Electronic Supplementary Material (ESI) for Molecular BioSystems

Supporting Information

Dynamic optimization of signal transduction via intrinsic disorder

L. Michel Espinoza-Fonseca*

Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN 55455,

*E-mail: mef@ddt.biochem.umn.edu

METHODS

Generation of conformational ensembles

The crystal structures of the ATP-CDK2-Cyclin A and $p27^{kip1}$-CDK2-Cyclin A[‡] were first subjected to energy minimization with NAMD 2.7.[1] These minimized structures were further used for the generation of conformational ensembles with CONCOORD.[2] CONCOORD is a simple, yet robust computational algorithm that generates protein conformations around a known structure based on geometric restrictions. To build the conformational ensembles, we generated 500 structures for each complex using a maximum number of 10,000 iterations per structure. Additionally, we set a value of 10 Å so that structures do not violate the predefined bounds by more that this value in total. The program also includes an algorithm that discards structures with interatomic bumps, which was enabled during the generation of the structures.

Principal component analysis

Principal component analysis is a powerful technique aimed to reduce the dimensionality of a given data set. In the case of protein, this technique provides a description of the motion of the atoms (i.e., Cα atoms).[3] If we consider only the Cα atoms, the *N*-residue trajectory can be considered as a vector function of time *t*, namely

[‡] Only residues Lys25-Arg96 of $p27^{kip1}$ were solved in the crystal structure of the $p27^{kip1}$-CDK2-Cyclin A complex.

$$\mathbf{r}(t) = \left[r_{1x}(t), r_{1y}(t), ..., r_{Nz}(t) \right]^{\mathrm{T}}$$

of size $f = 3N$, containing the Cartesian coordinates at time t for residue 1 in the x-direction, residue 1 in the y-direction, . . . , up to residue N in the z-direction. The ij-entry C_{ij} of the covariance matrix C is the covariance of the positions for two degrees of freedom i and j, namely, $C_{ij} = \left\langle \left(r_i(t) - \left\langle r_i \right\rangle_t \right) \left(r_j(t) - \left\langle r_j \right\rangle_t \right) \right\rangle_t$ where $\left\langle X \right\rangle_T$ is the time average over the whole trajectory. Although the structural ensemble generated with CONCOORD is not time-dependent, we considered each structure as a snapshot of a trajectory for the analysis. Because the method relies on the average position of the atoms in the ensemble, this approach does not affect the results of the analysis. Principal component analysis diagonalizes C by solving $\Lambda = \mathrm{T}^{\mathrm{T}} C \mathrm{T}$, so as to obtain the diagonal matrix Λ with the diagonal entries being the eigenvalues ranked by magnitude. The c^{th} column of the transformation matrix T is the c^{th} eigenvector $\mathbf{v}_c$, that is, $\mathrm{T} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_f]$. Principal component analysis of the trajectories was performed using GROMACS,[4] and only Cα atoms were considered for the analysis of each conformational ensemble. The porcupine plots describing the motions extracted from principal component analysis were created using VMD.[5]

References

(1) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J Comput Chem* **2005**, *26*, 1781-1802.

(2) de Groot, B. L.; van Aalten, D. M.; Scheek, R. M.; Amadei, A.; Vriend, G.; Berendsen, H. J. *Proteins* **1997**, *29*, 240-251.

(3) Amadei, A.; Linssen, A. B.; Berendsen, H. J. *Proteins* **1993**, *17*, 412-425.

(4) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. *J Comput Chem* **2005**, *26*, 1701-1718.

(5) Humphrey, W.; Dalke, A.; Schulten, K. *J Mol Graph* **1996**, *14*, 33-38, 27-38.

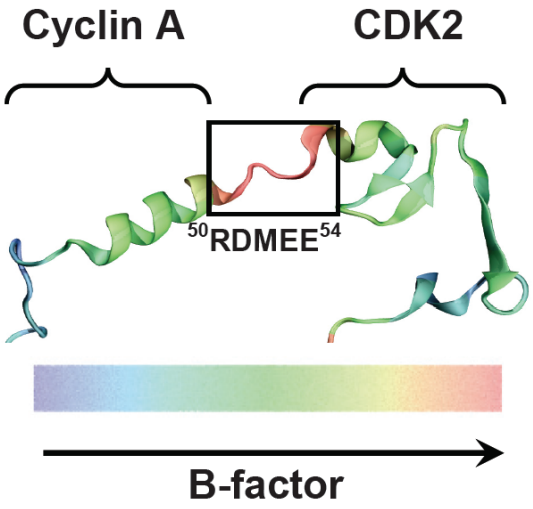

Figure S1. Ribbon representation of the structure of $p27^{kip1}$ bound to CDK2-CA. $p27^{kip1}$ is colored according to the experimentally determined backbone B-factors. The regions that bind to CDK2 and cyclin A are indicated by keys.