, 2012.
22. Zhou HX, Wlodek ST, McCammon JA, Conformation gating as a mechanism for enzyme specificity, *Proc Natl Acad Sci USA* **95**:9280–9283, 1998.
23. Zhou HX, Bates PA, Modeling protein association mechanisms and kinetics, *Curr Opin Struct Biol* **23**:887–893, 2013.
24. Zhou HX, Intrinsic disorder: signaling via highly specific but short-lived association, *Trends Biochem Sci* **37**:43–48, 2012.
25. Ladbury JE, Lemmon MA, Zhou M, Green J, Botfield MC, Schlessinger J, Measurement of the binding of tyrosyl phosphopeptides to SH2 domains: A reappraisal, *Proc Natl Acad Sci USA* **92**:3199–3203, 1995.
26. Chen CH, Piraner D, Gorenstein NM, Geahlen RL, Post CB, Differential recognition of syk-binding sites by each of the two phosphotyrosine-binding pockets of the Vav SH2 domain, *Biopolymers* **99**:897–907, 2013.
27. Pascal SM, Singer AU, Gish G, Yamazaki T, Shoelson SE, Pawson T, Kay LE, Forman-Kay JD, Nuclear magnetic resonance structure of an SH2 domain of phospholipase C-gamma 1 complexed with a high affinity binding peptide, *Cell* **77**:461–472, 1994.
28. Zhou HX, Pang X, Lu C, Rate constants and mechanisms of intrinsically disordered proteins binding to structured targets, *Phys Chem Chem Phys* **14**:10466–10476, 2012.
29. Hemsath L, Dvorsky R, Fiegen D, Carlier MF, Ahmadian MR, An electrostatic steering mechanism of Cdc42 recognition by Wiskott-Aldrich syndrome proteins, *Mol Cell* **20**:313–324, 2005.
30. Ahmad M, Gu W, Helms V, Mechanism of fast peptide recognition by SH3 domains, *Angew Chem Int Ed Engl* **47**:7626–7630, 2008.
31. Zhou HX, Kinetics of diffusion-influenced reactions studied by Brownian dynamics, *J Phys Chem* **94**:8794–8800, 1990.
32. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA, Electrostatics of nanosystems: Application to microtubules and the ribosome, *Proc Natl Acad Sci USA* **98**:10037–10041, 2001.
33. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C, Comparison of multiple amber force fields and development of improved protein backbone parameters, *Proteins* **65**:712–725, 2006.
34. Homeyer N, Horn AH, Lanig H, Sticht H, AMBER force-field parameters for phosphorylated amino acids in different protonation states: Phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine, *J Mol Model* **12**:281–289, 2006.
35. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen M-y, Pieper U, Sali A, Comparative protein structure modeling using MODELLER, *Curr Protoc Protein Sci*, Chap 2 unit 2.9, 2007.