Supporting Information

Bacot-Davis et al. 10.1073/pnas.1411098111

SI Materials and Methods

L_E Phosphorylation Assays. Recombinant GST-L_E (EMCV) proteins and mutational derivatives T_3A, T_4A, T_9A, $T_{15}A$, $Y_{27}F$, $Y_{32}F$, $Y_{36}F$, $Y_{41}F$, $T_{47}A$, $T_{47}E$, and $Y_{41}F/T_{47}A$ were expressed and purified as previously described (1–3). Each protein was dialyzed into buffer [25 mM Hepes (pH 7.3), 150 mM KCl, and 2 mM DTT] and stored at −80 °C. Concentrations were determined with BCA protein assay kits (Thermo Scientific). The cell-free phosphorylation assays were essentially as previously described (4). GST (85 pmol) or GST-L_E (85.71 pmol) was incubated with buffer alone, CK2 (10 U; New England Biolabs), or CK2 (10 U) plus Syk (10.3 U; SignalChem) in the manufacturers' reaction buffers supplemented with 5.0 μCi [γ-^{32}P] ATP (3,000 Ci/mmol, 10 mCi/mL). After incubation at 37 °C for 60 min, samples were loaded for SDS/PAGE fractionation. To evaluate the Syk (only) reactions, the proteins were pretreated with CK2 and (cold) ATP before the addition of Syk and [γ-^{32}P] ATP. The resolved gels were silver stained and then exposed to phosphor screens for band visualization (GE Healthcare).

L_M for NMR. Unlabeled GST-L_M (Mengo) fusion protein was expressed in *E. coli* as previously described (5). Bacterial cultures contained 25 μM $ZnCl_2$ for proper protein folding. The expressed protein included a thrombin cleavage site for GST-tag removal. [^{15}N/^{13}C]-L_M0P was produced from BL-21 (DE3) cells transformed with pGST-L_M at 16 °C in [^{15}N/^{13}C] M9 medium [42.3 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 8.5 mM NaCl, 18.3 mM $^{15}NH_4Cl$, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.2% ^{13}C -D-glucose (wt/vol), 50 μg/mL kanamycin, pH 7.3] before induction with isopropyl-β-D-thiogalactopyranoside (IPTG, 1 mM). Cells were collected at an OD_{600} of 2.7–3.2. Harvest was on GSTrap FF columns, where the GST tags were removed before elution by reaction with thrombin protease as previously described (3). The affinity chromatography was followed by gel filtration using a Sephacryl S-100 column (GE Healthcare) and finally anion exchange with a Mini Macro-Prep High Q cartridge (Bio-Rad). The protein was concentrated using an Amicon Ultracentrifuge device (Millipore), treated with 0.25 mM EDTA for 5 min at 25 °C and then refolded by dialysis (2 L, 20 mM Hepes, pH 7.5, 100 mM KCl, 2 mM DTT, 0.25 mM $ZnCl_2$, 12 h, 4 °C). The protein was then dialyzed twice more into NMR buffer (20 mM Hepes, pH 7.5, 150 mM KCl, 2 mM $MgCl_2$, 5 mM DTT, 0.04% NaN_3, 12 h, 4 °C) before storage at −80 °C. The molecular weight of [^{15}N/^{13}C]-L_M0P was determined by matrix-assisted laser desorption ionization-MS (MALDI-MS) using a Bruker BIFLEX III mass spectrometer. Protein purity (>95%) was determined by SDS/PAGE followed by silver stain. Care was taken at all steps to use NMR-grade, metal-free reagents.

L_M Phosphorylation. [^{15}N/^{13}C]-L_M0P was purified by gel filtration, concentrated, and then incubated with CK2 alone (10 U) or with CK2 (10 U) followed by Syk (10.3 U) in a reaction buffer supplemented with 200 μM [^{31}P]ATP. The buffers were as provided by the manufacturers. Reactions were at 37 °C for 2.5 h. After phosphorylation, [^{15}N/^{13}C]-L_M(1P/2P) was dialyzed (10 mM Bis-Tris propane, pH 7.4, 50 mM NaCl, and 2 mM DTT) and purified by anion exchange using a Mini Macro-Prep High Q cartridge (Bio-Rad) over a 20-column volume salt gradient (50–500 mM NaCl) to remove the kinases. The proteins were treated with 0.25 mM EDTA for 5 min at 25 °C, refolded (as above), dialyzed into NMR buffer (as above), and then stored at −80 °C.

Ran for NMR. Plasmids encoding Hexa-His-Xpress–tagged human Ran GTPase (His-Xp-Ran) were a gift from Mary Dasso (National Institutes of Health, Bethesda, MD). Unlabeled protein was expressed in BL21 cells as previously described (3). [^{15}N/^{13}C] preparations were similar, except the cells were grown at 30 °C in M9 medium as described for [^{15}N/^{13}C]-L_M, with 50 μg/mL ampicillin instead of kanomycin. Initial protein purification steps (labeled or unlabeled) were as previously described (3), using a two-tier process of HisTrap HP (GE Healthcare) affinity chromatography followed by gel filtration using a Sephacryl S-100 column (GE Healthcare). If for use in NMR, the samples were treated with EDTA (5 mM, 30 min, 25 °C) and dialyzed (2 h, 25 °C) into NMR buffer (2 L, 20 mM Hepes, pH 7.4, 100 mM KCl, 2 mM $MgCl_2$, 2 mM DTT, and 0.04% NaN_3), followed by a second dialysis into fresh NMR buffer (overnight, 4 °C). Care was taken at all steps to use NMR-grade, metal-free reagents. Ran prepared this way (259 aa) retains the expression tag (43 aa) at the amino terminus of the full-length protein (216 aa). Recombinant GST-RCC1 (*X. laevis*) was purified as previously described (6) and then dialyzed into NMR buffer.

NMR Determinations. NMR data were collected at 25 °C using 280-μL samples in a 5-mm Shigemi tube. The protein concentration for labeled (^{15}N/^{13}C) or unlabeled L_M(0P/1P/2P) and Ran was 0.5 mM in the independent determinations. For Ran:L_M0P complexes, each protein was at 0.5 mM (one labeled and one unlabeled), and the samples were supplemented with (unlabeled) GST-RCC1 (1.4 nmol). The resolved spectra, including [^{1}H-^{15}N] HSQC, [^{1}H-^{13}C] HSQC, HBHA(CO)NH, CBCA(CO)NH, C(CO)NH, HC(CO)NH, HC(C)H-TOCSY, 3D ^{15}N-NOESY (t_{mix} = 120 ms), and 3D ^{13}C-NOESY (t_{mix} = 120 ms) were collected on a Bruker DRX-600 spectrometer equipped with an ^{1}H, ^{13}C, ^{15}N, ^{31}P three-axis gradient cryogenic probe. Standard NMR terminology includes NOESY (nuclear Overhauser effect spectroscopy), NHCABA (carbon alpha, carbon beta, amide spectroscopy), CBCA(CO)NH (carbon beta, carbon alpha, carbonyl spectroscopy), HSQC (heteronuclear single quantum coherence spectroscopy), TOCSY (total correlated spectroscopy), CARA (computer-aided resonance assignment).

Data Processing. Fig. S1 shows a summary flowchart for all data manipulations. The collected NMR data were processed using NMR-Pipe software (7), followed by peak picking and spin-system determination with CARA software (8). Cross-reference of ^{15}N-HSQC, ^{13}C-HSQC, CBCACONH, HBHACONH, CCONH, HCCONH, and HCCH-TOCSY spectra from uniformly labeled ^{15}N/^{13}C proteins assigned backbone and side-chain atoms (e.g., Figs. S2 and S4). TALOS+/RAMA+ generated dihedral angle constraint files (9) for input into CYANA (10) for structure calculations (Figs. S2 and S5*B*). The *-ref* comment alongside X-ray–determined structures of Ran generated upper and lower references for TALOS+ dihedral angle constraints in conjunction with chemical shifts assignments for NMR-resolved Ran (PDB ID code 2MMC) and Ran:(L_M0P). Cross-correlation of 3D HCCH-TOCSY, ^{13}C-NOESY, and ^{15}N-NOESY spectra, collected using a mixing time set to 120 ms for all triple-labeled data acquisition, assigned NOESY connectivity with CARA, SPARKY, CYANA, and CS-Rosetta (8, 10–14). Nonstandard amino acids and refinements (Table S3) were finalized by using VMD-X-PLOR (15). The quality of each generated structure was analyzed for restraint and geometry violations using the Duke University MolProbity web server (16, 17). All L_M datasets (71 aa) recorded the (4 aa)

amino-terminal extensions. The Ran datasets omitted tag-related peaks and numbered the protein (216 aa) according to its native sequence. Additional information for many of these technical processes is presented in the next sections.

Residual Dipolar Coupling. $^{1}H/^{15}N$ couplings for solution-state protein samples were measured using $^{1}J_{NH}$ modulation experiment, as previously described (18), with the addition of an evolution period during 2D ^{15}N-HSQC data collections.

NMR-PIPE NOESY Processing. The command lines used to process $^{15}N/^{13}C$-NOESY spectra following data collected on a 600-MHz Bruker spectrometer are provided in Dataset S1.

Dihedral Angle Constraints. Residue atoms were manually assigned using CARA (8). CARA wrote *.tab files for input into TALOS+/RAMA+, which generated de novo *.aco dihedral angular constraints and secondary structure element files. Peak lists and peak tables were imported from CARA into SPARKY (11) for figure visualization and assignment verification with PINE-SPARKY (19).

Automatic CYANA NOESY Assignments with CS-Rosetta Convergence. Structure calculations were conducted by running CYANA 3.0 (e-NMR) with peak intensities from three NOESY spectra: TALOS+ dihedral angle constraints, residual decoupling restraints, and initial, manually assigned NOE restraints as input files (10). After several runs using TALOS+-derived dihedral angle constraints and increasingly convergent high-quality CS-Rosetta and CYANA-derived NOE restraints, final automated NOESY assignments were generated in CYANA with the seeding NOEs. Blind CS-Rosetta structure determination was also conducted with final CYANA-derived dihedral angle and NOE constraint files (14). CS-Rosetta models were similar to CYANA-determined structures. For the L_M structures, the 10 final CYANA-generated coordinate sets (10) were too dynamic for PDB deposition, so in each case, the low energy model 1 was further examined with the Xplor-NIH package NAMD (15) energy minimization and collective variable analysis (PLUMED) to calculate the final, refined 10 states that were compatible with and deposited at PDB.

Structure Generation Using CYANA. TALOS+ *.aco torsion angle constraint files, CYANA-derived *.upl residual dipolar coupling, distance restraint files, and final assigned NOESY peak lists were used as input into CYANA to calculate 50 structures to output the 10 final low-energy states. Combinations of random restraints were used to improve structure qualities as above and as previously detailed (14). PRO-CHECK and AQUA through ADIT-NMR were used to validate the final 10 NMR-determined PDB-deposited structures (20).

Docking and Bioinformatics. TALOS+ algorithms (7) were used to define α and β motifs within the determined structures (Fig. S5). The lowest energy NMR states for L_M0P and Ran, as determined from the docked complexes, were submitted to HADDOCK via the public web portal (21). No constraints were specified. Docking interfaces for the lowest energy complex were evaluated online using PDBePISA resources (www.ebi.ac.uk/pdbe/pisa/) and the PIC (22). RMSDs for comparative states or pairwise structures used the "align" function of PyMol (23), specifying only the backbone c+n+ca+o atoms (Table S4). Structure display was by PyMol or Chimera (24).

1. Porter FW, Bochkov YA, Albee AJ, Wiese C, Palmenberg AC (2006) A picornavirus protein interacts with Ran-GTPase and disrupts nucleocytoplasmic transport. *Proc Natl Acad Sci USA* 103(33):12417–12422.
2. Porter FW, Palmenberg AC (2009) Leader-induced phosphorylation of nucleoporins correlates with nuclear trafficking inhibition by cardioviruses. *J Virol* 83(4):1941–1951.
3. Bacot-Davis VR, Palmenberg AC (2013) Encephalomyocarditis virus Leader protein hinge domain is responsible for interactions with Ran GTPase. *Virology* 443(1):177–185.
4. Basta HA, Bacot-Davis VR, Ciomperlik JJ, Palmenberg AC (2014) Encephalomyocarditis virus leader is phosphorylated by CK2 and syk as a requirement for subsequent phosphorylation of cellular nucleoporins. *J Virol* 88(4):2219–2226.
5. Cornilescu CC, Porter FW, Zhao KQ, Palmenberg AC, Markley JL (2008) NMR structure of the mengovirus Leader protein zinc-finger domain. *FEBS Lett* 582(6):896–900.
6. Nemergut ME, Macara IG (2000) Nuclear import of the ran exchange factor, RCC1, is mediated by at least two distinct mechanisms. *J Cell Biol* 149(4):835–850.
7. Delaglio F, et al. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277–293.
8. Keller RLJ (2004) *The Computer Aided Resonance Assignment Tutorial* (Cantina Verlag, Germany).
9. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44(4):213–223.
10. Güntert P (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol* 278:353–378.
11. Goddard TD, Kneller DG (2008) *SPARKY 3*. (University of California, San Francisco).
12. Bradley P, Misura KMS, Baker D (2005) Toward high-resolution de novo structure prediction for small proteins. *Science* 309(5742):1868–1871.
13. Shen Y, et al. (2008) Consistent blind protein structure generation from NMR chemical shift data. *Proc Natl Acad Sci USA* 105(12):4685–4690.
14. Lange OF, et al. (2012) Determination of solution structures of proteins up to 40 kDa using CS-Rosetta with sparse NMR data from deuterated samples. *Proc Natl Acad Sci USA* 109(27):10873–10878.
15. Schwieters CD, Clore GM (2001) The VMD-XPLOR visualization package for NMR structure refinement. *J Magn Reson* 149(2):239–244.
16. Davis IW, et al. (2007) MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35(web server issue):W375-83.
17. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12–21.
18. Tjandra N, Grzesiek S, Bax A (1996) Magnetic fields dependence of nitrogen-proton J splittings in 15N-enriched human ubiquitin resulting from relaxation interference and residual dipolar coupling. *J Am Chem Soc* 118(26):6264–6272.
19. Lee W, Westler WM, Bahrami A, Eghbalnia HR, Markley JL (2009) PINE-SPARKY: Graphical interface for evaluating automated probabilistic peak assignments in protein NMR spectroscopy. *Bioinformatics* 25(16):2085–2087.
20. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: {Programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8(4):477–486.
21. de Vries SJ, van Dijk M, Bonvin AMJJ (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc* 5(5):883–897.
22. Tina KG, Bhadra R, Srinivasan N (2007) PIC: Protein Interactions Calculator. *Nucleic Acids Res* 35(web server issue):W473-6.
23. Anonymous (2008) *The PyMOL Molecular Graphics System* (Schrödinger, LLC), Version 1.1r.1.
24. Pettersen EF, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.

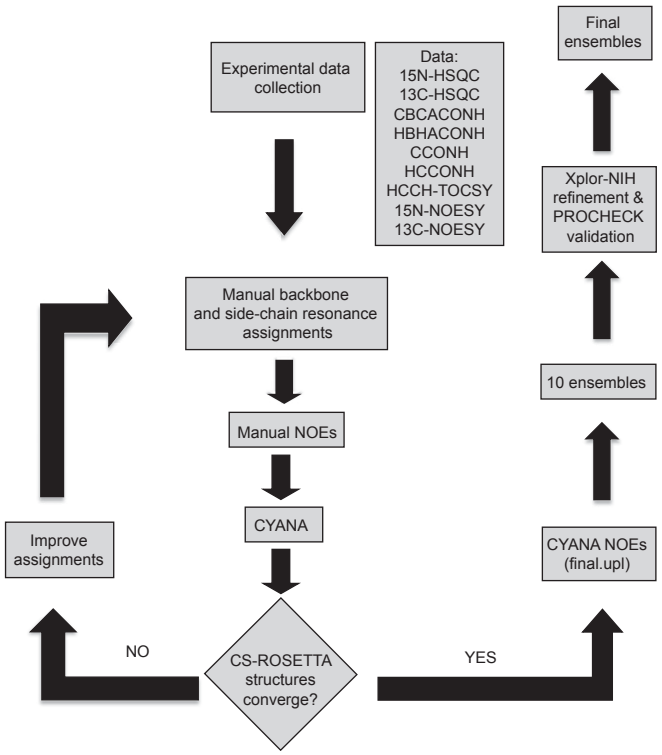

Fig. S1. Work flow of NMR structure determinations. Restraints files were generated from chemical shift assignments using TALOS+/RAMA+, CYANA, and CS-Rosetta suites for each determination, as indicated.

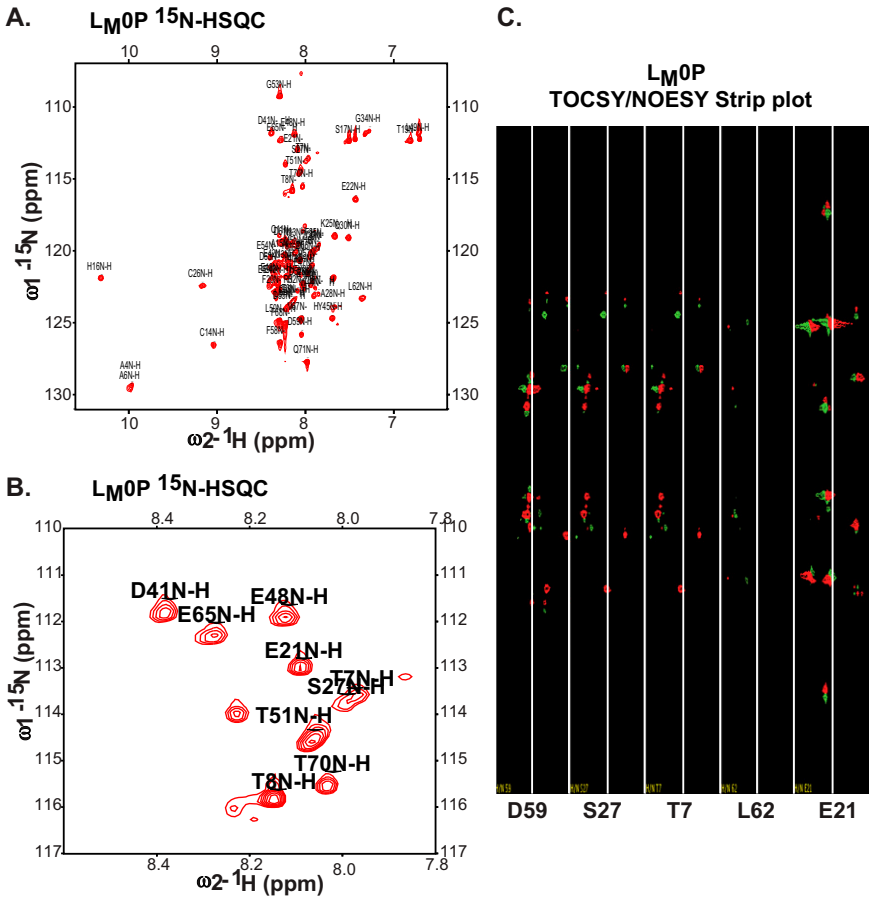

Fig. S2. Nuclei assignments for L_M0P. (*A*) 2D ^{15}*N*-HSQC assignments were used to determine ^{15}*N*-NOESY J-coupling connectivity. All assigned peaks are labeled. (*B*) Section of resonances of the 2D ^{15}*N*-HSQC of L_M0P, from *A,* magnified for clarity. (*C*) TOCSY-NOESY connectivity strips for L_M0P residues; D59, S27, T7, L62, and E21 are shown as examples.

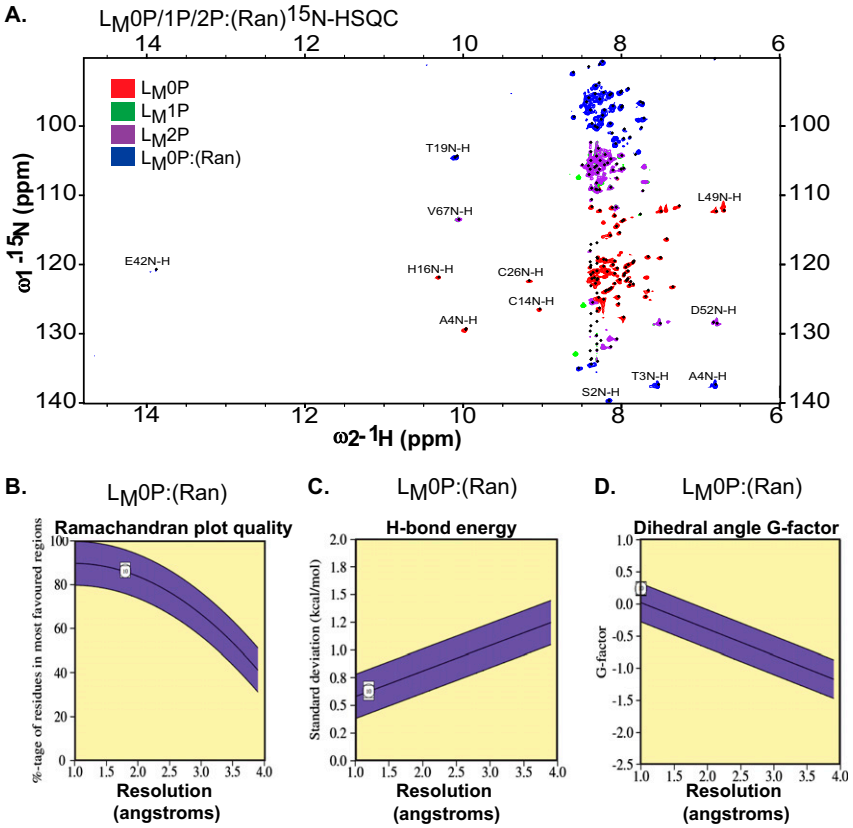

Fig. S3. L_M(0P/1P/2P) stereochemical parameters. (*A*) 2D ^{15}*N*-HSQC of L_M0P, L_M1P, L_M2P, and L_M0P:(Ran). To avoid obscuring visualization of the superimposed datasets, only a few peaks are shown as labeled. A fully annotated version of this image is available from the authors. (*B*) A 1.8-Å resolution Ramachandran plot quality. (*C*) Hydrogen bond energy SD of 0.6 compared with a typical value of 1.3. (*D*) Dihedral angle G-factor of 0.2 is within the favored region of dihedral angle conformations. For *B*–*D*, white box is current structure.

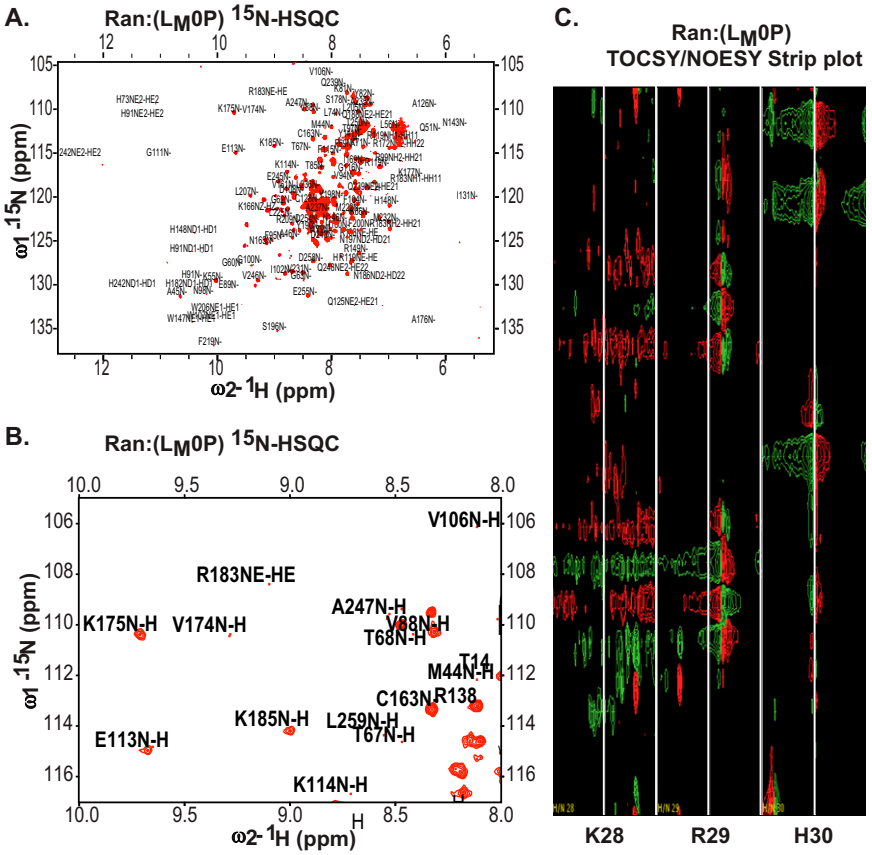

Fig. S4. Nuclei assignments for Ran:(L_M0P). (*A*) 2D ^{15}*N*-HSQC assignments of Ran, titrated with L_M0P, were used to determine ^{15}*N*-NOESY J-coupling connectivity. To avoid obscuring the complete visualization of most peaks, some assigned labels have been removed. A fully annotated version of this image is available from the authors. (*B*) A section of Ran:(L_M0P) 2D ^{15}*N*-HSQC residue resonances from *A* is magnified for clarity. (*C*) TOCSY-NOESY connectivity strips for Ran:(L_M0P) residues K28, R29, and H30 are shown as examples.

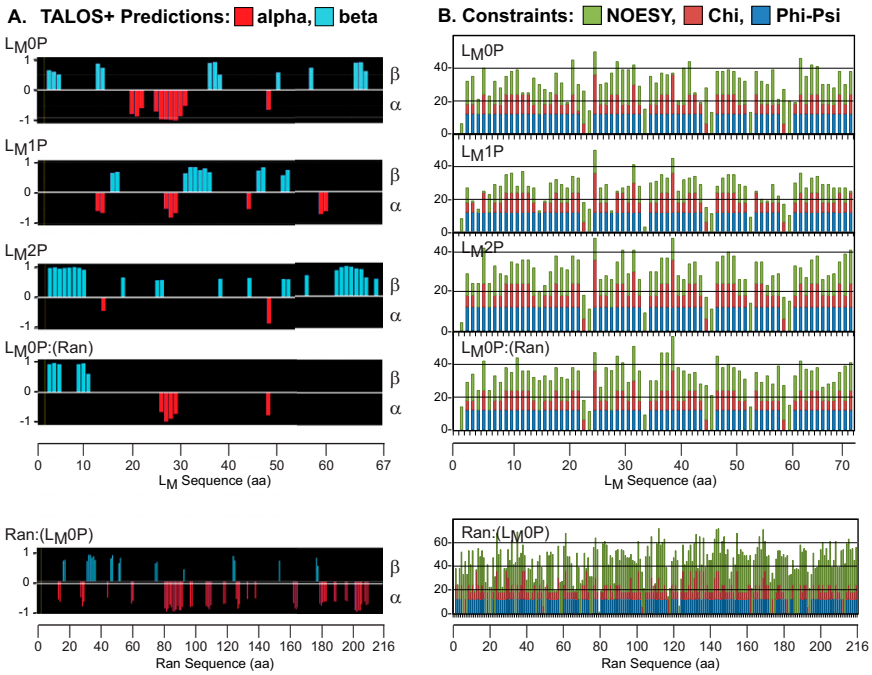

Fig. S5. (*A*) TALOS+ refinement. TALOS+/RAMA+ used the random coil index (RCI) method and Artificial Neural Network (ANN) to predict the secondary structures for Ran residues. Positive values (aqua) represent β-sheet structure predictions, negative values (red) represent α-helix structure predictions, and values of zero represent random coil conformations. Chemical shift-derived values are plotted for L_M0P, L_M1P, L_M2P, L_M0P:(Ran), and Ran:(L_M0P) according to each sequence. (*B*) As per the workflow chart in Fig. S1, constraints for each structure were generated from chemical shift assignments using TALOS+/RAMA+, CYANA, and CS-Rosetta suites. Observed NOESY (green), Chi (red), and Phi-Psi (blue) constraints are plotted as stacked bars for each residue.

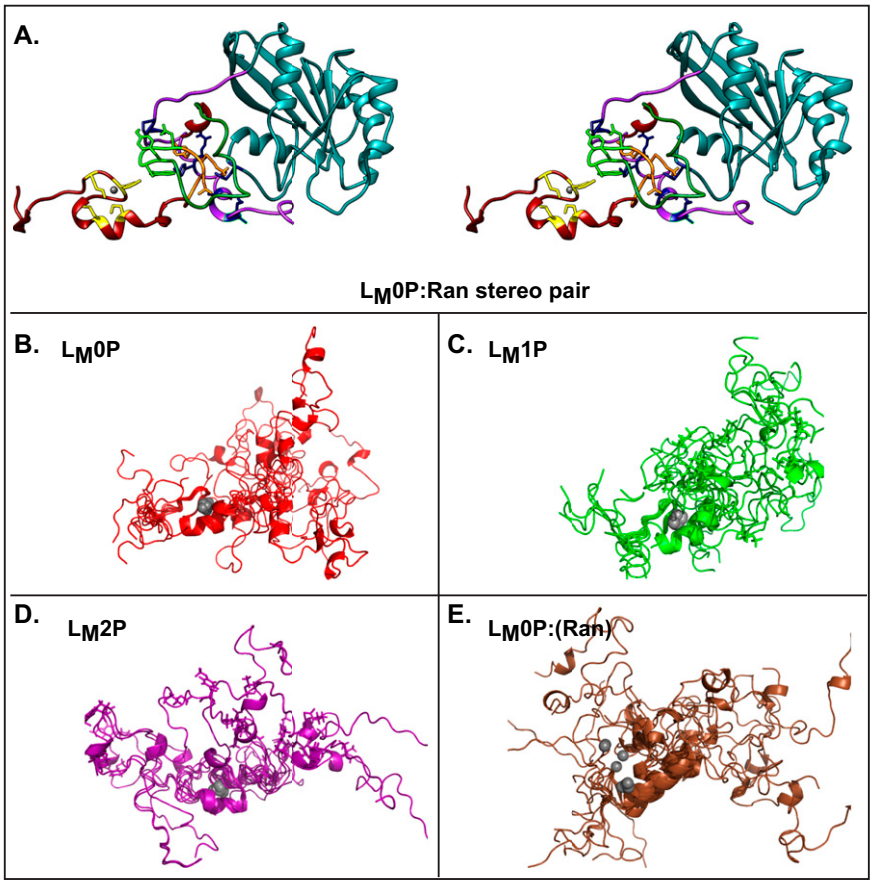

Fig. S6. (*A*) Stereo images of Ran:L_M0P complex. Illustration from Fig. 4*C* is reproduced as a stereo pair. (*B–E*) The 10 low-energy CYANA-generated coordinate files for each L_M determination are shown superimposed according to the common zinc finger motifs, using the PyMol align function. These figures illustrate the original sampling of dynamic ensembles. The initial states were then further culled into more a refined, low-energy, related series of PDB-acceptable ensembles (i.e., Figs. 2 and 4), as described in *SI Materials and Methods*.

Table S1. GST-L_E phosphorylation sites

	Kinase[†]		
Substrate*	CK2	Syk[‡]	CK2+Syk
GST	–	–	–
GST-L_E	+	+	+
GST-L_E T_3A	+	+	+
GST-L_E T_4A	+	+	+
GST-L_E T_9A	+	+	+
GST-L_E T_{15}A	+	+	+
GST-L_E Y_{27}F	+	+	+
GST-L_E Y_{32}F	+	+	+
GST-L_E Y_{36}F	+	+	+
GST-L_E Y_{41}F	+	–	+
GST-L_E T_{47}A	–	–	–
GST-L_E T_{47}E	–	+	+
GST-L_E Y_{41}F/T_{47}A	–	–	–

*Recombinant GST-L_E and mutant derivatives were prepared as in *Materials and Methods*.
[†]Reactions with these enzymes and [^{32}P]ATP gave strongly labeled proteins (+) as in or failed to label (–).
[‡]Reactions recording [^{32}P]ATP incorporation with Syk (only) were preceded by reactions with CK2 in the presence of unlabeled ATP.

Table S2. Ran NOESY distance restraints

	Number of NOESY peaks			Number of CYANA restraints				
					Long range*			
Target	Picked	Manual	CYANA	Restraints	SUP† = 1	SUP† < 1	Violated distant restraints	Violated angle restraints
L_M0P	885	11	874	256	18	19	0	0
L_M1P	580	8	572	129	14	15	0	1
L_M2P	804	0	804	353	11	11	0	0
L_M0P:(Ran)	1059	1	1058	397	11	11	0	0
Ran:(L_M0P)	5636	192	5444	600	82	53	0	5

Summary of NOESY cross-peaks used in L_M0P and Ran:(L_M0P) assignments with CYANA. Columns indicate cross-peaks for each respective group. CYANA semiautomated peak picking followed initial manual assignments. CYANA-generated NOE distance restraint reliabilities fell from 0 to 1. Combinations of random restraints were used to improve structure qualities as previously detailed (14).
*Long-range restraints $|i\text{-}j| \geq 5$.
†SUP, reliability of constraints as assigned by CYANA from 0 to 1.

Table S3. Structure quality

Crystallography equivalent resolution	H-bond energy (Å)	Dihedral angles G-factor (Å)	Ramachandran plot quality (Å)	H-bond mean parameter
L_M0P	1.5	0.2/1.0	1.8	0.7
L_M1P	1.7	0.1/1.0	2.3	0.7
L_M2P	1.1	0.1/1.0	2.2	0.6
L_M0P:(Ran)	1.2	0.2/1.0	1.8	0.6
Ran:(L_M0P)	1.5	0.4/1.0	1.0	0.7

PROCHECK/AQUA suites (20) through ADIT-NMR, assessed each structure quality as a final validation before PDB and BMRB deposition.

Table S4. Ran:(L_M0P) relative to crystallographic determinations

Ran PDB	Bound GNP	Length	All 1–216	Core 8–176	COOH 177–216	P-loop 16–25	Switch 1 32–45	Switch 2 66–79
Å RMSD vs. Ran:(L_M0P)-state 1								
1I2M	0	8–176	3.9	3.9	—	0.3	3.5	3.2
1BYU	GDP	9–177	12.7	4.0	—	0.2	4.1	3.2
3GJ0	GDP	1–207	12.6	4.0	7.1	0.2	4.2	3.2
1K5G	GTP	8–213	4.6	0.4	3.6	0.3	0.3	0.3
3GJX	GTP	9–179	1.6	1.6	—	0.2	0.3	3.4
3EA5	GTP	6–179	1.9	1.8	—	0.2	0.4	3.4
2BKU	GTP	9–177	1.7	1.7		0.2	0.5	3.4
1RRP	GTP	8–211	4.9	1.1	4.3	0.3	0.6	2.3
Ran:(L_M0P) States 2–10								
Average	0	1–216	4.5	0.2	4.9	0.1	0.2	0.1
Variance	—	—	0.4	0.0	1.2	0.0	0.0	0.0

Backbone atoms of Ran (n+ca+c+o), within the indicated PDB files, were compared with Ran:(L_M0P)-state-1 using the PyMOL align function over the indicated residues. Similar alignments assessed variance among all pairwise Ran:(L_M0P) state 1–10 coordinates. RMSD values rounded to 0.1 Å. Length is the resolved residues within each file.

Other Supporting Information Files

Dataset S1 (DOCX)