Molecular Mechanism of Selective Recruitment of Syk Kinases by the Membrane Antigen-Receptor Complex

SUPPLEMENTARY MATERIALS

Peter J. Bond[†] and José D. Faraldo-Gómez*

Max Planck Institute of Biophysics &
Cluster of Excellence 'Macromolecular Complexes'
60438 Frankfurt am Main, Germany

*Corresponding author:
jose.faraldo@biophys.mpg.de

[†]Current address:
Unilever Centre for Molecular Science Informatics,
University of Cambridge, UK

May 10, 2011

SUPPLEMENTARY METHODS

Energetics of EF loop conformational change in the cSH2 domain

The Gibbs free-energy difference between the open (o) and closed (c) states of the cSH2 domain is defined as:

$$\Delta G_{\mathrm{oc}} = -k_{\mathrm{B}}T \ln \frac{Q_{\mathrm{c}}}{Q_{\mathrm{o}}} + P(V_{\mathrm{c}} - V_{\mathrm{o}}) \qquad [1]$$

where Q and V respectively denote the partition function and volume of the molecular system in either state. It can be shown, however, that the contribution of the pressure-volume term is negligible ($P\Delta V << k_{\mathrm{B}}T$), because the volume of the system is essentially unchanged before and after the conformational change in the protein (Figure A). Thus ΔG_{oc} can be written as a ratio of configurational integrals:

$$e^{-\beta \Delta G_{\mathrm{oc}}} = \frac{\int_{\mathbf{X} \in \text{closed}} e^{-\beta H(\mathbf{X},\mathbf{P})} d\mathbf{X}\, d\mathbf{P}}{\int_{\mathbf{X} \in \text{open}} e^{-\beta H(\mathbf{X},\mathbf{P})} d\mathbf{X}\, d\mathbf{P}} \qquad [2]$$

where $H = K(\mathbf{P}) + U(\mathbf{X})$ is the Hamiltonian of the system, $\mathbf{X}$ and $\mathbf{P}$ are the coordinates and momenta of the constituent atoms, and $\beta = 1/k_{\mathrm{B}}T$. Since both integrals are identical with regard to the momenta,

$$e^{-\beta \Delta G_{\mathrm{oc}}} = \frac{\int_{\mathbf{X} \in \text{closed}} e^{-\beta U(\mathbf{X})} d\mathbf{X}}{\int_{\mathbf{X} \in \text{open}} e^{-\beta U(\mathbf{X})} d\mathbf{X}} \qquad [3]$$

It is convenient to rewrite this expression as:

$$e^{-\beta \Delta G_{\mathrm{oc}}} = \frac{\int_{\mathbf{X} \in \text{closed}} e^{-\beta U(\mathbf{X})} d\mathbf{X}}{\int_{\mathbf{X} \in \text{closed}} e^{-\beta (U(\mathbf{X}) + u_{\mathrm{c}}(\mathbf{X}))} d\mathbf{X}} \times \frac{\int_{\mathbf{X} \in \text{closed}} e^{-\beta (U(\mathbf{X}) + u_{\mathrm{c}}(\mathbf{X}))} d\mathbf{X}}{\int_{\mathbf{X} \in \text{open}} e^{-\beta (U(\mathbf{X}) + u_{\mathrm{o}}(\mathbf{X}))} d\mathbf{X}} \times \frac{\int_{\mathbf{X} \in \text{open}} e^{-\beta (U(\mathbf{X}) + u_{\mathrm{o}}(\mathbf{X}))} d\mathbf{X}}{\int_{\mathbf{X} \in \text{open}} e^{-\beta U(\mathbf{X})} d\mathbf{X}} \qquad [4]$$

The last term in this equation can be interpreted as the free-energy cost of introducing a biasing potential $u_{\mathrm{o}}(\mathbf{X})$ that constraints the molecular system to remain in open state; the first term is the analogous quantity for the closed state, sign-inverted. The middle term is the work required to drive the system from the (constrained) open sate to the (constrained) closed state. Therefore,

$$\Delta G_{\mathrm{oc}} = -\Delta G_{\mathrm{c}} + \Delta W_{\mathrm{oc}} + \Delta G_{\mathrm{o}} \qquad [5]$$

We first focus on ΔW_{oc}, which we rewrite as:

$$\Delta W_{\mathrm{oc}} = \frac{\int_{\mathbf{X} \in \text{closed}} e^{-\beta U_{\mathrm{c}}(\mathbf{X})} d\mathbf{X}}{\int_{\mathbf{X} \in \text{open}} e^{-\beta U_{\mathrm{o}}(\mathbf{X})} d\mathbf{X}} \qquad [6]$$

i.e. we have defined $U_{\mathrm{c}} = U + u_{\mathrm{c}}$ and $U_{\mathrm{o}} = U + u_{\mathrm{o}}$. To obtain ΔW_{oc} we use Thermodynamic Integration (TI) method; this a well-established scheme to calculate the reversible work associated with a change in the potential energy function U of a molecular system – in our case from U_{o} to U_{c}. To do so, the potential energy function must be written as a

function of a continuous parameter λ, such that $U(\lambda) = U_o$ if $\lambda = 0$ and $U(\lambda) = U_c$ if $\lambda = 0$. The reversible work (or free-energy difference) is then

$$\Delta W_{oc} = \int_0^1 d\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda \qquad [7]$$

where $\langle\ \rangle$ denotes a simulation average at a given value of λ.

In the cSH2-domain calculations, we define $U(\lambda)$ in terms of a collective variable RMSD(λ) that describes the conformation of the EF loop with respect to a reference structure:

$$U(\lambda) = U + \frac{k}{2}\left[\mathrm{RMSD}(\lambda)\right]^2 = U + \frac{k}{2}\frac{1}{N}\sum_{i=1}^{N}\left[\phi_i - \phi_i^\lambda\right]^2 \qquad [8]$$

Here, ϕ_i represents each of the N peptide-bond dihedral-angles involved in the conformational change in the EF loop, and ϕ_i^λ denotes the set of intermediate values each dihedral adopts in between the two end-points, $\phi_{i,o}$ and $\phi_{i,c}$. That is,

$$\phi_i^\lambda = (1-\lambda)\phi_{i,o} + \lambda\phi_{i,c} \qquad [9]$$

The resulting values of ΔW_{oc} are shown in Fig. 2, plotted as a function of λ.

We now proceed to demonstrate that this is the only significant contribution to the total change in free energy, i.e. $\Delta G_{oc} \sim \Delta W_{oc}$, because the first and last terms in the right-hand side of Eq. 5 approximately cancel each other. To see this, let us consider the last term:

$$e^{-\beta\Delta G_o} = \frac{\int_{\mathbf{X}\in\,\mathrm{open}} e^{-\beta(U(\mathbf{X})+u_o(\mathbf{X}))}d\mathbf{X}}{\int_{\mathbf{X}\in\,\mathrm{open}} e^{-\beta U(\mathbf{X})}d\mathbf{X}} = \frac{\int_{\Phi\in\,\mathrm{open}} e^{-\beta(W(\Phi)+u_o(\Phi))}d\Phi}{\int_{\Phi\in\,\mathrm{open}} e^{-\beta W(\Phi)}d\Phi} \qquad [10]$$

Here, we have rewritten the configurational integral in Eq. 4 in terms of the potential of mean-force $W(\Phi)$ along the dihedral angles influenced by the constraining potential u_o, i.e. $\Phi = (\phi_1, \phi_2 \ldots \phi_N)$. Analysis of unbiased simulations in either the open or closed state shows that the probability distributions associated with $W(\Phi)$ are approximately described by Gaussian functions, with average $\bar{\phi}_i$ and variance σ_i (Table B):

$$e^{-\beta W(\Phi)} \cong \prod_{i=1}^{N} C_i \exp\left\{-\frac{(\phi_i - \bar{\phi}_i)^2}{2\sigma_i^2}\right\} \qquad [11]$$

This approximation allows for an analytical solution to Eq. 10. For the denominator:

$$\int_{\Phi\in\,\mathrm{open}} e^{-\beta W(\Phi)}d\Phi = \prod_{i=1}^{N}\int d\phi_i\, C_{i,o}\exp\left\{-\frac{(\phi_i - \bar{\phi}_{i,o})^2}{2\sigma_{i,o}^2}\right\} \approx \prod_{i=1}^{N} C_{i,o}\,\sigma_{i,o}\sqrt{2\pi} \qquad [12]$$

The numerator is slightly more complicated, as it also includes the biasing potential,

$$u_o = \sum_{i=1}^{N} \tfrac{1}{2} k(\phi_i - \hat{\phi}_{i,o})^2 \qquad [13]$$

Note that these harmonic functions are slightly off-centered with respect to the unbiased probability distributions, i.e. $\hat{\phi}_{i,\mathrm{o}} \neq \overline{\phi}_{i,\mathrm{o}}$, because $\hat{\phi}_{i,\mathrm{o}}$ correspond to the crystal-structure values (Table B). Nonetheless, an analytical solution is again possible considering the Gaussian-product rule:

$$\begin{aligned}\int_{\Phi\in\,\mathrm{open}} e^{-\beta(W(\Phi)+u_{\mathrm{o}}(\Phi))}d\Phi &= \prod_{i=1}^{N}\int d\phi_i\, C_{i,\mathrm{o}} \exp\left\{-\frac{(\phi_i-\overline{\phi}_{i,\mathrm{o}})^2}{2\sigma_{i,\mathrm{o}}^2}\right\}\exp\left\{-\beta\frac{k}{2}(\phi_i-\hat{\phi}_{i,\mathrm{o}})^2\right\} \\ &= \prod_{i=1}^{N}\int d\phi_i\, C'_{i,\mathrm{o}} \exp\left\{-\frac{(\phi_i-\phi'_{i,\mathrm{o}})^2}{2\sigma'^2_{i,\mathrm{o}}}\right\} \approx \prod_{i=1}^{N} C'_{i,\mathrm{o}}\,\sigma'_{i,\mathrm{o}}\sqrt{2\pi}\end{aligned} \qquad [14]$$

where the following definitions have been introduced for clarity:

$$\begin{aligned}\phi'_{i,\mathrm{o}} &= \frac{\overline{\phi}_{i,\mathrm{o}}+\beta k\sigma_{i,\mathrm{o}}^2\,\hat{\phi}_{i,\mathrm{o}}}{1+\beta k\sigma_{i,\mathrm{o}}^2}, \\ \sigma'_{i,\mathrm{o}} &= \sqrt{\frac{\sigma_{i,\mathrm{o}}^2}{1+\beta k\sigma_{i,\mathrm{o}}^2}}, \\ C'_{i,\mathrm{o}} &= C_{i,\mathrm{o}}\exp\left\{-\frac{\beta k}{1+\beta k\,\sigma_{i,\mathrm{o}}^2}(\overline{\phi}_{i,\mathrm{o}}-\hat{\phi}_{i,\mathrm{o}})^2\right\}\end{aligned} \qquad [15]$$

Substituting the expressions in Eqs. 12, 14 and 15 into Eq. 10, we obtain:

$$e^{-\beta\Delta G_{\mathrm{o}}} = \prod_{i=1}^{N}\exp\left\{\frac{-\beta k(\overline{\phi}_{i,\mathrm{o}}-\hat{\phi}_{i,\mathrm{o}})^2}{1+\beta k\sigma_{i,\mathrm{o}}^2}\right\}\frac{1}{\sqrt{1+\beta k\sigma_{i,\mathrm{o}}^2}} \qquad [16]$$

Analogously for the closed-state constraint u_{c} (first term in Eqs. 4 and 5):

$$e^{-\beta\Delta G_{\mathrm{c}}} = \prod_{i=1}^{N}\exp\left\{\frac{-\beta k(\overline{\phi}_{i,\mathrm{c}}-\hat{\phi}_{i,\mathrm{c}})^2}{1+\beta k\sigma_{i,\mathrm{c}}^2}\right\}\frac{1}{\sqrt{1+\beta k\sigma_{i,\mathrm{c}}^2}} \qquad [17]$$

The values of $\overline{\phi}_{i,\alpha}$, $\sigma_{i,\alpha}$, and $\hat{\phi}_{i,\alpha}$ for all the dihedral-angles considered in the calculation of ΔG_{oc} are shown in Table B, alongside the resulting values of ΔG_{o} and ΔG_{c} for both the ZAP-70 and Syk cSH2 domains. As mentioned above, this analysis demonstrates that $|\Delta G_{\mathrm{o}} - \Delta G_{\mathrm{c}}| < k_{\mathrm{B}}T$. That is, the (mostly entropic) cost of introducing the dihedral-angle constraint $u(\Phi)$ in either the open or closed state of the EF-loop is similar. In conclusion, the approximation $\Delta G_{\mathrm{oc}} \sim \Delta W_{\mathrm{oc}}$ implicitly adopted in Fig. 2 is valid.

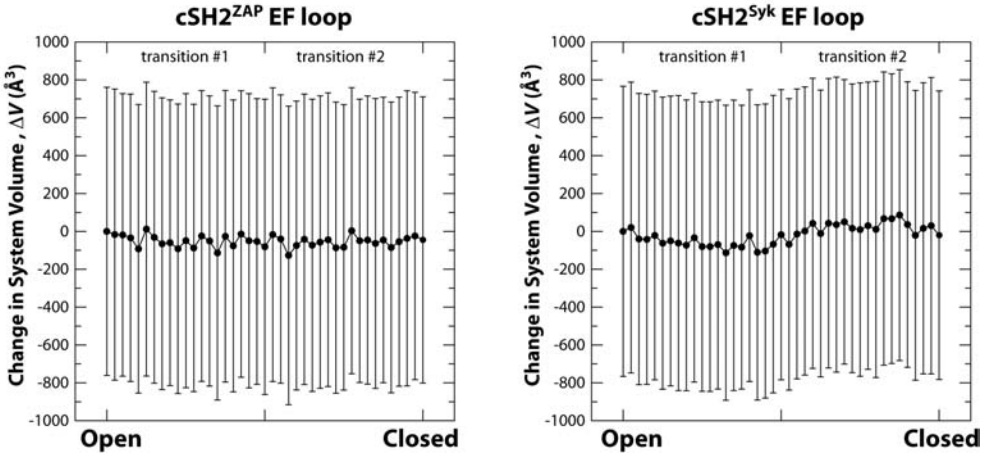

Figure A. Change in the volume of the simulation system during gating of the cSH2 domain EF loop. The data plotted derives from the TI simulations of the gating transition for the cSH2 domains of either ZAP-70 or Syk (Figs. 2, S1). Average volumes were computed for every window along the transition, and ΔV was obtained by subtracting the value for the open state. Error bars represent the natural fluctuations in the volume expected under NPT conditions. As the plots demonstrate, ΔV is in the order of 100 Å^3 or less – a value significantly smaller than the thermal fluctuations. At a pressure P = 1 atm, the work associated with such a small change is negligible; specifically, $P\Delta V < 1.5$ cal/mol << $k_B T$.

cSH2^{ZAP}	$\overline{\phi}_{i,o}$	$\sigma_{i,o}$	$\hat{\phi}_{i,o}$	$\overline{\phi}_{i,c}$	$\sigma_{i,c}$	$\hat{\phi}_{i,c}$
ϕ_{226}	-100.7	22.1	-103.7	96.8	17.4	95.1
Ψ_{226}	-171.0	15.2	-170.2	-4.2	13.4	-0.9
ϕ_{225}	67.9	11.0	67.2	-80.8	14.2	-84.8
Ψ_{224}	121.9	9.0	121.0	-22.2	11.3	-24.1
	ΔG_o = 6.44 kcal/mol			**ΔG_c = 6.60 kcal/mol**		
cSH2^{Syk}	$\overline{\phi}_{i,o}$	$\sigma_{i,o}$	$\hat{\phi}_{i,o}$	$\overline{\phi}_{i,c}$	$\sigma_{i,c}$	$\hat{\phi}_{i,c}$
ϕ_{226}	-99.4	24.3	-107.0	91.0	13.5	95.1
Ψ_{226}	-183.4	17.8	-169.4	14.7	19.0	-0.9
ϕ_{225}	64.6	10.4	67.4	-86.4	17.7	-84.8
Ψ_{224}	123.2	9.2	124.4	-17.1	22.2	-24.1
	ΔG_o = 7.04 kcal/mol			**ΔG_c = 7.58 kcal/mol**		

Table B. Average and variance of the unbiased probability distributions of the dihedral-angles employed to simulate the gating of the cSH2 domain, for either the open or closed states of the EF loop (in degrees), in ZAP-70 and Syk. The dihedral-angle reference values used in constrained TI simulations (at $\lambda = 0$ or $\lambda = 1$) are provided alongside. From both sets of values, the free-energy cost associated with imposing such constraints may be computed, according to Eqs. 16 and 17 ($N = 4$, $\beta^{-1} = 0.6$ kcal/mol, $k = 2400$ kcal/mol rad^{-2}).

SUPPLEMENTARY FIGURES

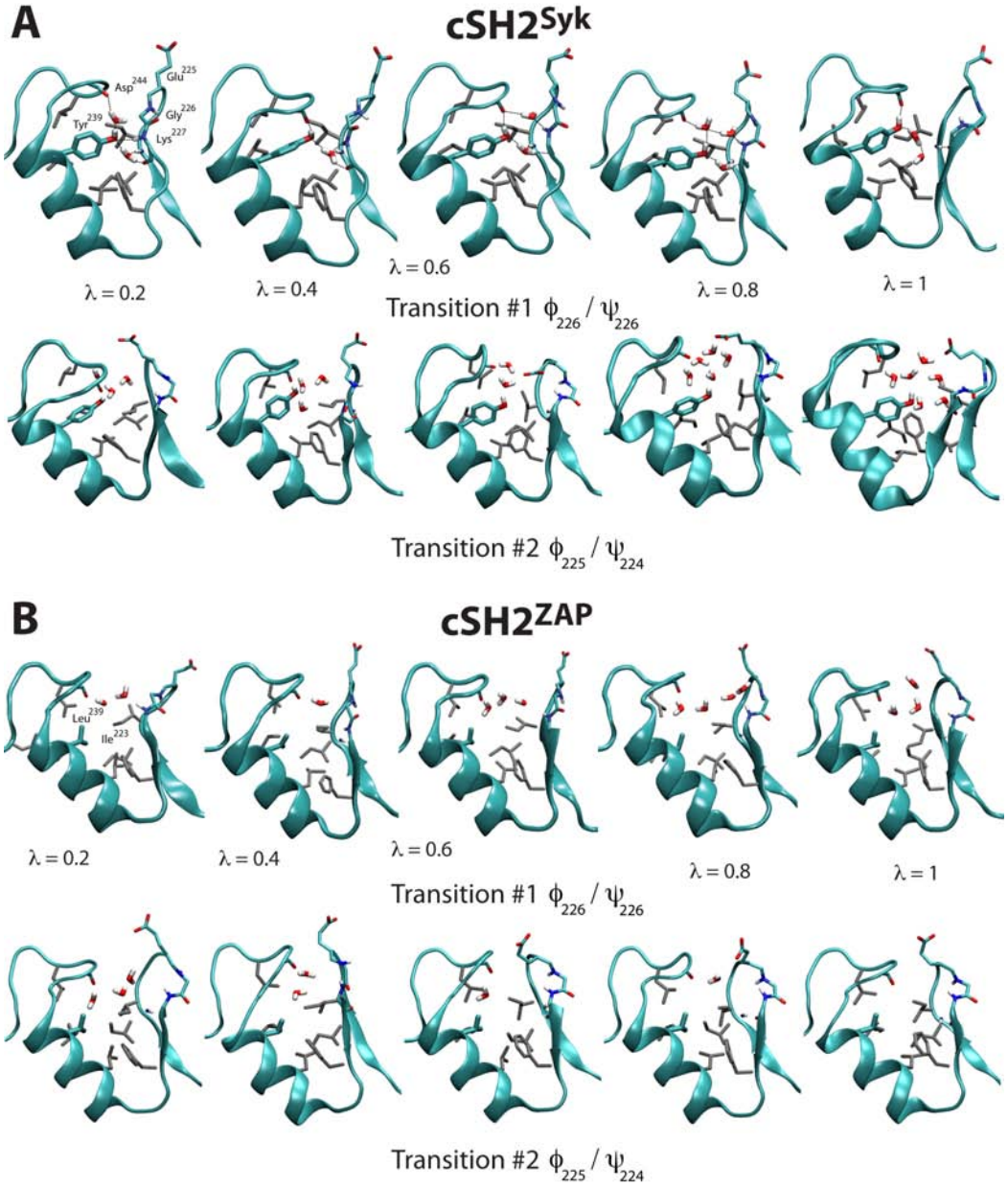

Figure S1. Analysis of EF-loop gating calculations. **(A)** Snapshots of the cSH2^{Syk} pY+3 specificity binding pocket along the conformational pathway defined by (top) the dihedral-angle pair ψ^{226} and ϕ^{226}, and (bottom) the dihedral-angle pair (ψ^{224}, ϕ^{225}). (**B**) Analogous snapshots for cSH2^{ZAP}. The cSH2 domains are shown in cartoon format, highlighting the protein-water interaction network in the binding pocket; conserved hydrophobic residues shown in grey. Closure of the EF loop in Syk is structurally hindered by persistent hydration of Tyr^{239}; by contrast, loop closure in the ZAP-70 cSH2 proceeds readily as water is expelled from its more hydrophobic pocket.

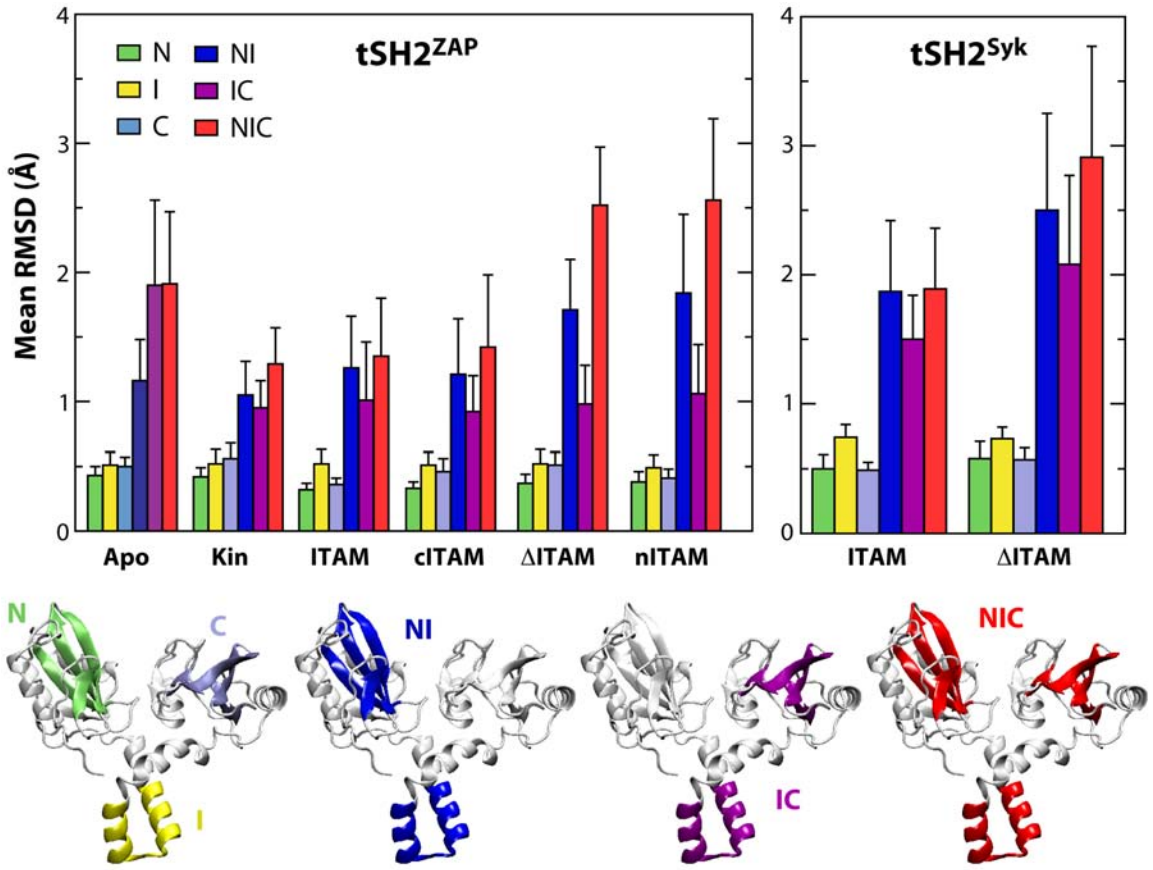

Figure S2. Conformational drift in the simulated ensembles. The RMSD of the main-chain atoms, compared to the corresponding crystal structure, is shown for selected regions within each protein in the simulations of ZAP-70 (left graph) and Syk (right graph). The histogram bars are color-coded as indicated below the graphs, by equivalently colored protein segments in the tandem. These are: the core β-strands of the nSH2 domain (green, labelled *N*; residues 33-38, 45-51, 56-62, 69-70); the core β-strand of the cSH2 domain (ice-blue, labelled *C*; residues 186-191, 197-203, 208-214, 221-222); the core α-helical region of the inter-SH2 linker (yellow, labelled *I*; residues 121-142); and combinations of these groups, labelled *NI* (dark blue), *IC* (purple), and *NIC* (red). Means and standard deviations are over the ensemble of simulations, after discarding the initial 5 ns of each trajectory.

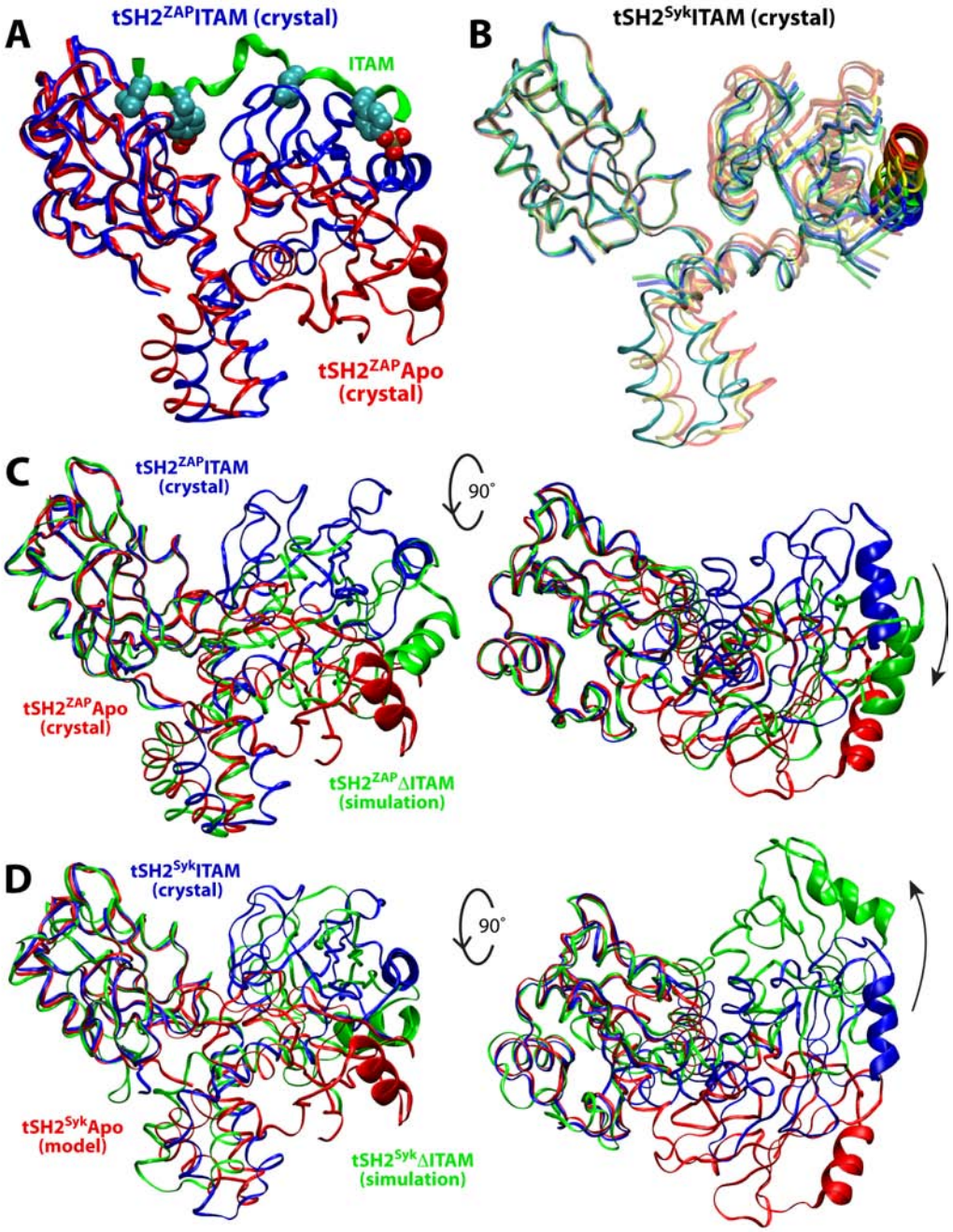

Figure S3. Representative experimental and simulated structural ensembles of the ZAP-70 and Syk tandems, in each case following structural fits onto the nSH2 domain. In all figures, helix αA in the cSH2 domain is emphasized in thick cartoons format for reference. **(A)** The ZAP-70 crystal structure (Hatada, Nature 1995) tandem (blue) bound to an ITAM peptide (green) is superimposed onto the apo state (Folmer, Biochemistry 2002) (red). The pairs of pY and pY+3 moieties in the ITAM peptide are shown in space-filling format. In **(B),** the six SH2-tandem molecules of the ITAM-bound Syk asymmetric unit are shown (Futterer, J. Mol. Biol. 1998), with the ITAM peptide removed for clarity. In **(C),** a representative simulation snapshot of $tSH2^{ZAP}\Delta ITAM$ (from the region of maximum probability density in Figure 4B, main text) is shown in green, overlaid on the ITAM-bound (blue) and apo (red) ZAP-70 crystal structures. Note how the simulated conformation has drifted *towards* the apo state. In **(D)**, a representative simulation snapshot of $tSH2^{Syk}\Delta ITAM$ (from the region of maximum probability density in Figure 4D, main text) is shown in green, overlaid on the ITAM-bound Syk crystal structure (blue) and a model of the Syk apo state (red) based on the ZAP-70 crystal structure. Note how this conformation has drifted *further away* from the putative apo state.

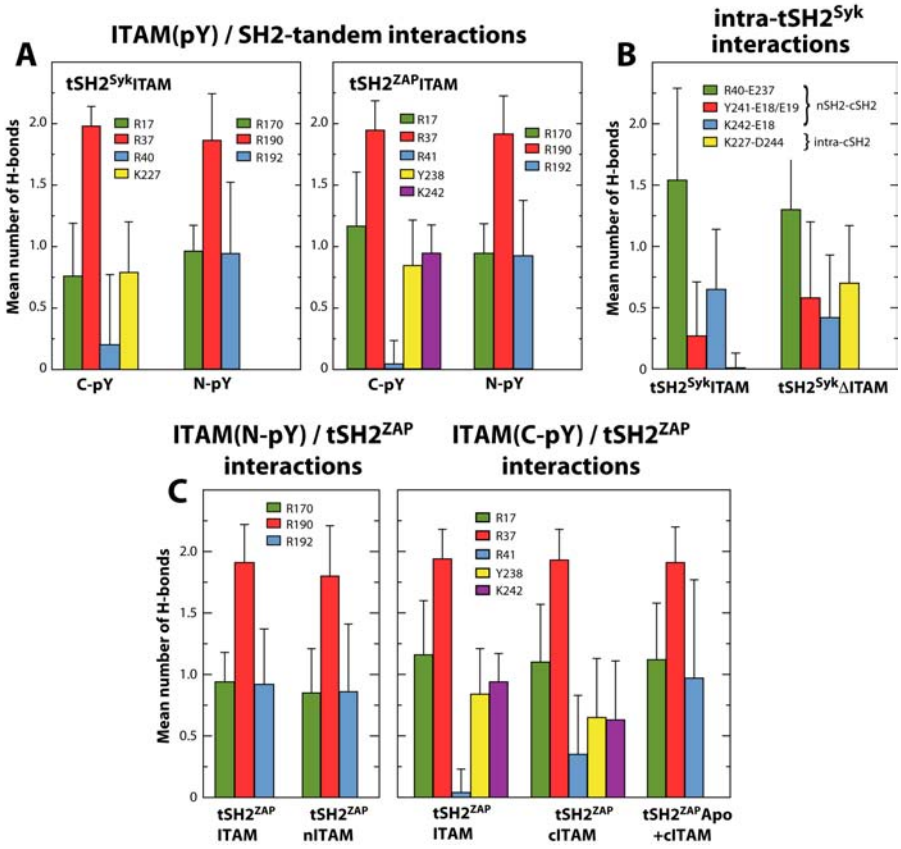

Figure S4. Key interactions between ITAM and the SH2 tandem. Means and standard deviations are obtained for each simulation, after discarding the initial 5 ns. **(A)** A comparison of key protein-phosphate interactions in ITAM-bound states of ZAP-70 and Syk. The mean number of H-bonds are shown between ITAM N-terminal (N-pY) or C-terminal (C-pY) pTyr phosphates and selected SH2 side-chains in the ITAM-bound simulations of ZAP-70 (right graph) and Syk (left graph). The C-pY and N-pY interactions are illustrated graphically in Fig. 5 and in Fig. S4D, respectively. Both ZAP-70 and Syk share a set of conserved arginine-phosphate interactions, localized in each SH2 domain. In ZAP-70, H-bonds are also stable between C-pY and Lys242 (and Tyr238) from the opposing SH2 domain, as inferred in the crystal structure. In Syk, Lys227 similarly H-bonds to C-pY, despite being disordered in its corresponding crystal structure. **(B)** A comparison of key protein-protein interactions in the ITAM-bound state of Syk, versus those formed after removing ITAM (illustrated graphically in Fig. 5). None of these inter-SH2 interactions are observed in the ZAP-70 simulations, due to a series of charged-to-hydrophobic amino acid substitutions (Fig. 5). When ITAM is dissociated from the Syk tandem, the inter-SH2 H-bonds remain, and help to maintain a degree of coupling between the SH2 domains. In addition, Lys227, previously bound to C-pY, rotates away from its ligand-bound location to interact with Asp244 in the cSH2 domain, removing the excess positive charge from the SH2-SH2 interface. **(C)** A comparison of key protein-phosphate interactions in the various ITAM-bound ZAP-70 constructs. The mean number of H-bonds between ITAM N-pY or C-pY phosphates and selected SH2 side-chains are shown. For N-pY (left graph), three conserved arginines provide a similar coordination network in the fully ITAM-bound (illustrated graphically in Fig. S4D) or nITAM-bound states. In contrast, whilst conserved Arg17 and Arg37 of the nSH2 domain always interact with C-pY (right graph), the third coordination site for phosphate is contributed by either domain in alternative bound states. Specifically, in the apo-conformation, cITAM-bound ZAP-70 tandem ($tSH2^{ZAP}$Apo+cITAM), the SH2 domains are uncoupled, and Arg41 from the nSH2 reorients to interact with C-pY (illustrated graphically in Fig. S4E). However, upon full ITAM binding ($tSH2^{ZAP}$ITAM), C-pY replaces this interaction, and now interacts with Lys242 (and Tyr238) from the opposing cSH2 domain, as observed in the crystal structure (illustrated graphically in Fig. 5). This new interaction is important to stabilize the conformational change in the tSH2 tandem. Consistently, the cITAM-bound ZAP-70 tandem in its ITAM-bound conformation ($tSH2^{ZAP}$cITAM) exhibits C-pY interactions intermediate between the other two states. This leads to a molecular mechanism of ITAM engagement in which first cITAM is captured by the isolated nSH2 domain, with C-pY interacting with Arg17, Arg37, and, transiently, Arg41. Subsequently, the possibility of exchanging Arg41 with Lys242 in the cSH2 domain stabilizes the reorientation of this domain; both SH2 domains thus become coupled via the C-pY group in ITAM.

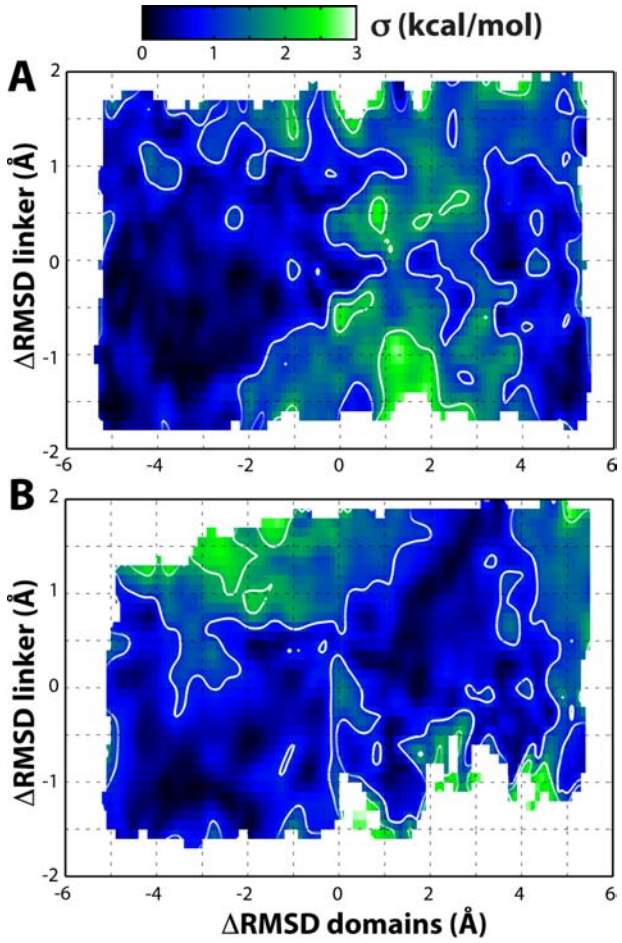

Figure S5. Estimated statistical uncertainty of the conformational free-energy landscapes calculated for the ZAP-70 SH2 tandem, in the absence **(A)** or presence **(B)** of a C-terminal ITAM peptide fragment (Fig. 6). This uncertainty is estimated from the standard deviation of three free-energy surfaces analogous to those shown in Fig. 6, calculated independently. Contours indicate changes of 1 kcal/mol.

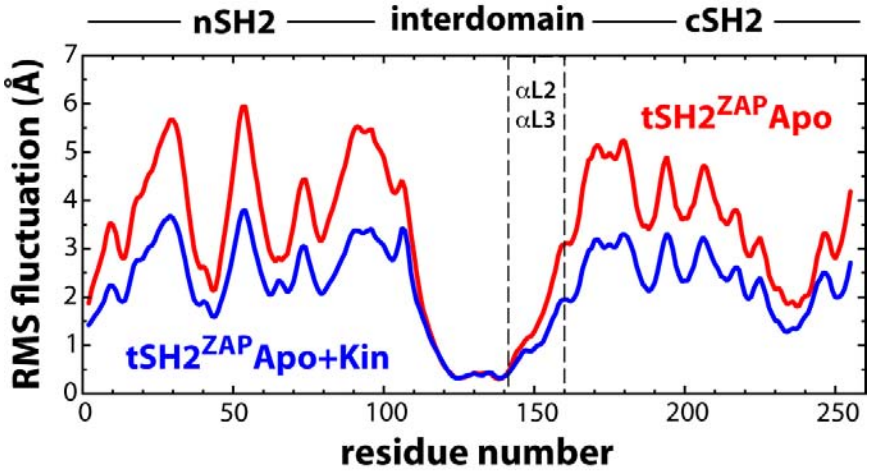

Figure S6. Root-mean squared fluctuation (RMSF) of the ZAP-70 apo SH2 tandem about the average structure, either isolated ($tSH2^{ZAP}$Apo) or in the kinase-bound ($tSH2^{ZAP}$Apo+Kin) state. On average, the RMSF profile for the isolated tandem exhibits a ~50 % increase in magnitude outside of the inter-domain linker region in comparison with the kinase-bound system. Note the conformational restriction imposed by the kinase domain in the αL2-αL3 region. The mean RMSF for each simulation ensemble was calculated for the Cα-atoms, after discarding the initial 5 ns of each trajectory, and least-squares fitting all snapshots to the average structure.

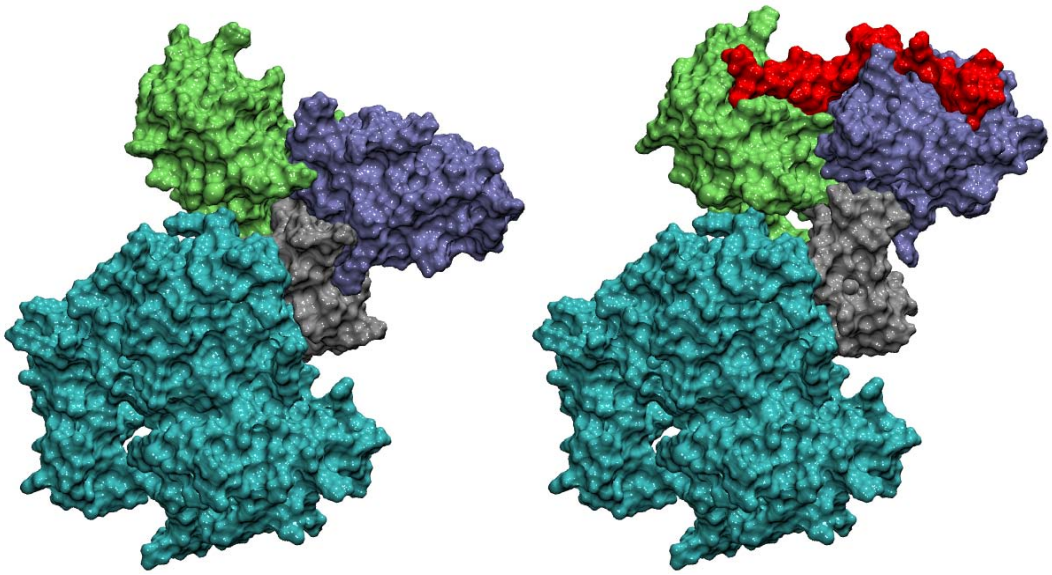

Figure S7. Left, crystal structure of the autoinhibited ZAP-70 kinase complex; catalytic and regulatory domains are colored as in Fig. 1. Right, molecular model where the ITAM-bound state of the SH2 tandem replaced the unliganded form. Note the rearrangement of tandem upon ITAM binding is not self-evidently incompatible with the catalytic-domain interface, e.g. the cSH2 domain reorients away from the kinase, while the ITAM binding sites are located remotely from the interface.