

Multidomain assembled states of Hck tyrosine kinase in solution

Sichun Yang^{a,2}, Lydia Blachowicz^a, Lee Makowski^b, and Benoît Roux^{a,b,1}

^aDepartment of Biochemistry and Molecular Biology, University of Chicago, 929 East 57th Street, Chicago, IL 60637; and ^bBiosciences Division, Argonne National Laboratory, Argonne, IL 60439

Edited* by John Kuriyan, University of California, Berkeley, CA, and approved July 15, 2010 (received for review April 5, 2010)

An approach combining small-angle X-ray solution scattering (SAXS) data with coarse-grained (CG) simulations is developed to characterize the assembly states of Hck, a member of the Src-family kinases, under various conditions in solution. First, a basis set comprising a small number of assembly states is generated from extensive CG simulations. Second, a theoretical SAXS profile for each state in the basis set is computed by using the Fast-SAXS method. Finally, the relative population of the different assembly states is determined via a Bayesian-based Monte Carlo procedure seeking to optimize the theoretical scattering profiles against experimental SAXS data. The study establishes the concept of basis-set supported SAXS (BSS-SAXS) reconstruction combining computational and experimental techniques. Here, BSS-SAXS reconstruction is used to reveal the structural organization of Hck in solution and the different shifts in the equilibrium population of assembly states upon the binding of different signaling peptides.

coarse-grained model | Bayesian analysis | folding | SAXS | simulation

Tyrosine kinases of the Src-family are large multidomain allosteric enzymes implicated in the signaling pathways regulating cell growth and proliferation (1). The key role that Src kinases play in the onset of many human diseases, particularly cancer, makes them important targets for therapeutic intervention (2).

All Src kinases share a common structural organization comprising the SH3 and SH2 binding domains followed by a highly conserved catalytic domain connected by flexible linkers (1). Crystal structures have offered a detailed view of the down-regulated forms of Hck and c-Src (3, 4), characterized by a compact assembled form stabilized by auto-inhibitory intramolecular interactions between the N and C terminus of the catalytic domain with the SH3 and SH2 modules, respectively. The SH2 and SH3 modules of the Src-family tyrosine kinases play dual roles: both are involved in the auto-inhibitory intramolecular interactions down-regulating kinase activity, but can also serve as possible binding receptors for cellular signals leading to kinase activation.

In a broadly accepted paradigm, Src inactivation/activation is pictured in terms of a simple two-state process involving an assembled (inactive) and a disassembled (active) conformation, i.e., once any of the intramolecular interaction is released, the kinase domain switches to a disassembled catalytically active state targeting down-stream cellular substrates (5–8). Crystal structures (3, 4) and mutational studies complemented by molecular dynamics (MD) simulations (9) are consistent with this view. However, one crystal structure of c-Src in a partially activated state in which the SH2-SH3 modules are reassembled in a different orientation with respect to the catalytic domain (10), suggests that the situation could be considerably more complex. In this regard, it seems likely that different signals could promote different assembly states, leading to differently configured activated kinases. While it is understood that Src increases its activity in response to molecular signals that interfere with the interactions stabilizing the assembled complex, little is known about the spatial organization of the up-regulated active form of Src. A quantitative characterization of the structural organization of the multidomain Src in solution under different conditions is critical

for a better understanding of the molecular basis of the allosteric response.

Probing the spatial organization of a large and flexible multidomain macromolecular complex such as Src is challenging. The protein is expected to display substantial conformational dynamics, giving rise to a large number of possible conformations separated by small energy differences. Capturing any of those in a crystallographic state would be inherently difficult, and ideally one would like to observe the protein conformation in solution, where it is not affected by lattice packing. In principle, small-angle X-ray solution scattering (SAXS) is an experimental technique that can be used for mapping the three-dimensional organization of multidomain proteins in solution (11–15). Different conformations give rise to distinct scattering patterns containing information on the overall architecture and domain-domain separation within the complex. However, it is not straightforward to determine the three-dimensional shape directly from the low-resolution information provided by the SAXS data. More importantly, there are fundamental limitations associated with the naive notion of a unique molecular shape in the case of flexible multidomain complexes, which may display considerable conformational dynamics.

To address those issues, we present an integrative approach, the basis-set supported SAXS (BSS-SAXS) reconstruction, that aims at combining SAXS data with a simplified coarse-grained (CG) model to characterize the assembly states of the flexible multidomain Hck in solution. Simulations of the CG model are used to extensively explore and sample the accessible conformational space of the multidomain complex. The configurations are clustered into a small number of distinct putative assembly states, which are then used as a basis-set to analyze SAXS data. Comparison of the calculated SAXS patterns from all assembly states with experimental data makes it possible to determine the population fraction of each state of Hck under various conditions. The results indicate that multiple assembly states are generally present, and that the population equilibrium among these states responds to both mutations and the presence of signaling peptides binding to the SH2 or SH3 domains.

Results

Classes of Hck Assembly States. The simplified CG model of Hck, though extremely simplified, incorporates the main generic features of the protein: (*i*) the native three-dimensional conformations of the SH3, SH2, and catalytic domains from the X-ray

Author contributions: S.Y. and B.R. designed research; S.Y. performed research; S.Y. and L.B. contributed new reagents/analytic tools; S.Y. analyzed data; and S.Y., L.M., L.B., and B.R. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

¹To whom correspondence should be addressed. E-mail: roux@uchicago.edu.

²Present address: Center for Proteomics and Bioinformatics and Department of Pharmacology, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1004569107/DCSupplemental.

structures, (ii) the topological connectivity of the polypeptide chain via flexible linkers, and (iii) the excluded volume from the residue-residue core repulsion. The CG model also accounts for the intramolecular interactions involved in the stabilization of the down-regulated assembled form (3): the tyrosine-containing C-terminal “Ctail” of the kinase domain with the SH2 domain, and the polyproline SH2-kinase “linker” with the SH3 domain. As described in *SI Appendix*, both interactions were incorporated in the CG model, and adjusted to match the K_d from experiments for the SH2-binding Ctail peptide, and the SH3-binding polyproline peptide. The active and inactive conformations of the catalytic domain are also included in the multistate model as described previously (16, 17) to account for the ability of full-length Hck to adopt different conformations. To ensure proper sampling of various assembly conformations, simulations using different initial conditions were conducted, ranging from fully assembled to disassembled states, from increased to decreased binding interactions, and from destabilizing stable configurations to lowered transition barriers.

A large number of configurations were generated using MD simulations based on the CG model. A two-step clustering scheme was used to organize the large amount of information generated by the simulations of the CG model into a manageable form. Initially, 25 structural clusters were determined from the trajectories from a structural clustering based on residue-residue distances similarity criterion (16, 17). More detail on the clustering is given in *SI Appendix: Fig. S1*. Then, average scattering patterns were calculated for members of each cluster using the Fast-SAXS method (18), and the clusters were further regrouped into nine assembly states on the basis of a scattering pattern similarity criterion. The relationship between the nine states according to this metric is shown diagrammatically in *SI Appendix: Fig. S1*, and a representative conformation from each state is shown in Fig. 1. It should be noted that the configurations within those states are regrouped according to the similarity of their scattering patterns, and that they may display conformational differences. This set of nine assembly states, extracted from the simplified CG simulations, is then used as a basis-set for extracting the multidomain organization of Hck in solution from the experimentally observed SAXS patterns.

Bayesian-Based Monte Carlo Analysis of SAXS Data. The calculated profile I_{calc} is an average over the $N_s = 9$ states, weighted by the fractional population $\{P_i\} \equiv P_1, \dots, P_{N_s}$ of state i ,

$$I_{\text{calc}}(q) = \sum_{i=1}^{N_s} P_i I_i(q), \quad [1]$$

where $I_i(q)$ is the theoretical scattering profile of state i (see examples in *SI Appendix: Fig. S2*). The goal of the analysis is to determine the set of N_s fractional population $\{P_i\}$ from the SAXS data. To this end, we define a dimensionless scoring function, $\chi^2[\{P_i\}]$, measuring the agreement between the calculated and experimental SAXS profiles,

$$\chi^2[\{P_i\}] = \sum_{q_{\min}}^{q_{\max}} \frac{1}{\sigma^2(q)} (\log I_{\text{calc}}(q) - \log I_{\text{exp}}(q) - \Delta)^2, \quad [2]$$

where q_{\min} and q_{\max} are the lower and upper limit of q -range of the experimental profile of $I_{\text{exp}}(q)$, respectively, and Δ is the offset between $\log I_{\text{calc}}$ and $\log I_{\text{exp}}(q)$ at $q = q_{\min}$. The estimated experimental SAXS uncertainty, $\sigma(q)$, is assumed to be $[\delta \log I_{\text{exp}}(q)]^{1/2}$. Assuming independent experimental errors and a uniform prior, the probability distribution $\exp(-\chi^2)$ is sampled in the space of the model parameters, $\{P_i\}$, according to a Bayesian-Based Monte Carlo (BMC) algorithm. The procedure yields the fractional population P_i for the N_s states, along with their uncertainties extracted the fluctuations.

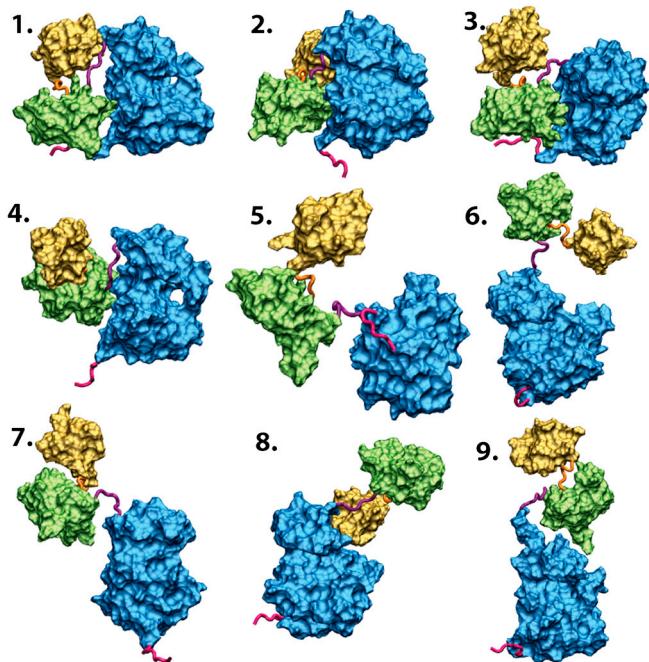


Fig. 1. Multiple assembly conformational states adopted by the multidomain Hck from a CG model. Representatives of these nine scattering states, ranging in architecture from fully to partially assembled to disassembled states, in size from compact to extended forms, and in interdomain separation from assembled to fully or partially disassembled states. The catalytic domain is in blue, SH2 in green, and SH3 in yellow. States 1 and 8 are similar to the assembled Hck (pdb entry 1QCF) and the partially active c-Src (pdb entry 1Y57) structures, respectively. More information about the states is in *SI Appendix: Figs. S4–S12*.

Tests with the BSS-SAXS method indicate that the scattering pattern in the low- q regions, up to about 0.15 \AA^{-1} contains most of the information that controls the structural differences in Hck (see *SI Appendix: Fig. S3A, C*). Including higher- q SAXS data has little effect on the average values of P_i , although it reduces the uncertainty of the P_i (see *SI Appendix: Fig. S3B, D*). Additional tests with the BSS-SAXS/BMC simulations show that the procedure yields the correct population P_i (see *SI Appendix: Fig. S3E, F*), although large increases in the estimated SAXS uncertainty can lead to inaccuracies (see *SI Appendix: Fig. S3G, H*). The present approach has some similarities with other methods, such as EOM (19) and BILBOMD (20), which were also designed to go beyond standard strategies based on a single conformation to fit SAXS data. For example, all three methods primarily rely on a large set of random conformers. Both EOM and BILBOMD use a Monte Carlo genetic algorithm to select a small number of specific conformers out of the large set to best-fit the SAXS data, which are attributed equal statistical weight, whereas BSS-SAXS uses a BMC algorithm to determine the fractional population P_i for the states in a scattering basis-set. Overall, the tests show that the BSS-SAXS analysis can provide quantitative estimates of the fractional population P_i of Hck in solution.

BSS-SAXS Characterization of Hck Assembly. SAXS data were collected for the wild-type Hck, and for the high-affinity Ctail mutant (Hck-YEEI) in the presence of two types of external peptide ligands: (p2) a high-affinity phosphorylated tyrosine-containing SH2-binding peptide (6); (p3) a high-affinity SH3-binding PPII peptide (21 and 22). All the experimental scattering profiles are shown in Figs. 2, 3. For the sake of completeness, a simple Guinier analysis was performed for all the SAXS data shown in Figs. 2 and 3, and the radii of gyration R_g were extracted by fitting the scattering patterns in the range up to 0.05 \AA^{-1} . The results are reported in the Figure captions.

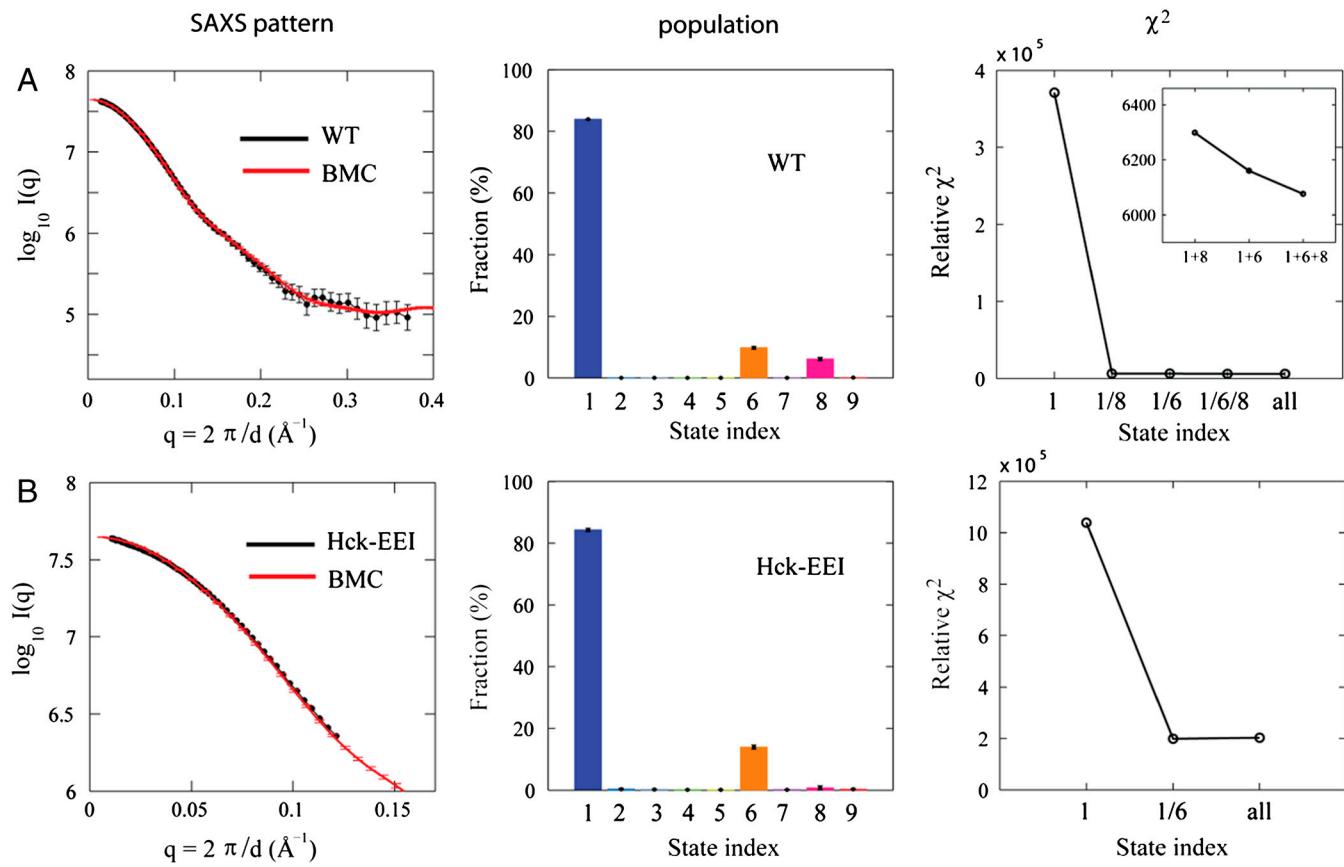


Fig. 2. Analysis of wild-type Hck and high-affinity Ctail mutant (Hck-YEEI) in solution. The scattering pattern (left), the assembly state population (center), and the relative χ^2 scores for different combinations (right) are shown. (A) SAXS data for the wild-type Hck (black) and BMC simulated scattering profile (red) are shown. R_g is 28.1 ± 0.2 Å. All the SAXS profiles of $I(q)$ are shown in a logarithmic scale, where $q = 2\pi/d$ and d is the Bragg spacing in scattering. Relative χ^2 scores show that the inclusion of states 6 and 8 has improved the fit to the SAXS data. The relative χ^2 score for the crystal structure is 7.6×10^7 . (B) SAXS data for the Hck-EEI mutant (black) and BMC simulated scattering (red) are shown. R_g is 27.1 ± 0.2 Å. Inclusion of state 6 greatly improves agreement with the SAXS data (right).

The experimental SAXS data from wild-type Hck in the absence of external peptides is shown in Fig. 2A. It is observed that the SAXS profile calculated from the population derived from the BMC analysis (red) accurately reproduces the experimental data (black). The fit is achieved with the population of assembly states determined via the BMC analysis. The latter indicates that the enzyme is predominantly in the assembled state 1, with the population of 82%. The more weakly populated states include the fully disassembled state 6 and the partially disassembled state 8 with SH2-dissociated but SH3-bound similar to the partially disassembled crystal structure of c-Src (10). As shown in Fig. 2A, this combination of states best fits the SAXS data. The experimental SAXS data from the high-affinity Ctail mutant YEEI is shown in Fig. 2B. In this case too, the assembled state 1 dominates the three-dimensional organization, with the population of 83%.

To investigate how ligand binding affects the equilibrium assembly of the multidomain complex, SAXS data were also collected from Hck with two types of peptide ligands. As shown in Fig. 3A, there is a substantial shift of wild-type Hck from a compact state to disassembled states in the presence of 250 μ M of the high-affinity phosphorylated SH2-binding peptide p2. Comparison of Figs. 2A and 3A shows that the population of the compact state 1 decreases from 82% to 22%, while that of the disassembled state 6 increases from 8% to 29%. Repeating this experiment for the Hck-YEEI Ctail mutant yields a very different result. The BMC analysis given in Fig. 3B indicates that, for the Hck-YEEI mutant, the population of the assembled state is reduced only to 62% upon addition of the peptide p2. Here, the disassembled and compact state 6, where SH2 and SH3

are disassociated but the overall Hck architecture remains compact, is also populated.

The effect of the SH3-binding peptide p3 is shown in Fig. 3C. BMC analysis of the SAXS data indicates that p3 destabilizes the assembled down-regulated state 1 and promotes the disassembled compact state 6. Comparison of Figs. 2A and 3C shows that the population of state 1 decreases from 82% to 7% and that of state 6 increases from 8% to 50% in the presence of peptide p3. Also, the population of the SH3-displaced states 2 and 3 increases, reaching a total combined population of about 39%. Cross-correlation analysis of the population of states 2 and 3, however, suggests that they are highly correlated and that determining their relative weight is at the limit of resolution of the present BMC analysis.

Lastly, the spatial organization of Hck was examined in the presence of both the p2 and p3 peptides simultaneously. The results of the BMC analysis of SAXS data, shown in Fig. 3D, demonstrate that any remnant of the assembled down-regulated state is wiped out upon the binding of p2 and p3. Here, the fully disassembled states 5 and 6 are representative of the spatial organization. Both states are disassembled when Ctail is released from SH2 and SH3 is disassociated, but they differ in overall size: state 5 is extended (*SI Appendix: Fig. S8*) and state 6 exhibits a compact form (*SI Appendix: Fig. S9*). Consistent with the BSS-SAXS analysis, R_g reaches its largest value (31.7 Å) under these conditions.

Discussion

The SAXS data is consistent with an increasing disassembled state for Hck as it is perturbed by the signaling peptides. In

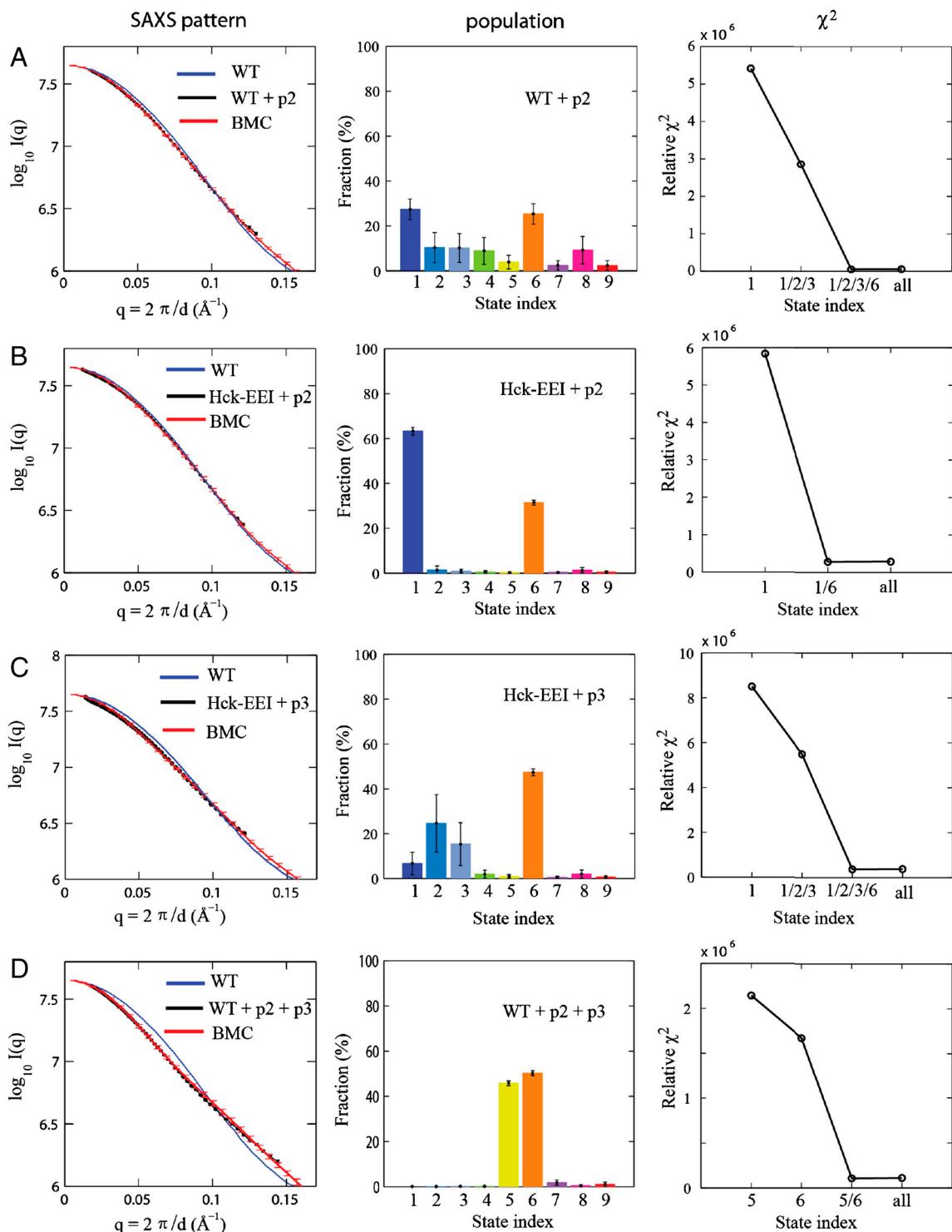


Fig. 3. The wild-type and mutant Hck in solution in the absence and presence of SH2-binding (p2) and SH3-binding (p3) peptides. In all cases, the scattering pattern (left), the assembly state population (center), and the relative χ^2 scores for different combinations (right) are shown. (A) SAXS data of Hck in the presence of p2 at a concentration of 250 μM (black) and the BMC simulated profile (red) are shown. Addition of p2 causes a shift in the equilibrium toward a coexistence of assembled and various disassembled states. R_g is $29.3 \pm 0.2 \text{ \AA}$. Relative χ^2 scores for different solutions of the state combination. (B) SAXS data of Hck-YEEI in the presence of p2 at a concentration of 250 μM (black) and the BMC simulated profile (red) are shown. R_g is $28.1 \pm 0.1 \text{ \AA}$. Furthermore, in contrast to the Hck-YEEI scattering in the absence of p2 (Fig. 2B), these results demonstrate that the addition of p2 to a Hck-YEEI solution causes an increase in population of the disassembled and compact state 6. (C) SAXS data of Hck-YEEI in the presence of p3 peptide at a concentration of 1 mM (black) and the BMC simulated profile (red) are shown. R_g is $29.0 \pm 0.3 \text{ \AA}$. (D) Addition of both p2 and p3 peptides shifts the equilibrium of Hck conformations toward disassembled states. SAXS data of the wild-type Hck in the presence of both p2 at a concentration of 250 μM and p3 at a concentration of 1 mM (black) and the BMC simulated profile (red). R_g is $31.7 \pm 0.2 \text{ \AA}$.

particular, the radius of gyration R_g extracted from the SAXS data via a Guinier analysis is typically smaller in the absence

of binding peptides (Fig. 2) than in their presence (Fig. 3). Also consistent with this result, the smallest value of 27.1 \AA is observed

for the EEI Ctail mutant of Hck (Fig. 2B), whereas the largest value of 31.7 Å is observed for the wild-type Hck in the presence of peptides p2 and p3 (Fig. 3D). Nevertheless, although a simple analysis of SAXS data based on the variations of R_g is indicative of broad structural changes, the scattering pattern does not, by itself, provide a straightforward characterization of the assembly of a complex multidomain protein such as Hck.

To determine quantitatively how various factors can affect the configurational landscape of Hck, we set out to estimate the relative population of a set of possible assembly states from the experimental SAXS data using a BMC analysis. This analysis leans on the configurations generated by simulating a simplified CG model. The set of nine assembly states extracted from those simulations serves as a basis-set covering the wide range of compact and loose configurations that Hck could possibly adopt. In the BSS-SAXS reconstruction, the set of nine assembly states constitutes a Bayesian prior in the analysis (23). One configuration from each of the assembly states is illustrated in Fig. 1. However, it is emphasized that each state actually comprises a large ensemble of configurations (SI Appendix: Figs. S4–S12). Such conformational heterogeneity is consistent with the existence of a broad continuum of accessible configurations, but it could also reflect our limited ability to resolve the conformation of a multidomain protein unambiguously using only the information provided by the scattering patterns. Without additional experimental data permitting a more detailed structural description, it is best to exert some caution when interpreting the results of the BSS-SAXS analysis in terms of “state” population.

The nine states in Fig. 1 can be resolved from SAXS data, i.e., differences in scattering intensity are larger than experimental uncertainty. Distinguishability of the states included in the set is measured from the uncertainty-weighted difference in average SAXS patterns $\langle I(q) \rangle$ (see also SI Appendix: Table S1). The set includes a wide range of different configurations, varying in their multidomain organization and overall architecture. State 1 is similar to the X-ray structure of Hck corresponding to the fully assembled down-regulated kinase (3, 4), while state 8 is analogous to the partially activated state observed in one X-ray structure of c-Src (10). Fluctuations are important and there are subtle, though important, differences among some states. For example, states 5 and 6 are both disassembled, but the former is more extended than the latter. State 2 and 3 correspond to SH3-displaced configurations, lacking the direct interaction between the SH3 and the PPII linker. In some disassembled configurations, the SH3 and SH2 domains maintain the relative orientation observed in the assembled form (state 1), consistent with a previous simulation study (24).

The result of the BMC analysis of SAXS data for the wild-type Hck is shown in Fig. 2A. The analysis indicates that the enzyme is predominantly in the assembled state 1, with a probability of 82%. As shown in the figure (left), the calculated scattering profile with the combination of states determined from the BMC analysis (red curve) is in excellent agreement with the SAXS data (black curve). Although the assembly state 1 is similar to the X-ray structure of Hck (3, 4), the ensemble of configurations includes dynamical fluctuations with rmsd on the order of 2–3 Å (see SI Appendix: Fig. S4). The scattering pattern calculated from the static X-ray structure alone does not fit the SAXS data accurately. In fact, the assembly state 1 (with dynamical fluctuations) also yields a poor score (Fig. 2A, right). The contribution from assembly states 6 and 8 is critical and improves the goodness of fit. In a previous SAXS study of c-Src, it was concluded that about 83% of the enzymes were in a conformation similar to the assembled form (25). In those experiments, Tyr₅₂₇ in the C terminus of the catalytic domain was phosphorylated by the kinase Csk, which is required for kinase inactivation. It is particularly noteworthy that the assembled form is predominant, even though Tyr₅₂₇ is unphosphorylated in the present

experiments. The interaction of the phosphorylated tyrosine with the SH2 module is critical to stabilize the assembled form of the kinase (3, 4). It is well known that dephosphorylation of Tyr₅₂₇ leads to kinase activation and truncation of the C terminus leads to a constitutively active kinase. Yet, the present results indicate that the assembled conformation is also predominant in solution (84%), even when the tyrosine in the Ctail is unphosphorylated. Thus, despite the expectation that Hck with an unphosphorylated Tyr₅₂₇ in the Ctail might adopt a multiplicity of disassembled states, the assembled conformation typical of the inactive kinase is predominant. The present observations challenge the notion that dephosphorylation of the Ctail immediately triggers a large conformational change toward a disassembled state leading to kinase activation.

Comparison of the results obtained on the wild-type and mutant kinase sheds an interesting light on the competition between the binding of the Ctail and an external p2 peptide, corresponding intramolecular and bimolecular association processes, respectively. The experimental SAXS data from the high-affinity Ctail mutant Hck-YEEI is shown in Fig. 2B. Here as well, the assembled state 1 dominates the three-dimensional organization. The lack of increase in the population of assembly state 1 is somewhat surprising because the high-affinity Ctail mutant is known to down-regulate the kinase (6), although a Guinier analysis indicates that R_g shifts from 28.1 Å to 27.1 Å (Fig. 2A and B). As shown in Fig. 3A and B, the presence of the SH2-binding peptide p2 destabilizes the assembled down-regulated state, the effects being more important for the wild-type Hck than for the Hck-YEEI Ctail mutant. The population of assembly state 1 is reduced to 22% for the wild-type Hck, while it is reduced only to 62% for the Hck-YEEI mutant. This destabilization is also reflected in the R_g of 29.3 Å and 28.1 Å, for the wild-type and Hck-YEEI mutant, respectively. The shift toward disassembled configurations is consistent with the increase in kinase activity observed in the presence of SH2-binding ligands (7). Nevertheless, the nonnegligible population of assembly state 1 remaining in the presence of high-affinity SH2-binding peptide is somewhat surprising. The peptide comprises a phosphorylated tyrosine and is present at a nearly saturating concentration of 250 μM. In spite of this saturation, the intramolecular SH2-Ctail interactions for both the Hck-YEEI mutant and, to a lesser extent, wild-type Hck compete against the binding of a very high-affinity peptide and protect the assembly state 1. The observation can be rationalized in terms of a simple equilibrium model between an assembled state A and a disassembled state D, with the assumption that only the latter can bind the external ligand. The equilibrium for the three state process, A ⇌ D + L ⇌ DL, is governed by $[D]/[A] = p$, and $[DL]/[D] = [L]/K_d$, where p is the natural propensity of occurrence of the D state, and K_d is the dissociation constant, respectively. The probability of the assembled state is then equal to, $P_A = (1 + p + p[L]/K_d)^{-1}$. Because p is smaller for the high-affinity Ctail mutant Hck-YEEI than for the wild-type Hck, the intramolecular interaction is able to override the strong bimolecular association ($p_{YEEI}[L]/K_d \ll p_{wt}[L]/K_d$), thus protecting the assembled down-regulated conformation.

The action of the SH3-binding peptide p3 is meant to mimic a signal such as produced by the Nef PxxP motif, which was previously used to induce kinase activation by releasing the SH3 module from the PPII linker (26). The analysis of the SAXS data for the Hck-YEEI mutant in the presence of p3 shown in Fig. 3C indicates that the binding of p3 increases the population of assembly states 2 and 3, corresponding to two possible SH3-displaced orientations (see also SI Appendix: Fig. S5 and Fig. S6). A number of in vitro and in vivo data including cell based assays and mutagenesis have also supported the concept of such an SH3-displaced model in kinase activation (5, 27). While the present results are clearly at the limit of resolution available from the

analysis of SAXS data, they support the concept of SH3-displaced states.

The classic view of Src activation has typically invoked some loose disassembled conformation, in which all autoinhibitory intramolecular interactions are released. The analysis of the SAXS data for the wild-type Hck mutant in the presence of both the p2 and p3 peptides shown in Fig. 3D yields a representation of such a state of Hck in solution. Of particular interest, the assembled state 1 is wiped out while the most populated states, 5 and 6, differ from those detected under the previous conditions. This result demonstrates that the multidomain protein has access to a wide basin of different conformations and that different conditions can shift the population of the states in various directions.

Conclusion

An integrated framework combining SAXS experiments and computational modeling was introduced to quantitatively characterize the assembly of Hck kinase in solution. The BSS-SAXS analysis provides a view of the rich possibilities of three-dimensional conformational organizations accessible to Src in solution, going beyond the down-regulated assembled or fully disassembled states, which may be associated with kinase activation. The method offers a complementary and powerful approach to characterize multidomain molecular assemblies in solution, especially when multiple conformations can coexist. An important finding from the analysis is that multiple assembly states are required to construct a proper basis-set covering the conformational landscape of Hck. Available crystal structures are not enough to match the experimental SAXS data. Of particular interest, a large fraction of the kinase conformations exist in the assembled down-regulated state, even when the Tyr₅₂₇ in the C terminus of the catalytic domain is not phosphorylated. Furthermore, different external signals lead to different distributions among the possible assembly states. These observations point to a molecular interpretation of the allosteric response of Hck to various biochemical signals that is in accord with the general concept of a shift in equilibrium populations (28, 29).

1. Martin GS (2001) The hunting of the Src. *Nat Rev Mol Cell Biol* 2:467–475.
2. Noble MEM, Endicott JA, Johnson LN (2004) Protein kinase inhibitors: insights into drug design from structure. *Science* 303:1800–1805.
3. Sicheri F, Moarefi I, Kuriyan J (1997) Crystal structure of the Src family tyrosine kinase Hck. *Nature* 385:602–609.
4. Schindler T, et al. (1999) Crystal structure of Hck in complex with a Src family-selective tyrosine kinase inhibitor. *Mol Cell* 3:639–648.
5. Moarefi I, et al. (1997) Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement. *Nature* 385:650–653.
6. Porter M, Schindler T, Kuriyan J, Miller WT (2000) Reciprocal regulation of Hck activity by phosphorylation of tyr527 and tyr416. effect of introducing a high affinity intramolecular SH2 ligand. *J Biol Chem* 275:2721–2726.
7. Goldberg GS, et al. (2003) Src phosphorylates cas on tyrosine 253 to promote migration of transformed cells. *J Biol Chem* 278:46533–46540.
8. Miller W (2003) Determinants of substrate recognition in nonreceptor tyrosine kinases. *Acc Chem Res* 36:393–400.
9. Young MA, Gonfloni S, Superti-Furga G, Roux B, Kuriyan J (2001) Dynamic coupling between the SH2 and SH3 domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation. *Cell* 105:115–126.
10. Cowan-Jacob SW, et al. (2005) The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* 13:861–871.
11. Koch MHJ, Vachette P, Svergun DI (2003) Small-angle scattering: a view on the properties, structures, and structural changes of biological macromolecules in solution. *Q Rev Biophys* 36:147–227.
12. Putnam CD, Hammel M, Hura GL, Tainer JA (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations, and assemblies in solution. *Q Rev Biophys* 40:191–285.
13. Lipfert J, Doniach S (2007) Small-angle X-ray scattering from RNA, proteins, and protein complexes. *Annu Rev Biophys Biomol* 36:307–327.
14. Forster F, et al. (2008) Integration of small-angle X-ray scattering data into structural modeling of proteins and their assemblies. *J Mol Biol* 382:1089–1106.
15. Hura GL, et al. (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat Methods* 6:606–612.
16. Yang S, Roux B (2008) Src kinase conformational activation: thermodynamics, pathways, and mechanisms. *PLoS Comput Biol* 4:e1000047.
17. Yang S, Banavali NK, Roux B (2009) Mapping the conformational transition in Src activation by cumulating the information from multiple molecular dynamics trajectories. *Proc Natl Acad Sci USA* 106:3776–3781.
18. Yang S, Park S, Makowski L, Roux B (2009) A rapid coarse residue-based computational method for X-ray solution scattering characterization of protein folds and multiple conformational states of large protein complexes. *Biophys J* 96:4449–4463.
19. Bernadó P, Mylonas E, Petoukhov MV, Blackledge M, Svergun DI (2007) Structural characterization of flexible proteins using Small-Angle X-ray scattering. *J Am Chem Soc* 129:5656–5664.
20. Pelikan M, Hura GL, Hammel M (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering. *Gen Physiol Biophys* 28:174–189.
21. Zarrinpar A, Bhattacharya RP, Lim WA (2003) The structure and function of proline recognition domains. *Science's STKE* 2003:1–10.
22. Lerner EC, et al. (2005) Activation of the Src family kinase Hck without SH3-linker release. *J Biol Chem* 280:40832–40837.
23. Sriraman S, Kevrekidis IG, Hummer G (2005) Coarse master equation from bayesian analysis of replica molecular dynamics simulations. *J Phys Chem B* 109:6479–6484.
24. Faraldo-Gomez J, Roux B (2007) On the importance of a funneled energy landscape for the assembly and regulation of multidomain Src tyrosine kinases. *Proc Natl Acad Sci USA* 104:13643–13648.
25. Bernadó P, Perez Y, Svergun DI, Pons M (2008) Structural characterization of the active and inactive states of Src kinase in solution by small-angle X-ray scattering. *J Mol Biol* 376:492–505.
26. Grzesiek S, et al. (1996) The solution structure of HIV-1 nef reveals an unexpected fold and permits delineation of the binding surface for the SH3 domain of Hck tyrosine protein kinase. *Nat Struct Biol* 3:340–345.
27. Lerner EC, Smithgall TE (2002) SH3-dependent stimulation of Src-family kinase autop phosphorylation without tail release from the SH2 domain in vivo. *Nat Struct Biol* 9:365–369.
28. Vendruscolo M, Dobson CM (2006) Structural biology: dynamic visions of enzymatic reactions. *Science* 313:1586–1587.
29. Henzler-Wildman K, Kern D (2007) Dynamic personalities of proteins. *Nature* 450:964–972.
30. Seeliger MA, et al. (2005) High yield bacterial expression of active c-Abl and c-Src tyrosine kinases. *Protein Sci* 14:3135–3139.
31. Yang S, Onuchic JN, Levine H (2006) Effective stochastic dynamics on a protein folding energy landscape. *J Chem Phys* 125:054910.
32. Best RB, Chen YG, Hummer G (2005) Slow protein conformational dynamics from multiple experimental structures: the helix/sheet transition of arc repressor. *Structure* 13:1755–1763.

Methods

Experiments. Wild-type HCK3D or the HCK3D.YEEI mutant were coexpressed in bacteria with YopH phosphatase and purified as described by Seeliger et al. (30). The SH2 binding peptide (referred to as p2; sequence EPQ[pY]EEPIKQ, where pY is phosphotyrosine) and the SH3 binding peptide (referred to as p3; sequence VSLARRPLPPLP) were synthesized in house. SAXS data were collected at the BioCAT-18ID from the Advanced Photon Source. Data were collected using a flow-cell kept at 10 °C with a series of 1 s exposures for the Hck protein sample, the buffer, and the empty capillary. SAXS profiles were determined by the subtraction of the buffer and the empty capillary scattering from the protein sample scattering. See more details in *SI Appendix*.

Multidomain Modeling of Hck. Extensive configurational sampling was carried out using a multistate CG model of Hck following methods similar to previous studies (16, 31, 32). Models of inactive and active states of the catalytic domain were built from the crystal structures of Hck (PDB entry: 1QCF) and c-Src (PDB entry: 1Y57), respectively (4, 10), and models of the SH3 and SH2 were taken from the inactive state of the down-regulated Hck (4). The models are built at the residue level, where each residue is represented by its Cα atom. For postdata analysis, a clustering and mapping method as described previously was applied to simulation data (17). The resulting configurations were clustered into 25 structural clusters, where a list of 89 interdomain pair wise residue distances were used in a K-means clustering calculation. For each cluster, a number of 100 randomly selected configurations was used to compute the theoretical scattering profiles using the Fast-SAXS method (18), yielding 25 profiles (and their standard deviations) $\log I_i(q) \pm \delta \log I_i(q)$ where $i = 1, \dots, 25$. A second level K-means clustering procedure was performed using a metric based on the similarity between calculated scattering patterns. This second step assigns the 25 structural clusters to $N_s = 9$ scattering classes or states (see *SI Appendix: Fig. S1*). See the *SI Appendix* for more details.

ACKNOWLEDGMENTS. We thank Markus Seeliger and John Kuriyan for providing the pHCK3D and YopH Duet plasmids and Richard Jones for synthesizing p2 and p3 peptides. SAXS experiments were carried out at the BioCAT of the Argonne National Laboratory with the help of Dr. Liang Guo. We thank Alper Dagcan for helping with SAXS experiments. This work was supported by the National Institutes of Health via Grant CA-093577 from the National Cancer Institute. The computations were supported by Teragrid under project MCA01S018 and by the San Diego Supercomputer Center.

Supplementary Information for “Assembled States of Hck Tyrosine Kinase in Solution”

Plasmids

The pHCK3D and YopH Duet plasmids were provided by Markus Seeliger and John Kuriyan. pHCK3D is a pET-28a vector (Novagen) modified to yield a tobacco etch virus (TEV) protease cleavable N-terminal hexahistidine tag. It contains a fragment of human HCK spanning the SH3, SH2, and catalytic domains, and the C-terminal tail (residues 85-531, chicken c-Src numbering). YopH Duet is a pCDFDuet-1 vector (Novagen) containing full-length YopH phosphatase from *Yersinia*. The pHCK3D.YEEI mutant was generated using the QuikChange Site Directed Mutagenesis kit (Stratagene) with pHCK3D as a template and mutagenic primers from Integrated DNA Technologies. This construct contains a high affinity C-terminal tail sequence (E₅₂₄SQYEEIP₅₃₁).

Protein expression and purification

Wild type HCK3D or the HCK3D.YEEI mutant were co-expressed in bacteria with YopH phosphatase and purified as described by Seeliger *et al*¹. The two plasmids containing the kinase and the phosphatase were co-transformed into *Escherichia coli* BL21(DE3) cells. 1-liter cultures were grown in Terrific Broth with 50 µg/ml each of kanamycin and streptomycin at 37°C to an OD₆₀₀ of 1.0, cooled for 1 hour with shaking at 18°C, and induced overnight at 18°C with 1mM IPTG. Cells were harvested by centrifugation at 4000 g for 10 minutes at 4 °C. Pellets were resuspended in 30 ml of 20 mM Tris (pH 8.0), 500 mM NaCl, 5% glycerol, 20 mM imidazole (NiA buffer) and

lysed by 4 cycles of homogenization at 15,000 psi using an Avestin homogenizer. The lysate was cleared by centrifugation at 125,000 g for 50 minutes at 4 °C. The supernatant was loaded onto 5 ml NiNTA resin (Qiagen) equilibrated in NiA buffer. After washing with 10 column volumes of NiA buffer, the protein was eluted with 3 column volumes of NiB buffer (NiA buffer plus 500 mM imidazole). The eluted protein was cleaved with 1 mg TEV per 25 mg kinase at 4 °C overnight while dialyzing against 2 liters of 20 mM Tris (pH 8.0), 100 mM NaCl, 5% glycerol, 0.5 mM EDTA, 0.5 mM TCEP in a 12-14 kDa molecular weight cutoff membrane. The dialyzed protein was diluted 1.3 fold with water, spun at 14,000 g for 10 minutes at 4°C, and loaded onto an ion exchange column (HiTrap Q FF, GE Lifescience) equilibrated with 20 mM Tris (pH 8.0), 0.5 mM TCEP (QA buffer). The kinase was eluted with a linear gradient of 0-50% QB buffer (QA buffer plus 1 M NaCl). The peak fractions were analyzed by SDS-PAGE and the fractions containing kinase were pooled. A 1.5 fold molar excess of the inhibitor PP1 (Tocris) was added to prevent aggregation and the pooled fractions were concentrated using an Amicon Ultra-15 centrifugal filter device with a 30-kDa molecular weight cutoff membrane. The concentrated sample was loaded onto a size-exclusion column (Superdex 200 10/300 GL, GE Lifecscience) equilibrated with 50 mM Tris (pH 8.0), 100 mM NaCl, and 0.5 mM TCEP. HCK was eluted at the volume expected for monomeric protein. No aggregation was detected in the void volume of the column. Sample concentrations were adjusted to 1 mg/ml for SAXS data collection.

Peptides

Peptides were synthesized in house. The SH2 binding peptide (referred to as p2) has the sequence EPQ[pY]EEIPIKQ (where pY is phosphotyrosine). The SH3 binding peptide (referred to as p3) has the sequence VSLARRPLPPLP. Stock solutions of p2 (800 μ M in Superdex column buffer) and p3 (8 mM in 50% DMSO) were stored in aliquots at -80°C until needed for SAXS data collection.

SAXS data collection

SAXS data were collected at the BioCAT-18ID from the Advanced Photon Source (APS), Argonne, IL. A flow-cell was used to deliver continuous flow through the capillary during data collection to reduce the radiation damage. The temperature of the capillary and sample cell was kept at around 10 °C. Data were collected with a series of 1-second exposures for the Hck protein sample, the buffer and the empty capillary. SAXS profiles were determined by the subtraction of the buffer and the empty capillary scattering from the protein sample scattering.

Modeling for the multi-domain assembly of Hck

To explore multiple conformations adopted by the multi-domain Hck complex, extensive sampling and modeling were carried out as follows. For convenience, three segments (Ile₈₂-Ser₁₄₂, Glu₁₄₄-Met₂₄₆, and Glu₂₅₅-Thr₅₂₃) in the human Hck are referred to as SH3, SH2, and catalytic-domain, respectively. For inter-domain interactions, the association was calibrated according to disassocia-

tion constants (K_d) between SH2 and a peptide with sequence of ESQYQQQP, and between SH3 and a peptide with sequence of SSKPQKWP. The former peptide resembles the $E_{524}SQYQQQP_{531}$ segment in the C-terminal tail of Hck, while the latter resembles the $S_{247}SKPQKWP_{254}$ sequence in the SH2-connecting polyproline linker region. The reported K_d 's are in the μM and tens of μM ranges for SH2-p2 and SH3-p3²⁻⁴, respectively. To implement these inter-domain interactions, umbrella sampling with harmonic biasing potentials was first performed to obtain a potential of mean force (PMF) of $\mathcal{W}(r)$ (where r are inter-residue distances between Asn₁₃₅ in SH3 and Gln₂₅₁ as in the former peptide and between Tyr₂₀₂ in SH2 and Tyr₅₂₇ as in the latter peptide). From $\mathcal{W}(r)$, a theoretical K_d was then calculated for the association between each domain and its binding peptide, $K_d = 1 / (4\pi \int r^2 e^{-\mathcal{W}(r)/k_B T} dr)$, where k_B is the Boltzmann factor and T is the system temperature. The resulting K_d 's of 1.2 μM and 27 μM , respectively, were implemented for the intramolecular interactions via the association with SH2 and SH3. To account for competitive binding of external peptides, these intramolecular interactions were also allowed to be on/off in the model. For example, when the interactions of SH2 and the C-tail are turned off, the C-tail is allowed to release from SH2 so that the binding interface of SH2 is available for an external SH2 ligand to bind to the complex. In addition, the SH2-tail interactions are also strengthened to account for its increased binding in a high-affinity mutant with a sequence of $E_{524}SQYEEIP_{531}$. In this case, a value of K_d of 0.56 μM between the SH2 and the high-affinity YEEI-containing peptide was used for the SH2 association in the complex.

For the catalytic domain, a multi-state coarse-grained model was applied to account for its internal transition⁵, where the energy function is built as a combination of two states, $\mathcal{E} =$

$-\log(e^{-\beta(\mathcal{E}_1+\delta)} + e^{-\beta\mathcal{E}_2})/\beta$ with $\beta = 1/(k_B T)$ ⁶. The parameter of δ can be viewed as a energy difference between two states, represented by \mathcal{E}_1 and \mathcal{E}_2 ⁵. In practice, δ was used as an advantage for sampling to accelerate the transition from one state to another. Models of inactive and active states of the catalytic domain were taken from the crystal structures of Hck (PDB entry: 1QCF) and c-Src (PDB entry: 1Y57), respectively^{7,8}. The average scattering profile calculated such assembly state is sufficiently different to be “SAXS-distinguishable”, as shown in Fig. 3 Similar to the catalytic domain, a coarse-grained representation for SH3 and SH2 domains was modeled after their structures^{5,9}. We used the structures of SH3 and SH2 taken from the inactive state of the down-regulated Hck⁷. The coarse-grained model was built at the residue level where each residue is represented by its C α atom. It includes bond, angle, and dihedral interactions between adjacent residues, repulsive interactions for residue pairs that are not in contact, and attractive interactions for long-range contacts⁹.

To sample the landscape of the Hck assembly, a total of 6 sets of simulations with the model and enhanced sampling techniques described above were carried out. Each of them was performed either by turning on/off the intramolecular interactions, by increasing the SH2-tail interactions accounting for mutants, or by adjusting δ for sampling ($\delta = 0$ when simulations start from the inactive compact or $\delta = -25$ kcal/mol when simulations start from a disassembled state). Langevin molecular dynamics were implemented with a friction coefficient of 50 ps⁻¹ and a time step of 0.01 ps, and each simulation has 20 μ sec. It results in a total of 120 μ sec simulation data for further analysis.

For post-data analysis, a clustering and mapping method as described previously was applied to simulation data¹⁰. The resulting configurations were clustered into 25 structural clusters, where a list of 89 inter-domain pairwise residue distances were used in a K-means clustering calculation. A resulting transition probability matrix was then constructed to monitor inter-cluster hopping. The use of this transition matrix is two-fold. First, an estimation about the probability of each structural cluster is yielded at a first-order approximation. Obtaining the probability distribution can help solve the problem of local oversampling where certain regimes of the landscape were heavily sampled locally but less favored globally. Second, a connectivity map connecting 25 structural clusters was constructed from the transition matrix, thus providing a measure of transitional pathways (see **Fig. 1**). More details on constructing this connectivity map can be found in previous publications^{5,10}.

SAXS computation and state assignment

A coarse-grained Fast-SAXS method was subsequently applied to compute theoretical SAXS profiles of all structural clusters¹¹. Fast-SAXS takes advantage of the coarse-grained nature of SAXS data, providing a rapid scattering determination from massive configurations generated from Hck simulations. In the Fast-SAXS calculations, each configuration was soaked in explicit water to account for the contribution from the hydration layer to the scattering. For each cluster, a number of 100 randomly selected configurations was used to compute the averaged scattering profiles (plus their standard deviations). Thus, a total of $n = 25$ scattering profiles ($\log I_i(q) \pm \delta \log I_i(q)$ where $i = 1, \dots, n$) were achieved.

To measure the scattering similarity between structural clusters, a next level K-means clustering procedure was performed. This scattering clustering was based on pair-wise scattering similarity,

$$S_{ij} = \left[\sum_{q_1}^{q_2} \frac{(\log I_i(q) - \log I_j(q))^2}{(\delta \log I_i(q)^2 + \delta \log I_j(q)^2)} \right]^{1/2} \quad (1)$$

where $q_1 = 0.01 \text{ \AA}$ and $q_2 = 0.6 \text{ \AA}$ were used and $\delta \log I_i(q)$ is the uncertainty of scattering intensity $\log I_i(q)$ of Cluster i . This scattering clustering results in the assignment of the 25 structural clusters into a total of 9 scattering states shown in Fig. 2. Each of the 9 scattering state includes multiple conformers, as shown in Figs. 5 to 13.

1. Seeliger, Markus A., Young, Matthew, Henderson, M. Nidanie, Pellicena, Patricia, King, David S., Falick, Arnold M., & Kuriyan, John (2005) *Protein Sci.* **14**(12), 3135–3139.
2. Zarrinpar, Ali, Bhattacharyya, Roby P., & Lim, Wendell A. (2003) *Sci. STKE* **2003**(179), 1–10.
3. Woo, Hyung-June & Roux, Benoit (2005) *Proc. Natl. Acad. Sci. USA* **102**(19), 6825–6830.
4. Lerner, Edwina C., Trible, Ronald P., Schiavone, Anthony P., Hochrein, James M., Engen, John R., & Smithgall, Thomas E. (2005) *J. Biol. Chem.* **280**(49), 40832–40837.
5. Yang, Sichun & Roux, Benoît (2008) *PLoS Comput. Biol.* **4**(3), e1000047.
6. Best, Robert B., Chen, Yng-Gwei, & Hummer, Gerhard (2005) *Structure* **13**, 1755–1763.
7. Schindler, Thomas, Sicheri, Frank, Pico, Alexander, Gazit, Aviv, Levitzki, Alexander, & Kuriyan, John (1999) *Mol. Cell* **3**, 639–648.
8. Cowan-Jacob, Sandra W., Fendrich, Gabriele, Manley, Paul W., Jahnke, Wolfgang, Fabbro, Doriano, Liebetanz, Janis, & Meyer, Thomas (2005) *Structure* **13**, 861–871.
9. Yang, Sichun, Onuchic, Jose N., & Levine, Herbert (2006) *J. Chem. Phys.* **125**(5), 054910.
10. Yang, Sichun, Banavali, Nilesh K., & Roux, Benoît (2009) *Proc. Natl. Acad. Sci. USA* **106**, 3776–3781.
11. Yang, Sichun, Park, Sanghyun, Makowski, Lee, & Roux, Benoît (2009) *Biophys. J.* **96**(11), 4449–4463.
12. Sicheri, Frank, Moarefi, Ismail, & Kuriyan, John (1997) *Nature* **385**, 602–609.

13. Lerner, Edwina C. & Smithgall, Thomas E. (2002) *Nat. Struct. Biol.* **9**(5), 365–369.
14. Cowan-Jacob, Sandra W., Fendrich, Gabriele, Manley, Paul W., Jahnke, Wolfgang, Fabbro, Doriano, Liebetanz, Janis, & Meyer, Thomas (2005) *Structure* **13**, 861–871.

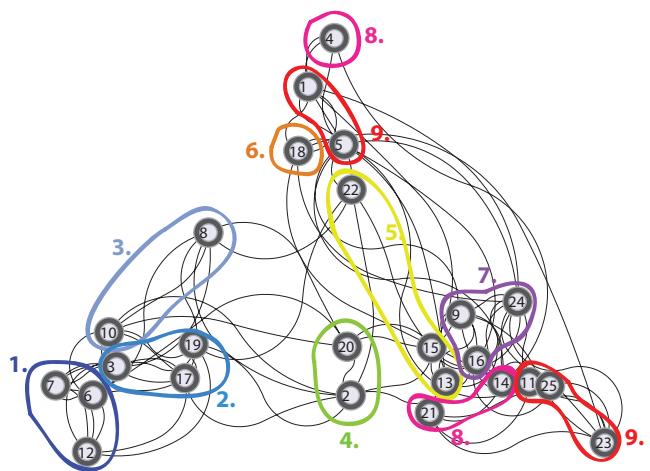


Figure 1: Graph representation of 25 structural clusters and their state assignment into 9 scattering states based on the scattering similarity.

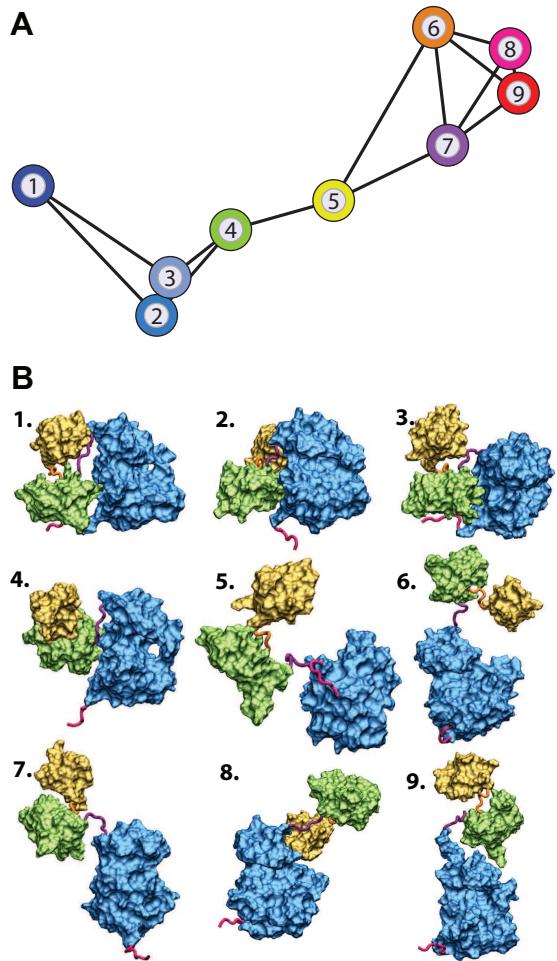


Figure 2: Multiple assembly conformational states adopted by the multi-domain Hck from a coarse-grained model. Graph representation of the network of Hck assembly states. The state-state distances are plotted according to a metric based on scattering difference computed by Eq. (1).

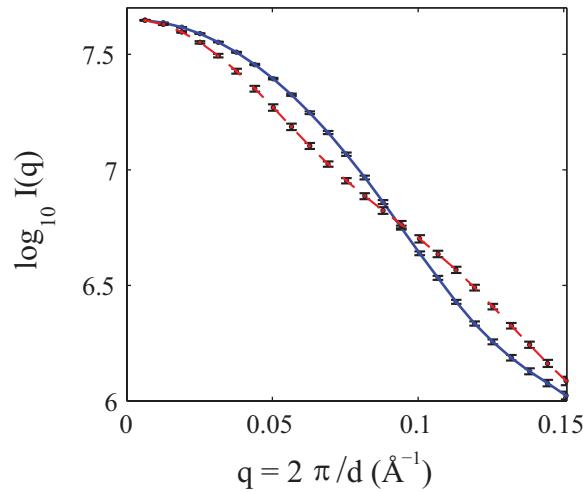


Figure 3: The configuration generated by the CG model can be organized into states exhibiting distinguishable SAXS profiles. Shown are the SAXS scattering profile from state 1 (blue solid line) and of state 8 (red dashed line) taken from Fig. 1. The SAXS profile $I(q)$ are shown in a logarithmic scale and q is the amplitude of momentum transfer. An efficient Fast-SAXS method was used to compute average scattering profiles of each state¹¹. The ensemble-averages (red and blue lines) and uncertainties (black bars) were obtained from 100 randomly selected configurations within each state, respectively. A complete list of χ^2 differences between states are given in Supplementary Table 1.

Table 1: A list of cross- χ^2 's calculated for all the scattering states using Eq. (2) with $q_{\min} = 0.0 \text{ \AA}$ and $q_{\max} = 0.15 \text{ \AA}$.

χ^2	1	2	3	4	5	6	7	8	9
1	0.0	113.3	78.4	81.5	164.8	214.9	210.0	356.3	418.9
2		0.0	20.6	52.3	128.4	194.2	177.7	299.3	347.9
3			0.0	38.6	101.2	168.3	145.0	232.4	267.1
4				0.0	48.4	129.2	94.5	158.1	169.5
5					0.0	109.8	60.8	127.4	130.7
6						0.0	68.2	39.4	49.8
7							0.0	53.7	41.1
8								0.0	25.3
9									0.0

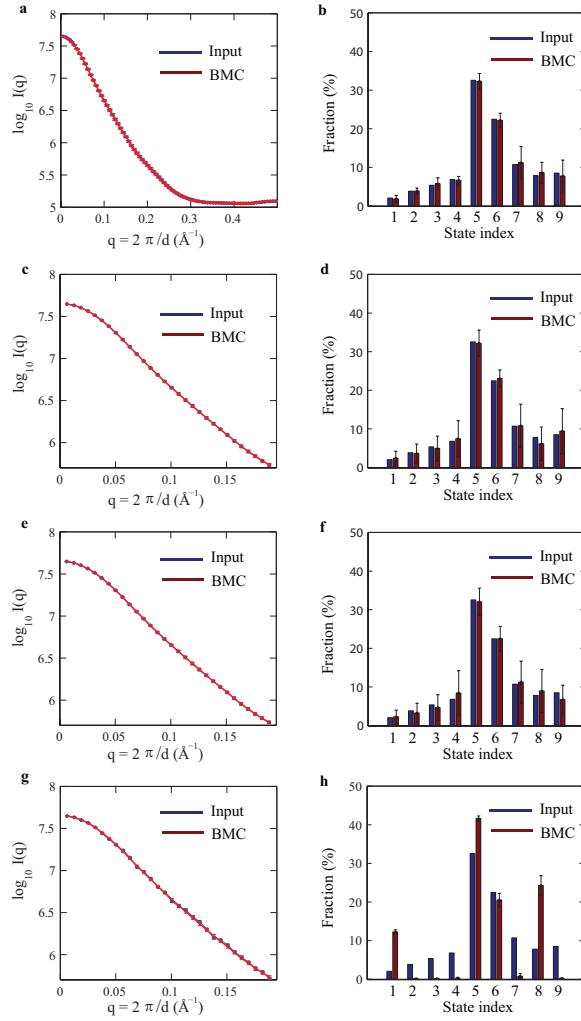


Figure 4: Theoretical validation of the Bayesian-based Monte Carlo (BMC) simulations. To test the BMC analysis, putative theoretical values of population fractions (blue stems) were used as inputs for the simulations. **(a, b)** Putative theoretical and BMC-simulated population fractions inferred from the full-set SAXS profiles and **(c, d)** from the low- q region. The inputs are in blue, while the simulated are in red. **(e, f)** Similar tests for the input scattering intensity $I(q)$ but with the addition of fluctuations in the amplitude of $\sigma(q)$ and **(g, h)** in the amplitude of $5\sigma(q)$, where $\sigma(q)$ is the uncertainty of $I(q)$.

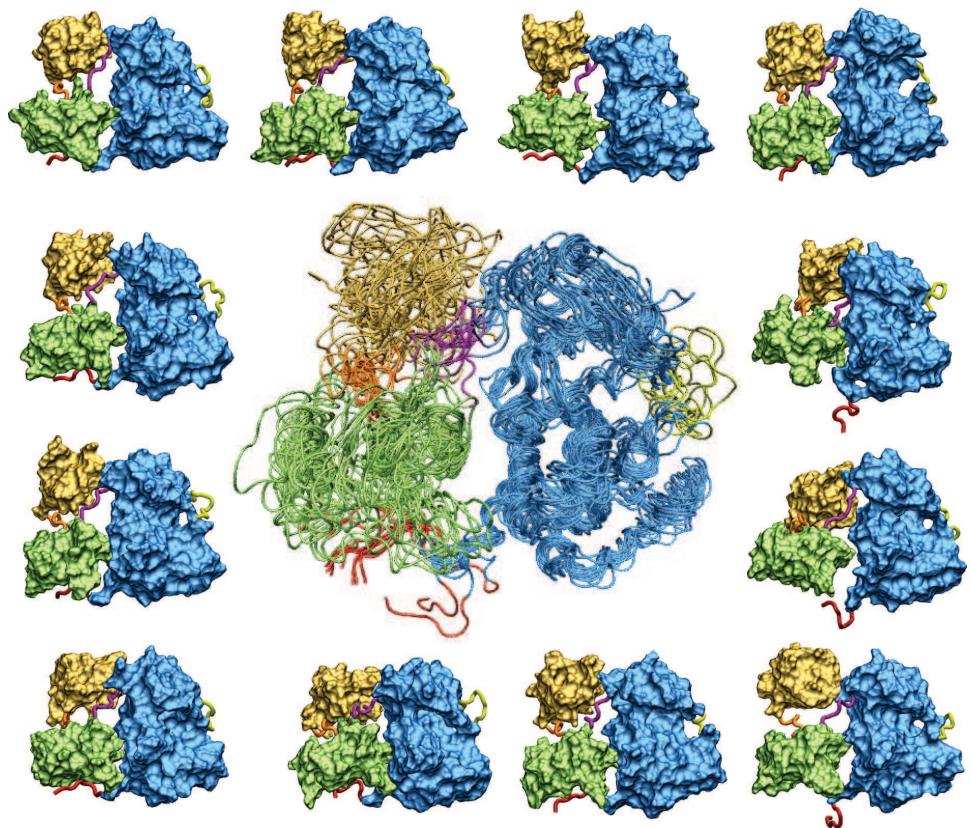


Figure 5: Snapshots and an ensemble of structures of state 1, which resemble the crystal structure of the fully assembled state (PDB entry: 1QCF)¹².

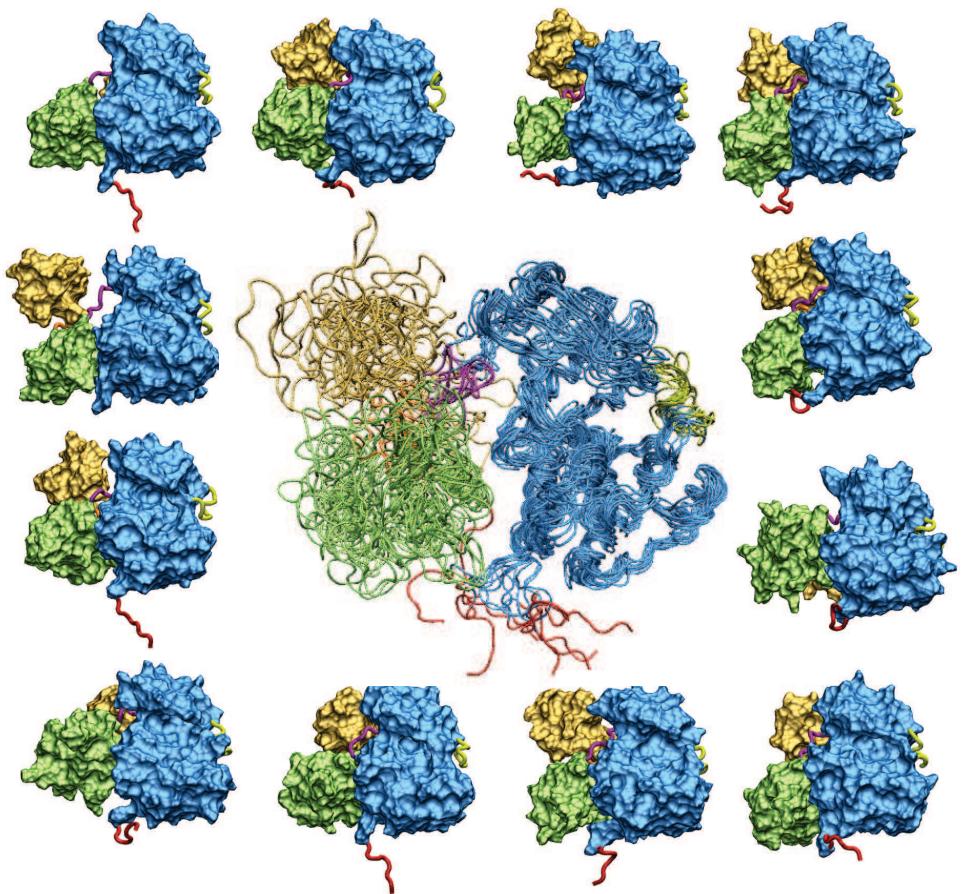


Figure 6: Snapshots and the ensemble of structures of state 2 (an SH3-displaced state with a detached C-tail). The observation of this SH3-displaced state is consistent with the previously reported hypothesis^{4,13}.

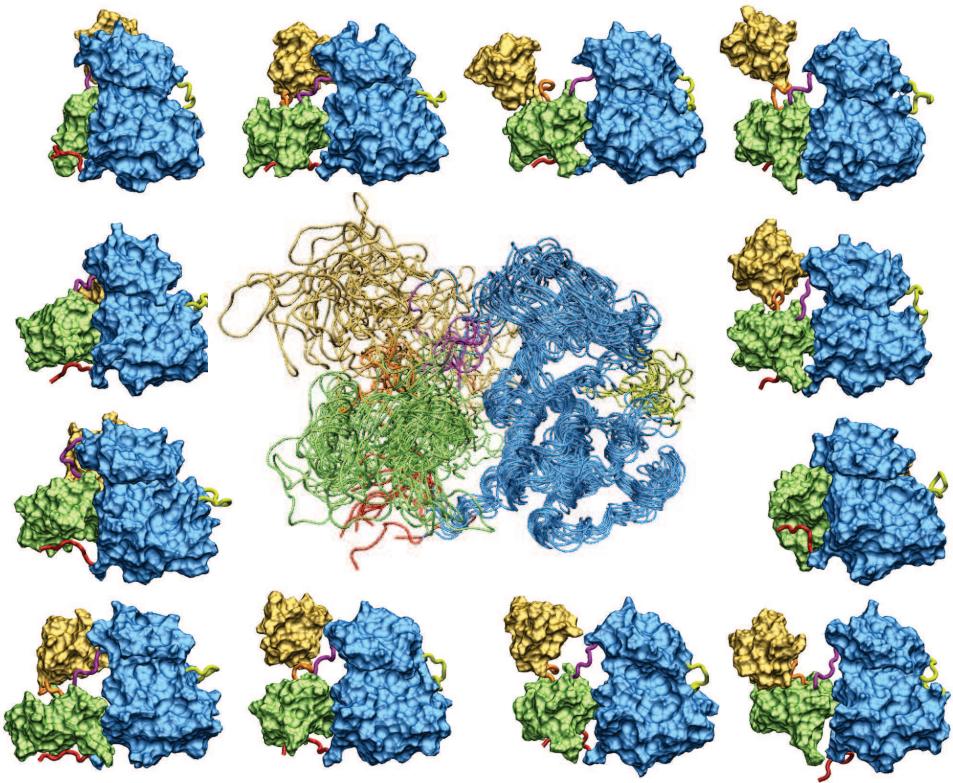


Figure 7: Snapshots and the ensemble of structures of state 3 (an SH3-displaced state with an associated C-tail).

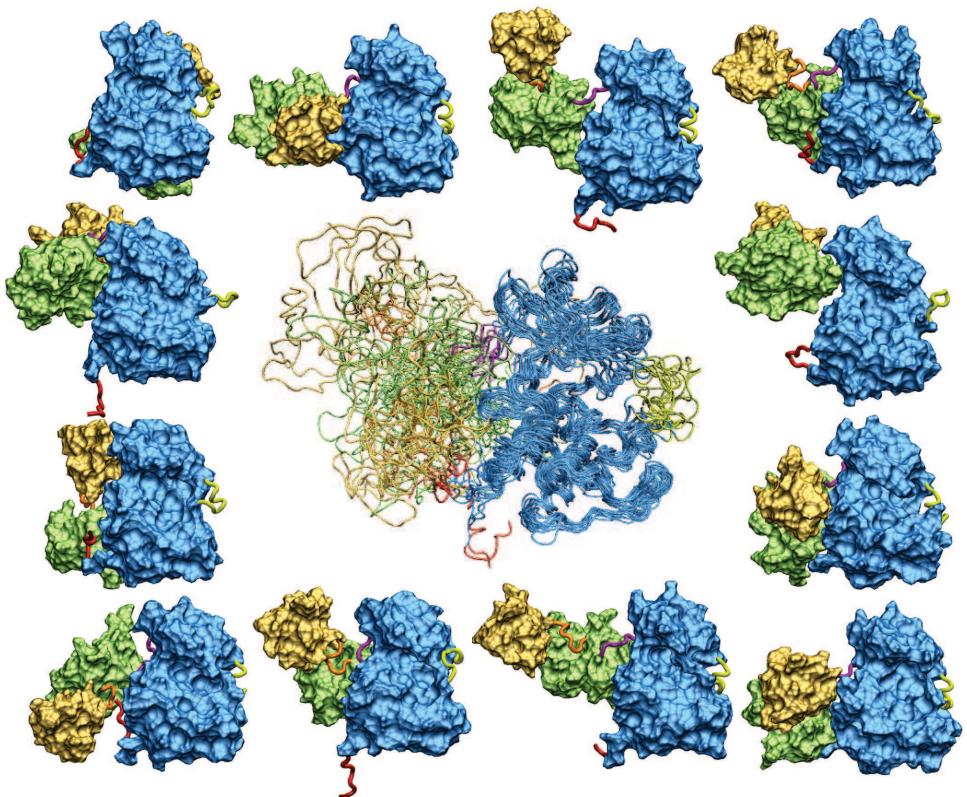


Figure 8: Snapshots and the ensemble of structures of state 4.

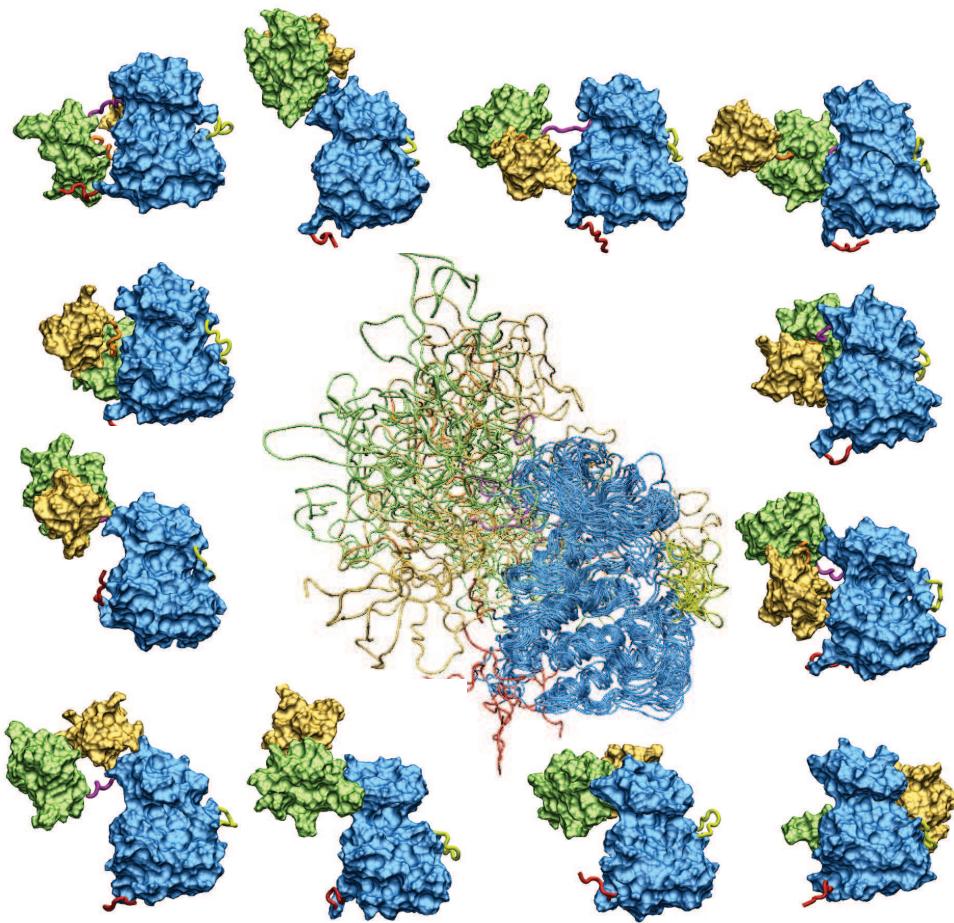


Figure 9: Snapshots and the ensemble of structures of state 5 (disassembled and extended).

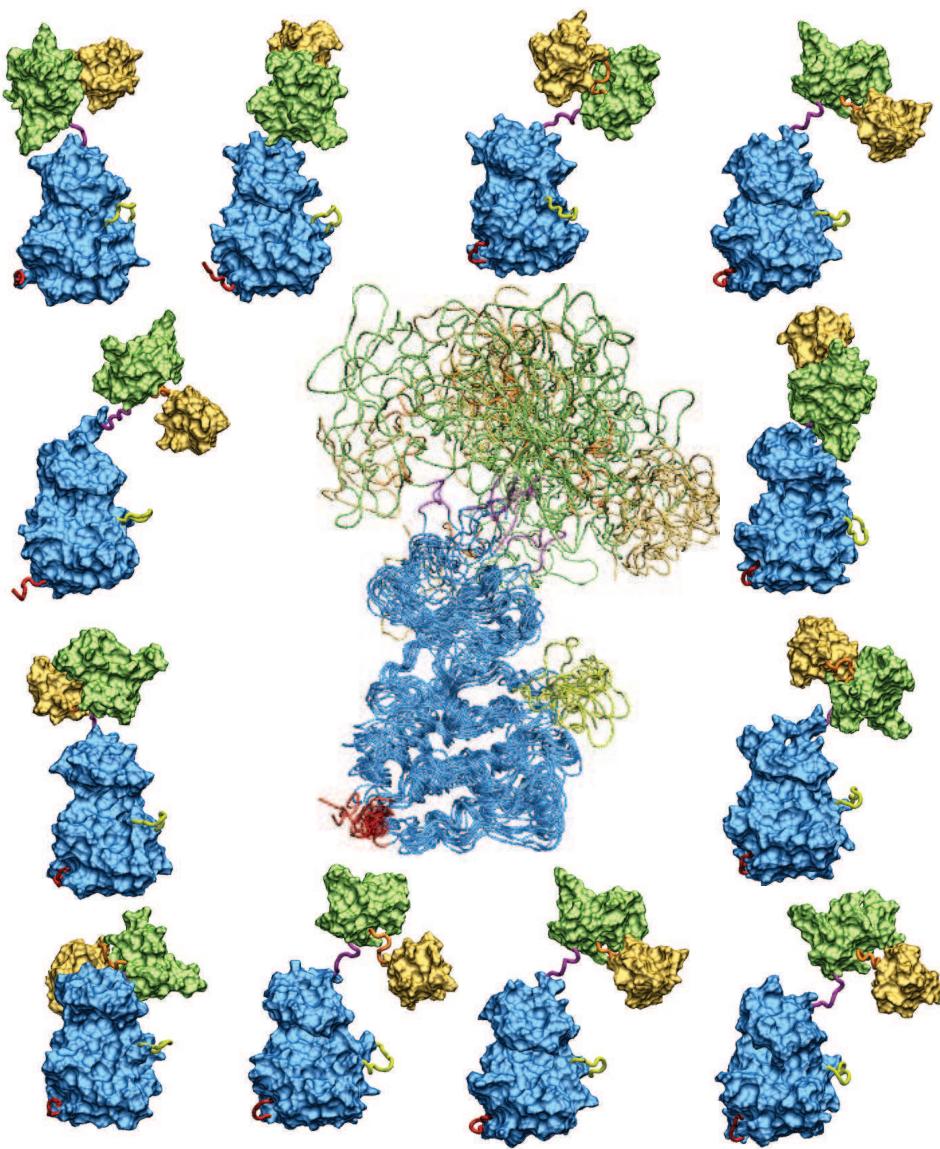


Figure 10: Snapshots and the ensemble of structures of state 6 (disassembled and compact).

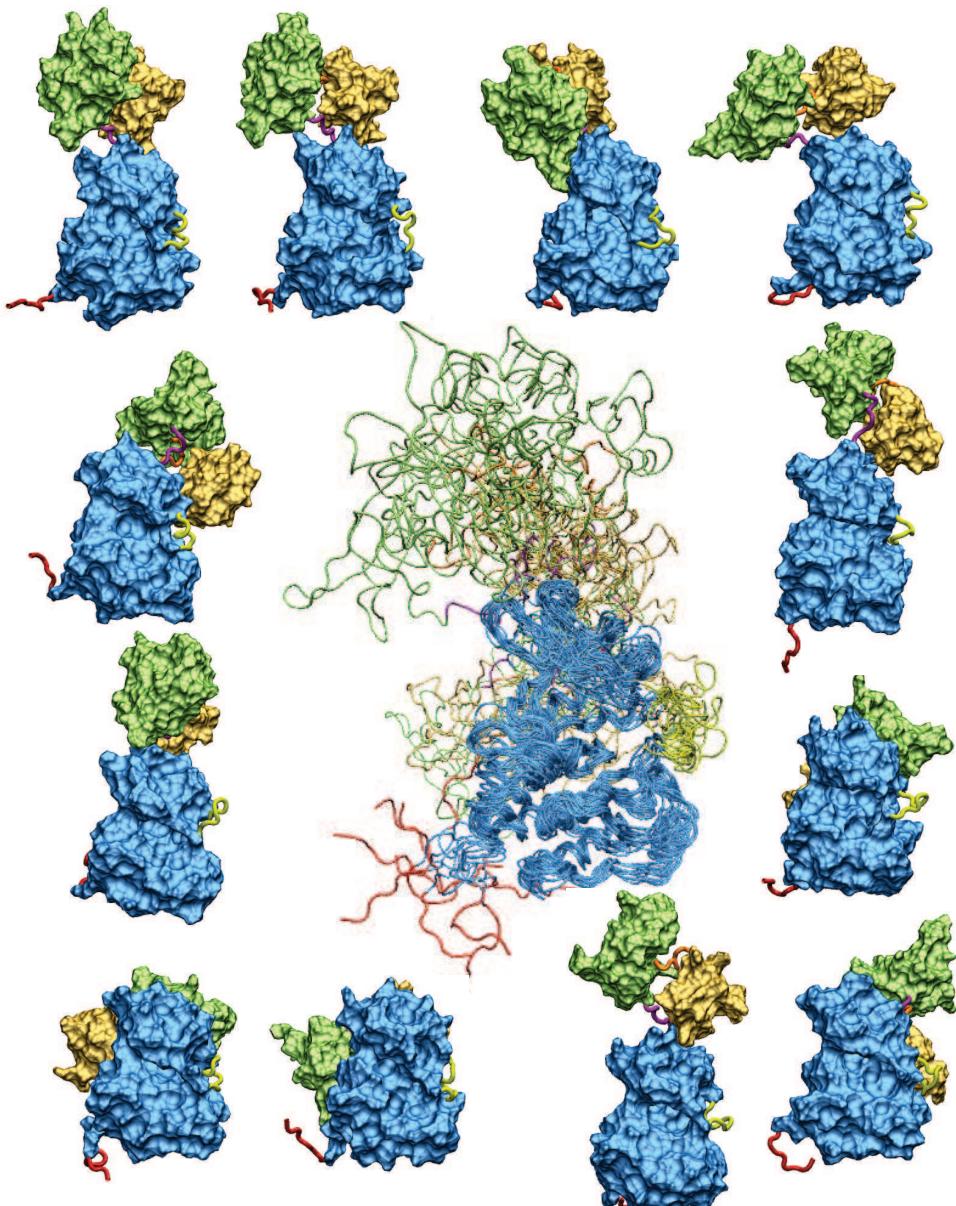


Figure 11: Snapshots and the ensemble of structures of state 7.

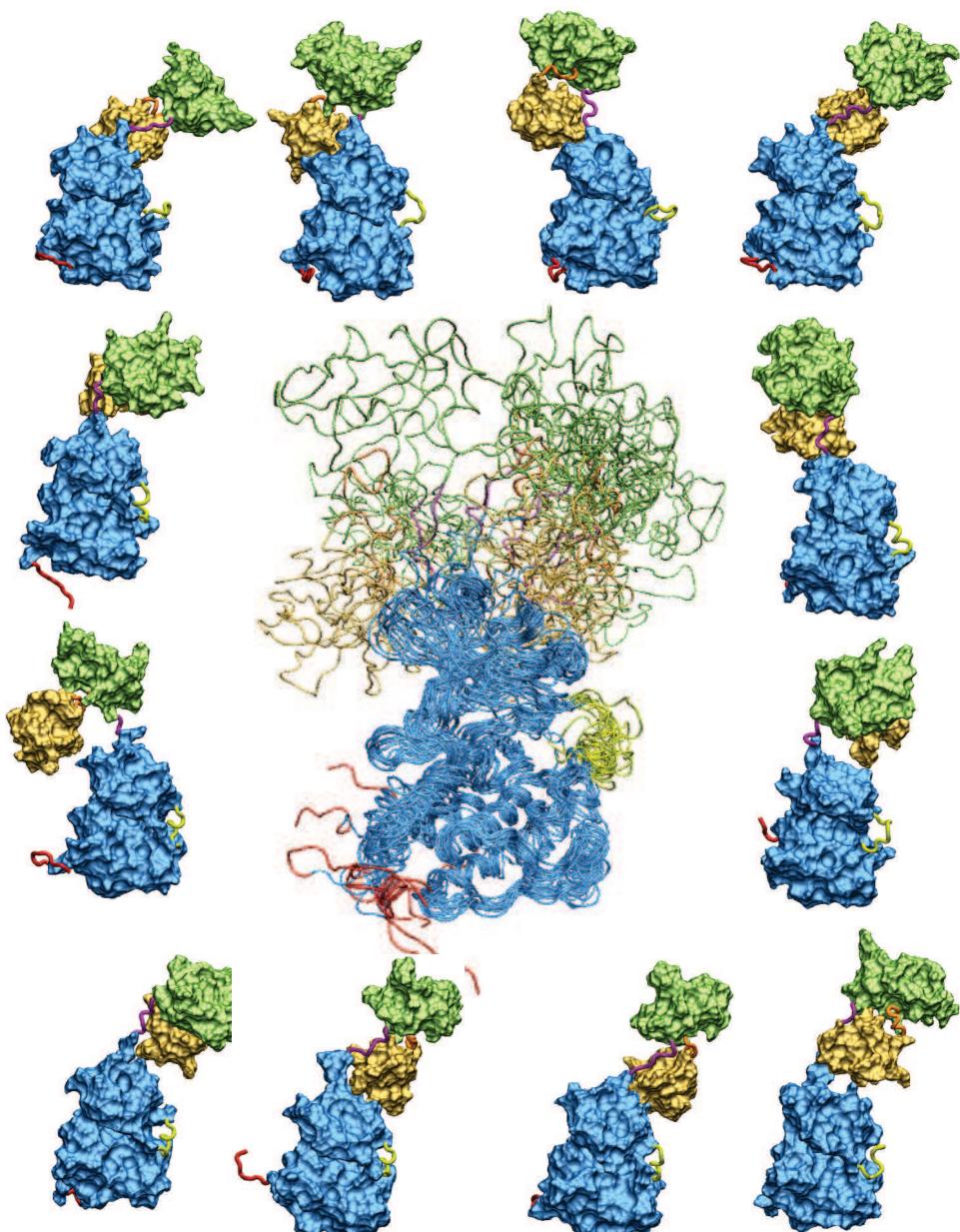


Figure 12: Snapshots and the ensemble of structures of state 8, which resembles the crystal structure of c-Src (1Y57.pdb)¹⁴.

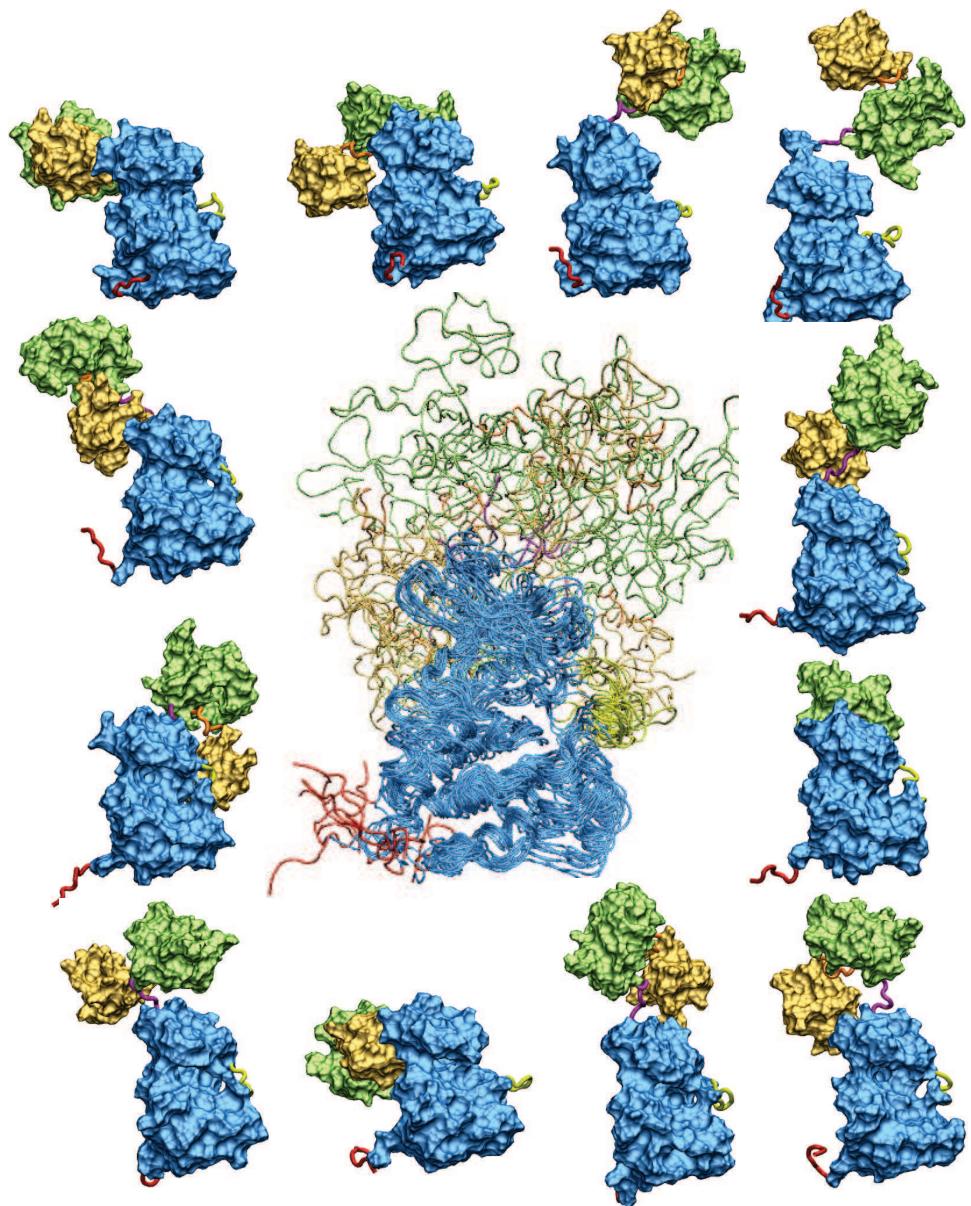


Figure 13: Snapshots and the ensemble of structures of state 9.

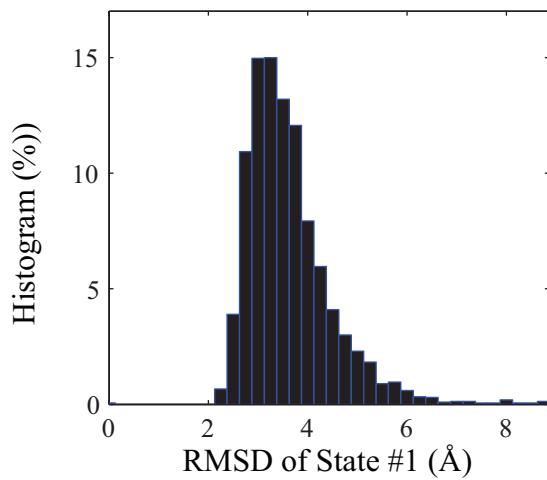


Figure 14: Histogram of the RMSD of the structures within state 1. A reference structure of the fully assembled structure (1QCF.pdb) was used for the RMSD calculations.

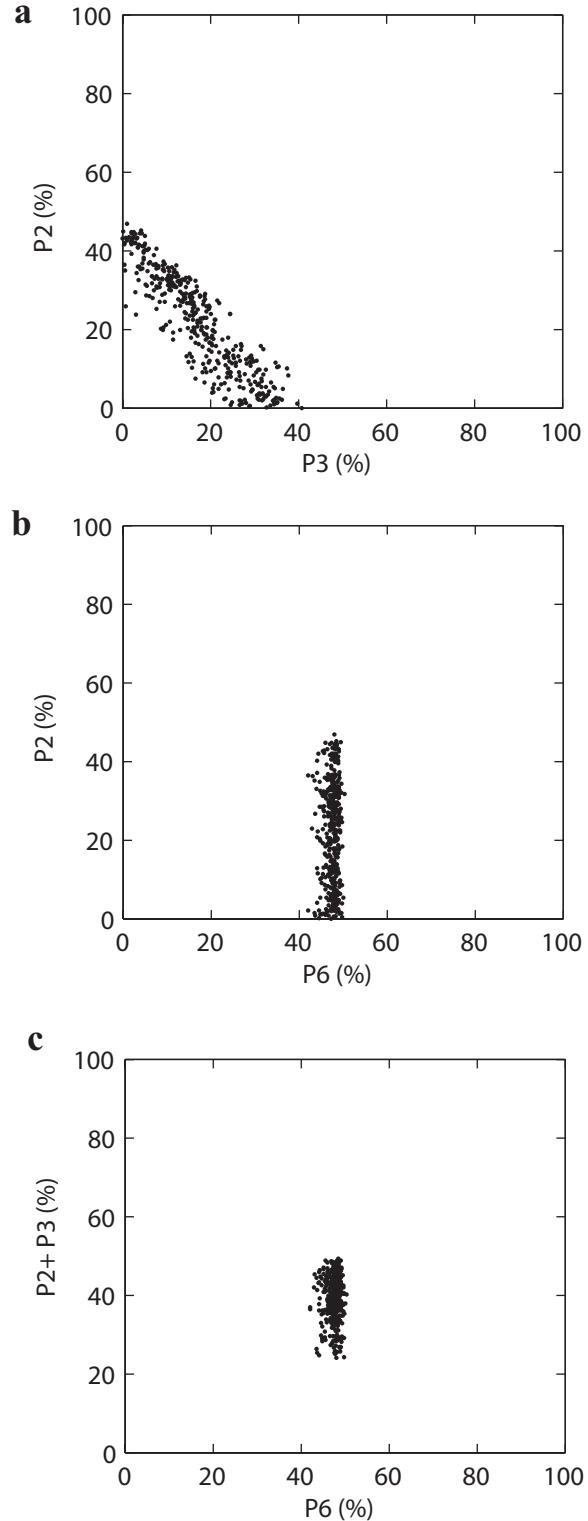


Figure 15: Correlation-function analysis of the population of states 2, 3, and 6 from a BMC simulation trajectory of SAXS data for the YEEI Ctail mutant in presence of p3 peptides (**Fig. 2e,f**). **(a)** Population of state 2 versus population of state 2. **(b)** Population of state 2 versus population of state 6. **(c)** A sum of the population of state 2 and of state 2 vs the population of state 6.

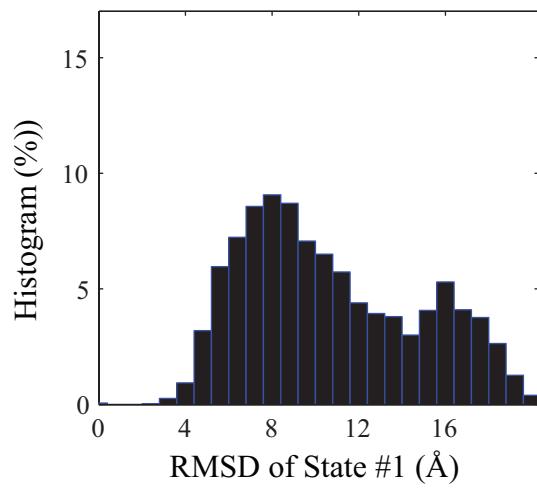


Figure 16: Histogram of the RMSD of the structures within state 8. A reference structure of the partially disassembled structure of c-Src was used for the RMSD calculations.