Solution structures of Mengovirus Leader protein, its phosphorylated derivatives, and in complex with nuclear transport regulatory protein, RanGTPase

Valjean R. Bacot-Davis[a], Jessica J. Ciomperlik[a], Holly A. Basta[a], Claudia C. Cornilescu[b], and Ann C. Palmenberg[a,b,1]

[a]Institute for Molecular Virology and [b]Department of Biochemistry, University of Wisconsin–Madison, Madison, WI 53706

Edited by Bert L. Semler, University of California, Irvine, CA, and accepted by the Editorial Board September 24, 2014 (received for review June 13, 2014)

Cardiovirus Leader (L) proteins induce potent antihost inhibition of active cellular nucleocytoplasmic trafficking by triggering aberrant hyperphosphorylation of nuclear pore proteins (Nup). To achieve this, L binds protein RanGTPase (Ran), a key trafficking regulator, and diverts it into tertiary or quaternary complexes with required kinases. The activity of L is regulated by two phosphorylation events not required for Ran binding. Matched NMR studies on the unphosphorylated, singly, and doubly phosphorylated variants of Mengovirus L (L_M) show both modifications act together to partially stabilize a short internal α-helix comprising L_M residues 43–46. This motif implies that ionic and Van der Waals forces contributed by phosphorylation help organize downstream residues 48–67 into a new interface. The full structure of L_M as bound to Ran (unlabeled) and Ran (216 aa) as bound by L_M (unlabeled) places L_M into the BP1 binding site of Ran, wrapped by the conformational flexible COOH tail. The arrangement explains the tight K_D for this complex and places the L_M zinc finger and phosphorylation interface as surface exposed and available for subsequent reactions. The core structure of Ran, outside the COOH tail, is not altered by L_M binding and remains accessible for canonical RanGTP partner interactions. Pull-down assays identify at least one putative Ran:L_M partner as an exportin, Crm1, or CAS. A model of Ran:L_M:Crm1, based on the new structures suggests L_M phosphorylation status may mediate Ran's selection of exportin(s) and cargo(s), perverting these native trafficking elements into the lethal antihost Nup phosphorylation pathways.

Leader protein | NMR | phosphorylation | RanGTPase | cardiovirus

The *Picornaviridae* family encompasses 26 genera and 46 species (1). Common to all isolates, the single-stranded, positive-sense RNA genome is characterized by a long ORF encoding 10–14 concatenated protein-coding genes. The replication cycle initiates as soon as this ORF is translated, and the resulting polyprotein is processed (co- and posttranslationally) into the required active components, which include seven to eight nonstructural proteins (NSPs) and three to four capsid proteins designated according to a standard "L-4-3-4" nomenclature (2). The Leader (L) proteins, when present, precede the capsid proteins (1ABCD) and all of the other NSPs (2ABC and 3ABCD). Most NSPs have vital roles in viral replication, but the L and 2A proteins are key determinants for antihost responses. The specific genes at these locales vary significantly among the genera and even among otherwise related species.

Unique to isolates in the *Cardiovirus* genus, the Leader gene encodes a small [67–76 amino acids (aa)] highly acidic protein (p*I*, 3.2–3.6) with very unusual properties. When L is expressed in cells by viral or recombinant introduction, it binds tightly (3 nM K_D) with 1:1 stoichiometry to the nuclear transport regulator, RanGTPase (3). Ran, a member of the Ras superfamily of GTPases, normally alternates between nuclear GTP- and cytoplasmic GDP-bound conformers, acting as a molecular switch for the coordinated transport of large molecules back and forth through the nuclear pores (4). However, when L binds Ran, the perverted complex recruits and activates a specific cohort of cellular kinases responsible for the L-induced hyperphosphorylation of Phe/Gly-containing nuclear pore proteins (Nups) (5–8). The consequence is rapid, potent inhibition of active nuclear-cytoplasmic trafficking. Because picornaviruses replicate in the cytoplasm, this inhibition is detrimental only to the cell. Among the measured results, there is antagonism of IFN transcription (9–12), impediment of cellular stress granule formation (13), and retention of cellular mRNA transcripts in the nucleus (12). These cumulative activities allow cardioviruses to negate almost all host antiviral innate immune responses and enhance their pathogenicity during infection.

The best studied L proteins, representing the *Encephalomyocarditis virus* (EMCV) species, are from EMCV-R (L_E) and Mengovirus (L_M) isolates. The species as a whole shares ∼95% identity here, but these strains differ by a single substitution ($L_{14}M$) in the 67-aa protein length (Fig. 1*A*). The change is in a conserved, amino-proximal CHCC zinc finger motif (aa 10–22), the structure of which was determined by NMR for the L_M protein (14). Technical difficulties hampered resolution of the complete protein, but a full coordinate set was recently completed [Protein Data Bank (PDB) ID code 2M7Y]. Outside of the zinc finger, the rest of that protein configured predominantly as random coil with a small β-hairpin in the COOH-proximal acidic domain (aa 37–61). The remaining interior residues, or

Significance

We describe here the first, to our knowledge, NMR structure of full-length RanGTPase protein. The protein is captured complexed to the Mengovirus Leader (L_M) protein. The pair, once bound, triggers a lethal hyperphosphorylation cascade of nuclear pore proteins, leading to enhanced virus replication and cell death. Structures for L_M in multiple phosphorylation states, and as bound by Ran, show induced fit reactive faces that putatively recruit and select relevant exportins and active kinases. Normal Ran cycling is irreversibly disrupted because L_M localizes to the RanBP1 site, excluding it from hydrolysis pathways. This unique modulation of Ran effector selection is, to our knowledge, the first structure description of nucleocytoplasmic trafficking perversion by a pathogen protein that targets Ran. Potential applications include antiviral drug targets and cancer cell division therapeutics.

Author contributions: V.R.B.-D. and A.C.P. designed research; V.R.B.-D., J.J.C., H.A.B., and C.C.C. performed research; V.R.B.-D., J.J.C., H.A.B., and C.C.C. contributed new reagents/analytic tools; V.R.B.-D., J.J.C., H.A.B., C.C.C., and A.C.P. analyzed data; and V.R.B.-D. and A.C.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. B.L.S. is a guest editor invited by the Editorial Board.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank (PDB), www.pdb.org (PDB ID codes 2M7Y, 2MMG, 2MMH, 2MMI, 2MMK, and 2MML) and Biological Magnetic Resonance Bank (BMRB), www.bmrb.wisc.edu (accession nos. 19854, 19084, 19855, 19857, and 19858).

[1]To whom correspondence should be addressed. Email: acpalmen@wisc.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1411098111/-/DCSupplemental.

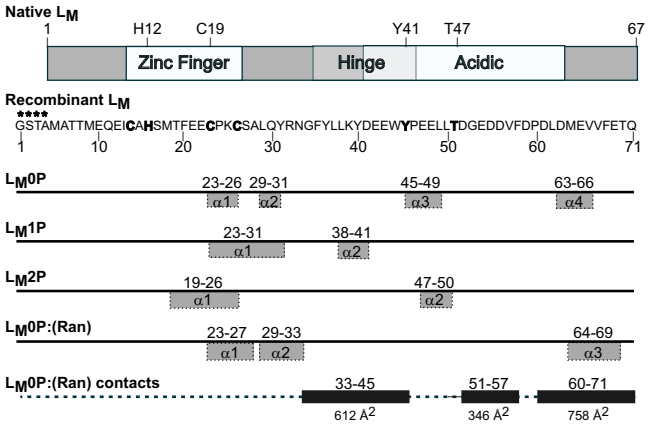

Fig. 1. L_M schematics. Protein map of native L_M shows motifs and the Y_{41}, T_{47} phosphorylation sites; the sequence of L_M as determined by NMR is 4 aa longer (****) than the native protein at the NH_2 end. NMR-determined α-helix motifs were defined by TALOS+ for respective structures (Fig. S5*A*). The remainder of each protein is random coil. The inclusive residue segments that shift on binding with Ran are indicated with the extent of their contact surfaces (Å^2) in the docked (Ran):L_M0P structure.

hinge region (aa 23–36), have been mapped as the primary contact point(s) for interactions with Ran, in what is presumed to be an induced-fit binding (15). In cells or via recombinant proteins, saturation binding with L_M is best achieved when Ran is aided by catalytic amounts of its cognate guanine nucleotide exchange factor 1 (RCC1), allowing it to morph between GTP- and GDP-bound conformers (3). Complicating a resolution of the full L-dependent antiviral mechanism are observations that L_E is itself phosphorylated during EMCV infection, in sequential reactions with casein kinase 2 (CK2) and spleen tyrosine kinase (Syk) at residues T_{47} and Y_{41}, respectively (16). Although not required for Ran interactions, the L_E modifications are clearly important to the virus because mutation at these same sites prevents subsequent Nup phosphorylation (5), suppresses NF-κB activation, and restricts infection-dependent IFN I stimulation (IRF-3 inhibition) (10–12).

The initial solution structure of L_M did not indicate how phosphorylation might affect the conformation of this protein, influence Ran binding, or contribute to activity of the Ran:L_M complex. Accordingly, we carried out, matched NMR studies on the unphosphorylated (L_M0P), singly phosphorylated (L_M1P, T_{47}), and doubly phosphorylated (L_M2P, Y_{41}/T_{47}) variants of recombinant L_M. In addition, the solution structures of L_M (labeled) as bound to Ran (unlabeled) and of Ran (labeled) as bound to L_M (unlabeled) were resolved by NMR and docked to each other. The combined datasets clearly define the Ran:L_M interfaces available for ternary interactions. Pull-down assays with GST-L_M and mutant phosphorylation derivatives, combined with previously resolved structures of Ran binding partners, predict the L_M phosphorylation interface, and the L_M zinc finger mediate Ran's selection of exportins like Crm1 or CAS and their respective (kinase?) cargos.

Results

L_E Phosphorylation Sites. The L_E (EMCV) and L_M (Mengo) proteins differ by a single substitution (L_{14} vs. M_{14}, respectively), but at the nucleotide level, convenient restriction sites make it easier to manipulate L_E rather than L_M sequences. The zinc finger motif of L_M has been described by NMR (14), but before extending this work to the fully phosphorylated protein, it seemed prudent to confirm the kinase specificities. A panel of 12 GST-L_E proteins was prepared, with alterations at every Thr and Tyr residue. The double mutation $Y_{41}F/T_{47}A$ was also included. The data, summarized in Table S1, confirmed previous reports (16) that CK2 uniquely recognizes T_{47}. This reaction is an obligate prerequisite to the single-site Syk phosphorylation at Y_{41}, a requirement that can be bypassed only if T_{47} position is substituted with a phosphomimetic aspartate ($T_{47}E$) or glutamate (16). The lack of phosphorylation at the Y_{41}/T_{47} mutated sites is not due to protein misfolding (15).

L_M NMR Determinations. Previous attempts to determine an L_M solution structure were confounded by contaminant heavy metals with affinity for the protein acidic domain (14). The problem was solved by treating samples with EDTA after the removal of the GST tag and then refolding by gradual addition of $ZnCl_2$. Dialysis removed exogenous zinc, a requirement for subsequent cosolubility with Ran. The preferred recombinant configurations extended the native L_M sequence (67 aa) by 4 aa (Gly-Ser-Thr-Ala) at the amino terminus (Fig. 1). Such extensions do not affect L_E activity (3, 7). Single (CK2) or double (CK2/Syk) phosphorylation reactions preceded the EDTA step.

L_M0P, L_M1P, and L_M2P samples ($^{15}N/^{13}C$) were investigated by high-field ^1H, ^{15}N, ^{13}C, ^{31}P NMR spectroscopy (*SI Materials and Methods* and Figs. S1–S4). Superimposition of 2D [^{15}N, ^1H]-HSQC spectra showed distinct peak changes, indicating global chemical shifts on protein phosphorylation (Fig. S3*A*). Residue-specific backbone assignments were obtained by cross-referencing the 2D [^{15}N, ^1H]-HSQC of all three proteins, as well as 3D HNCACB, 3D CBCA(CO)NH, 3D ^{15}N-NOESY, and 3D ^{13}C-NOESY spectra, using CARA analysis to verify sequential connectivity. For all proteins, backbone resonance assignments could be obtained for 100% of the 71 residues. The buffer conditions required for solubility, elevated the ^{31}P background signals to the extent that these particular peaks had to be normalized to maximum resonance levels to obtain good resolution. No above-background ^{31}P signal was identified in L_M0P protein samples, confirming that bacterial expression did not add phosphates. After treatment with CK2 or CK2/Syk, one or two additional major ^{31}P peaks were identified for L_M1P and L_M2P, at high resolution in isolated 1D ^{31}P ppm spectral regions, confirming the MS results showing that >80% of L_M1P was phosphorylated and >60% of L_M2P was doubly phosphorylated. The resolution statistics for all three proteins are summarized in Table 1. Fig. S5*B* shows restraints per residue. Tables S2 and S3 record determined values for distance restraint reliabilities and structure quality. The 10 lowest energy structures for L_M0P, L_M1P, and L_M2P are deposited in PDB (ID codes 2MMH, 2MML, and 2MMK). Their corresponding data are available from Biological Magnetic Resonance Bank (BMRB) (accession nos. 19084, 19858, and 19857). The bound zinc ion was modeled into the structures.

L_M0P/1P/2P Comparisons. In the residue numbering for recombinant L_M (native +4), the zinc finger (aa 14–26), hinge region (aa 35–45), and acidic domain (aa 41–65) form phenotypic landmarks defined by mutagenesis studies (15). T_{51} and Y_{45} are the phosphorylation sites (16). When L_M0P, L_M1P, and L_M2P were superimposed, the overall root mean square deviation (RMSD) was 12.8 Å for all backbone atoms. The majority of L_M0P was random coil, interspersed with four short helical segments (Figs. 1 and 2 *A* and *B*). As defined by TALOS+ algorithms (17), the α1 and α2 segments (aa 23–26 and 29–31) spanned the COOH-half of the zinc finger. The α3 and α4 motifs, in the hinge region (aa 45–49) and near the COOH tail (aa 63–66), were less well defined.

Surprisingly, when one or two phosphates were added, the overall proteins still configured largely as random coil (Fig. 1 and Fig. S6 *B–D*). Superimposed the L_M1P (Fig. 2 *C* and *D*) and L_M2P (Fig. 2 *E* and *F*) final states, oriented by their zinc fingers, had RMSD values of 11.8 and 13.7 Å, respectively. The phosphates did not reorganize their immediate locales, and changes in the COOH halves of the proteins were unexceptional. Instead,

MICROBIOLOGY

Table 1. NMR restraints and structural statistics

Measurements	L_M0P	L_M1P	L_M2P	(RAN): L_M0P	Ran: (L_M0P)
Total distance restraints	256	129	353	397	600
Number of torsion angle dynamics steps	5,000	5,000	5,000	5,000	5,000
Number of structures initial: 50	Final: 10	Final: 10	Final: 10	Final: 10	Final: 10
Hydrogen bonds	42	62	75	76	364
Total dihedral angle restraints	62	90	90	94	154
ϕ	31	45	45	47	77
ψ	31	45	45	47	77
Restraint violations					
Distance restraint violation > 0.2 Å	None	None	None	None	None
Angle restraint violation > 5.0°	None	1	None	None	5
Average RMSD (Å) among the 10 refined structures					
Residues	1–71	1–71	1–71	1–71	1–216
Backbone residues	2.0	1.2	3.1	1.0	4.8
Ramachandran statistics of 10 structures (% residues)					
Most favored regions	94.2	98.4	96.8	98.6	94.4
Additional allowed regions	5.8	1.6	3.2	0	2.8
Disallowed regions	0	0	0	1.4	2.8

All values were generated by CYANA for structure determination before Xplor-NIH refinement. Reported RMSD values rounded to 0.1 Å.

the upstream zinc finger regions now showed significantly less motion among the modeled structures, to the extent that in both cases, these regions stiffened into a single contiguous helix (Figs. 1 and 2 *G* and *H*). For L_M2P, the increased rigidity partially extended into the hinge and acidic domains, making all of the states more compact and easier to superimpose throughout their lengths. The zinc finger RMSD values were 0.6 and 1.6 Å, respectively. Although compressed topology was the most noticeable structural characteristic of progressive L_M phosphorylation, in none of the states, for any of the proteins, was there evidence of direct interactions among the defined phenotypic domains. Notably, however, L_M phosphorylation did affect residue H_{16} of the zinc finger, showing solution oscillations of 5–6 Å in L_M1P and 4–5 Å in L_M2P, relative to state-1. The H_{16} motility suggests alterations in additional zinc finger contacts and its putative associations as a direct result of L_M phosphorylation. As a rule, the zinc finger and acidic regions were separated in rough U-shaped conformations with independent faces, presumably available for different induced-fit binding partners.

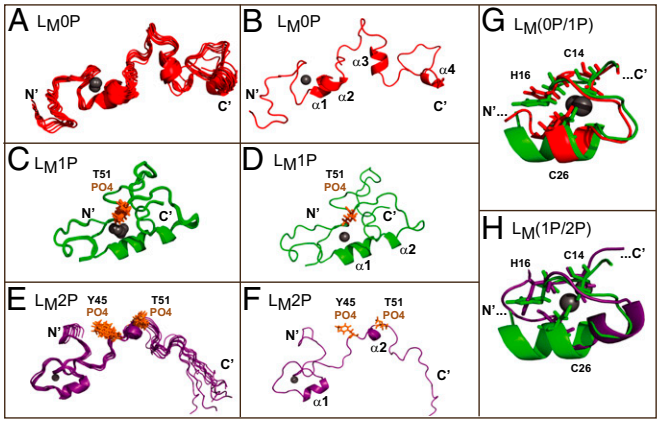

Fig. 2. Solution structures of L_M(0P/1P/2P). (*A*, *C*, and *E*) The 10 lowest-energy states for L_M0P, L_M1P, and L_M2P as free solution structures are shown. These ensembles are as deposited with PDB. (*B*, *D*, and *F*) The state-1, lowest energy structure for each protein is labeled with determined motifs. (*G*) Superimposition of the zinc-finger regions of L_M0P and L_M1P highlight observed rearrangements. (*H*) Similarly, superimposition of L_M1P and L_M2P zinc finger regions show conformational changes centering on the zinc binding domain, particularly H_{16}. In all panels, the zinc ion is a gray sphere.

Ran:(L_M0P) NMR Determinations. The main binding partner for cardiovirus L protein, RanGTPase, is insensitive to the phosphorylation status of L_M or L_E (16). Likewise, L_M(0P/1P/2P) is insensitive to the nucleotide status of Ran (GTP, GDP, and unbound) as long as the binding mix contains catalytic amounts of RCC1, a natural nuclear auxiliary factor that helps Ran morph among its conformers (3). Simultaneous resolution of a full Ran:L_M0P complex (216 and 71 aa) tests the practical limits of NMR, so paired combinations of labeled ($^{15}N/^{13}C$) and unlabeled proteins were analyzed in parallel, under identical conditions to the single L_M determinations. The native (unbound) solution structure of nucleotide-free Ran will be described in detail elsewhere (PDB ID code 2MMC and BMRB accession no. 19852). This dataset aided the assignment of 100% of the 216 Ran resonance peaks from the docked complex(es). The resolution statistics for Ran:(L_M0P) are summarized in Table 1. Tables S2 and S3 record determined values for distance restraint reliabilities and structure quality.

The 10 lowest energy states for Ran, as bound by L_M0P (PDB ID code 2MMG and BMRB accession no. 19854) showed a six-sheet β-propeller core structure, interspersed with nine α-helices (Fig. 3*C*), characteristic of other described crystal structures (18, 19). Relative to each other (Table S4), the core states (aa 8–176) were tight (RMSD, 0.3 Å), but the full protein value (RMSD, 4.6 Å) was higher because the COOH-tails (aa 177–216) in each state displayed as flexible, floppy arches (Fig. 4*A*). All these tails (RMSD, 4.9 Å) had the same central helix (aa 196–206), but none were similarly oriented relative to the core. Among Ran structures solved by crystallography, the COOH tail arrangements can vary according to nucleotide status and binding partner-induced shifts that may also involve the nucleotide-proximal phosphate binding P-loop (aa 16–25), Switch 1 (aa 32–45), Switch 2 (aa 66–79), and basic patch (aa 139–142) internal core segments (Fig. 3) (20). When L_M0P bound to Ran, the spectra recorded 36 changes of amide ^1H and ^{15}N chemical shifts within a defined subset of residues, including D_{18}, T_{21}, and K_{23} of the P-loop, E_{36} of the Switch I domain, Q_{69} and Y_{80} of the Switch II domain, H_{139} and R_{140} of the basic patch, and D_{211} and D_{213} of the acidic tail. Several of these locations, particularly in the COOH tail, were previously predicted by mutational mapping, as essential to L_M interactions (15).

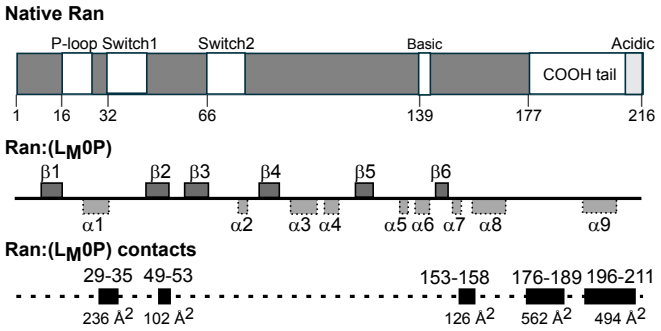

Fig. 3. Ran schematics. Native Ran protein is mapped with key activity switch elements. NMR-determined α-helix and β-sheet motifs were defined by TALOS+, for Ran as complexed with L_M0P (Fig. S5*A*). Inclusive residue segments that shift on binding with L_M0P are indicated with the extent of their contact surfaces (Å^2) in the docked Ran:(L_M0P) structure.

To best describe the solution structure of Ran as bound by L_M0P, the state-1 coordinates were aligned pairwise with representative PDB entries (Table S4). Only a few such structures have resolved COOH tails, so it was not unexpected that the full-length comparisons (all), or comparisons specific for this region (COOH), showed variability (RMSD, ~1.5–12.5 Å). The core region comparisons, however, more closely aligned Ran:(L_M0P) with the known GTP-dependent conformers (RMSD, ~1.5 Å) as opposed to GDP- or nucleotide-free forms (RMSD, ~3.9 Å). Among these, the core coordinates of PDB ID code 1K5G fit the Ran:(L_M0P) state-1 to a remarkable degree. The overall RMSD (0.4 Å) between these structures showed very low variability in all backbone residues, including the P-loop (0.2 Å), Switch 1 (0.3 Å), and Switch 2 (0.3 Å) segments. This particular dataset and the closely related PDB ID code 15KD entry describe Ran in complex with auxiliary factors RanBP1 and RanGAP, in a GTP-ground state and in a hydrolysis transition state mimic (21). The NMR solution structure of Ran, as bound by L_M0P, exhibits essentially the same core coordinates.

(Ran):L_M NMR Determinations. When the $^{15}N/^{13}C$ protein labels were switched in Ran:L_M0P complexes, the states of L_M0P as influenced by Ran showed multiple peak shifts relative to free L_M0P (Fig. 1 and Fig. S3*A*). The shifts included all residues in the hinge region, as had been anticipated from mutagenesis mapping (15). Also involved were regions from the carboxyl third of the protein. The zinc finger region did not change, maintaining the α1 and α2 helices. However, as with both phosphorylation datasets, the rest of (Ran):L_M0P now became more compact (Fig. 4*B*). The 10 low-energy states (PDB ID code 2MMI and BMRB accession no. 19855) remained predominantly random coil, with an average RMSD of 4.6 Å for backbone atoms. The fit with Ran was clearly induced by mutual binding.

The state-1, low-energy coordinate sets for Ran:(L_M0P) and (Ran):L_M0P were evaluated for fit according to GRAMM-X (22) and HADDOCK algorithms (23) without specified constraints. Previous (pseudo) dockings (15) pairing the initial solution structure of L_M (PDB ID code 2M7Y) with Ran (PDB ID code 1K5G) predicted the interactions at the Ran:BP1 binding face with the Ran COOH tail wrapping around L_M0P to hold it firmly onto this surface. The real solution structures indeed followed this pattern. For the HADDOCK outputs, the best cluster (models 1–4) had an RMSD of 0.5 Å, with an E-total of 305–343 kcal/mol for the interface. As depicted in the optimal energy model (model 1), L_M0P sits tightly on the top surface of Ran, without altering the Ran core, or approaching the nucleotide binding pocket (Fig. 4*C* and Fig. S6). The L_M0P hinge and acidic domains interact significantly with the proximal tip of the Ran COOH tail (aa 203–210), but the remainder of this segment is free to arch without steric hindrance, morphing and encircling central interaction residues of L_M0P (Fig. 4*D*). Fundamentally, this orientation looks very similar to Ran:BP1 complexes as they are presented in determined crystal structures (e.g., PD ID code 1K5G). The buried Ran:L_M0P interface covers ~1,700 Å^2, including about 28% of the L_M0P residues and 13% of the Ran residues (Figs. 1 and 3*D*). Extensive hydrogen bonding (>20×) and salt bridges (e.g., L_M0P K_{34} vs. Ran E_{34}) readily account for the 3 nM K_D (15). Important Ran:L_M0P contacts include T_{32}:L_{37}, A_{183}:D_{41}, P_{184}:Y_{45}, P_{185}:E_{42}, Q_{196}:Y_{36}, Y_{197}:Y_{40}; Y_{197}:W_{44}, A_{204}:G_{34}, A_{204}:L_{38}, and T_{207}:N_{33} (Fig. 4*E*). These protein placements are fully consistent with both determined Ran:(L_M0P) and (Ran):L_M0P NMR datasets and all resonance shifts relative to the unbound proteins (Tables S2 and S3). It explains the low RMSD for the core of bound Ran, flexibility of the Ran COOH tail, and requirements that the L_M0P zinc finger domain make no contacts with Ran that would prevent it from folding like the native L_M0P protein. In this configuration, both L_M0P phosphorylation sites are solvent exposed on the same face as the zinc finger, even though the loops which display and orient them form key Ran contacts.

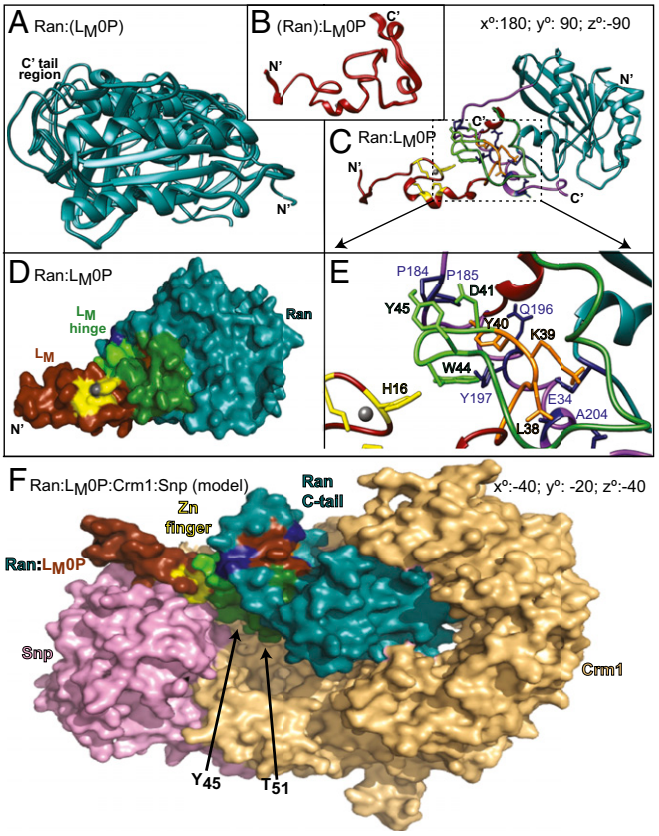

Fig. 4. Solution structure of Ran:L_M0P. (*A*) Ten low-energy states for Ran:(L_M0P). (*B*) Ten low-energy states for (RAN):L_M0P. All coordinates are as deposited with PDB. (*C*) HADDOCK-determined state-1 of Ran:(L_M0P) docked to state-1 (Ran):L_M0P. C′ tail region of Ran, wrapping around L_M0P is highlighted (purple). Ran rotation relative to *A* is indicated. Stereo version of this image is in Fig. S6. (*D*) Similar to *C*, Ran (blue), L_M0P (brown), zinc coordination residues (yellow), and hinge region contacts (green) are highlighted. (*E*) Similar to *C* and *D*, close-up shows orientations of key Ran:L_M0P interaction regions. (*F*) State-1 Ran:(L_M0P) coordinates were aligned and then substituted into PDB ID code 3GJX, a crystallographically determined complex of Ran (blue), Crm1 (tan), and snurportin1 (pink). Relative to *D*, the required rotation for Ran:L_M0P is indicated. The loop of Ran (blue) encircling L_M0P (brown, yellow, green) is the C′ tail. The L_M phosphorylation sites are buried in the Crm1 (T_{51}) and Snp (Y_{45}) interfaces.

Discussion

Cardiovirus L proteins are extraordinarily toxic to cells because their presence triggers massive hyperphosphorylation of Phe/Gly nuclear pore proteins (Nups). In cell-free assays with intact nuclei, within 5 min of the introduction of recombinant L_E-GST (or GST-L_E), there is complete inhibition of all active import and export of host proteins and RNA through the nuclear pore complexes (NPCs) (24). The discovery that L_E bound RanGTPase, the key regulator of nucleocytoplasmic trafficking (NCT), raised the initial possibility of putative stoichiometric inhibition. This idea was quickly discarded because Ran is an abundant protein (25), and only tiny amounts of L_E are required to trigger this effect. Instead, Ran:L_E binding is leveraged by consequent activation of a potent Nup phosphorylation cascade, the true cause of trafficking inhibition. This inhibition happens in infected cells even before the virus begins to replicate (6, 26). The cascade involves Erk1/2 and p38 kinases and is absolutely dependent on Ran:L_E interactions and also on the dual phosphorylation of L_E itself, an activity that is a prerequisite, and not a consequence, of the Nup modifications (7, 10, 15, 16, 27). Furthermore, the L_E zinc finger motif must remain intact and chelated to the metal for the protein to function (11). These points were clearly established with extensive activity assays, mutagenesis, and biochemical studies.

Because neither L_E nor Ran is a kinase, an obvious ensuing step must involve recruitment of one or more critical ternary/quaternary partners. The identification of these elements is underway. We are focusing on plausible pathways by which Erk1/2 and p38 can be diverted from their normal activities to act on Nups. However, because native Ran has many interaction partners, and L_E must be phosphorylated, sorting out precise steps is complicated. To aid in this process, as described here, we resolved the NMR solution structures of L_M, its phosphorylated derivatives, and the Ran:L_M0P complex. The studies had three goals: (*i*) determine whether phosphorylation significantly altered the structure of L_M; (*ii*) determine the format of Ran as bound by L_M, so that germane native binding partners could be evaluated; and (*iii*) determine the segments of L_M, not impacted by Ran, and therefore accessible to later interactions.

The L_M(0P/1P/2P) datasets showed this protein, in a free format, does not have a very organized secondary structure, except for the zinc finger domain (Fig. S6). Phosphorylation provided important constraints on the degree of random coil motion but did not by itself induce an overt restricted format. If conformation plays a role in L_M activity, outside of the zinc finger, it must be induced by the relevant binding partners. Indeed, when bound to Ran, L_M0P condensed and made specific contacts in the central hinge and acidic domains, via the same residues identified by mutagenesis (15, 24). Surprisingly, however, the amino third of the protein, including the zinc finger, and both internal phosphorylation sites were left solvent exposed. It was not expected that all these sites would localize to the same exposed face.

The Ran:L_M0P complex is the first solution structure for Ran and the first to describe the intact full-length protein. At least 45 Ran datasets have been collected and resolved by crystallography, but in almost all cases, 4–9 amino-terminal residues and >40 carboxyl-terminal residues are unresolved or were not included in the determinations. The entries differ in nucleotide-bound status and coresolution of diverse transport-related binding partners. By NMR, it was clear that L_M0P binding induced a Ran conformer almost identical to that assumed when Ran binds to BP1, a cytoplasmic auxiliary factor (e.g., PDB ID code 1K5G). Although the Ran:(L_M0P) complex was nucleotide free, the P-loop, Switch 1, and Switch 2 regions were set to the typical GTP formats, as they would be naturally, whenever Ran exits the nucleus, usually bound to an exportin, and makes initial cytoplasmic contacts with BP1 and RanGAP (21). L_M and BP1 do not share sequence similarity, yet on Ran, they occupy similar footprints and their binding is mutually exclusive (24). Not captured by crystallography, but very apparent by NMR, was the dynamic morphing of the Ran COOH tail over the top of this binding partner face. L_M0P made important contacts with the beginning and end of this segment, but the resolved states recorded considerable movement here for both proteins.

How then does this conformation allow the Ran:L_M complex to form its next interactions? Complete phosphorylation by CK2 (T_{51}) and Syk (Y_{45}) have been demonstrated after the pair is bound (16). The observed proximity of these sites to the zinc finger, as with the solution structures of L_M1P and L_M2P, might conceivably influence the rigidity or orientation of this domain when bound to Ran. More likely, it is a combination of all these factors on this exposed L_M face, working with Ran, now locked in a GTP format, that select the next partner. Our preliminary experiments suggest exportins, like Crm1 or CAS, are likely candidates. L_E-GST can extract both native proteins from HeLa cell extracts in reactions that show a strong dependence on the L_E phosphorylation status (Fig. 5). Mutations in either or both of the phosphorylation sites diminished the L_E binding. However, because these pull-downs are from extracts, it is not yet known whether similar exportin:L_E interactions are codependent, obligate, or independent of simultaneous Ran:L_E reactions. To work this out experimentally will require considerable validation of stepwise protocols, including the sequential addition of phosphates, the proper nucleotide status of Ran, demonstration of an active exportin conformation, and putative cargo inclusion.

As a guide for these parameters, we used molecular replacement algorithms to test a putative Ran:L_M0P docking into the context of a determined Ran:Crm1:cargo structure. The selected template (PDB ID code 3GJX) included snurportin1 as the Crm1 cargo (28). Obviously snurportin is not relevant to the Erk1/2 and p38 Nup phosphorylation pathways, but its location helps orient the participants. When native Ran binds an exportin, it must be in the GTP format (29), as it is for Ran:L_M0P, albeit in our complex, the nucleotide status is forced artificially by the L_M0P protein interactions. Substitution of the NMR-determined Ran:L_M0P for the crystallographically determined Ran into this structure (Fig. 4*F*) did not create steric clashes. All described Ran:Crm1 contacts were maintained (Table S4; PDB ID code 3GJX), and even each of the morphing NMR derivatives of the Ran COOH tail were without conflicts. Of importance, this enforced orientation of L_M0P placed the phosphorylation and zinc finger face into the immediately

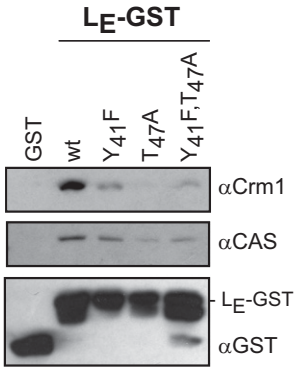

Fig. 5. Exportin pulldown by L_E-GST. L_E-GST proteins (or GST alone) were incubated with HeLa cell cytosol then reacted with glutathione-conjugated beads in assays identical to those previously described (24). The extracted proteins were fractionated by SDS/PAGE and then identified by Western analyses. The antibodies included αCrm1 (Abcam 24189; 1:2,000) αCAS (sc-1708, 1:500; Santa Cruz Biotechnology), and αGST (GE Healthcare; 1:10,000). L_E-GST proteins made this way (16) can show additional bands from alternative translational start sites. Similar pull-down assays were used in the original identification of L:Ran interactions (24), albeit protein detection was with different antibodies.

proximity of both Crm1 and snurportin surfaces. T_{51} points toward Crm1 and Y_{45}, plus the zinc finger is oriented toward the cargo. Obviously this model is only speculative, but it suggests working hypotheses that can now be tested. For example, the model predicts the L_M phosphorylation status may help determine whether Ran:L_M:Crm1 ternary (or quaternary?) complexes can be formed. It also predicts that the L_M zinc finger motif and its nearby phosphorylated residues may restrict or determine putative cargo selection, perhaps including the active Nup phosphorylation kinases themselves. These possibilities are under further investigation with binding, pull-down, mutagenesis, and reconstruction experiments.

Materials and Methods

L_E and L_M Proteins. Recombinant L_E-GST (EMCV) and mutated derivatives were expressed and purified from *Escherichia coli* as previously described (6, 15, 24). GST-L_M (Mengo) has also been previously described (14). The protein includes a thrombin cleavage site for GST-tag removal. Uniform, dual labeled [^{15}N/^{13}C]-L_M0P was produced from BL-21 (DE3) cells transformed with pGST-L_M. The required media, induction procedures, protein isolation (GST Trap column), removal of the GST tag, gel filtration, concentration, EDTA treatment, and refolding by dialysis are detailed in *SI Materials and Methods*. Protein purity (>95%) was determined by SDS/PAGE followed by silver stain. Phosphorylation with CK2 alone (L_M1P) or CK2 followed by Syk (L_M2P) is described in *SI Materials and Methods*. All labeled or unlabeled materials were assayed for molecular weight (SDS/PAGE, MALDI-MS), proper folding by ^{1}H spectra (15), and biological activity (24) before structure determination.

Ran Proteins. 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, as expressed in BL21 cells, was as previously described (15). Labeled [^{15}N/^{13}C] preparations were similar to L_M0P, except for the inclusion of ampicillin (50 μg/mL) in the medium. Ran purification, as summarized in *SI Materials and Methods*, has been previously described (15), but then, 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), followed by a second dialysis into fresh NMR buffer (overnight, 4 °C). 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 (*Xenopus laevis*) was purified as previously described (30) 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) L_M(0P/1P/2P) and Ran was 0.5 mM. When complexes were probed, 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} = 150 ms), and 3D ^{13}C-NOESY (t_{mix} = 140 ms) were collected on a Bruker DRX-600 spectrometer equipped with a ^{1}H, ^{13}C, ^{15}N, ^{31}P three-axis gradient cryogenic probe. The techniques and algorithms used to process the raw data are detailed in *SI Materials and Methods*. The information includes a process workflow chart (Fig. S1) backbone and side-chain assignments (Figs. S2 and S4), dihedral angle constraint files (Figs. S3 and S5*B*), structure calculations, nonstandard amino acids identification, motif location, processing command lines (Dataset S1), and data refinement (Fig. S5*A*). The quality of each generated structure (Table S3) was analyzed for restraint and geometry violations using the Duke University MolProbity web server (31, 32). The NMR restraints and structural statistics (Table 1) were generated by CYAN before Xplore-NIH refinement (33). 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. The lowest energy NMR states for L_M0P and Ran from the docked complexes were submitted to HADDOCK via the public web portal (23). 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 Protein Interactions Calculator (PIC) (34).

ACKNOWLEDGMENTS. We thank C.C.C. for providing access to the unpublished L_M dataset (PDB ID code 2M7Y). This work was supported by NIH Grant AI-17331 (to A.C.P.) and NIH Molecular Biosciences Training Grant PRJ 21 CV A347300 (to V.R.B.-D.). V.R.B.-D. was also supported by a Science and Medicine Graduate Research Scholars Fellowship from the University of Wisconsin–Madison. This study made use of the National Magnetic Resonance Facility at Madison, which is supported by NIH Grant P41GM66326. Equipment was purchased with funds from the University of Wisconsin–Madison; NIH Grants P41RR02301, P41GM66326, S10RR02781, S10RR08438, S10RR023438, S10RR025062 and S10RR029220; National Science Foundation Grants DMB-8415048, OIA-9977486, and BIR-9214394; and the US Department of Agriculture.

1. Adams MJ, King AM, Carstens EB (2013) Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2013). *Arch Virol* 158(9):2023–2030.
2. Rueckert RR, Wimmer E (1984) Systematic nomenclature of picornavirus proteins. *J Virol* 50(3):957–959.
3. Petty RV, Palmenberg AC (2013) Guanine-nucleotide exchange factor RCC1 facilitates a tight binding between the encephalomyocarditis virus leader and cellular Ran GTPase. *J Virol* 87(11):6517–6520.
4. Wennerberg K, Rossman KL, Der CJ (2005) The Ras superfamily at a glance. *J Cell Sci* 118(Pt 5):843–846.
5. Lidsky PV, et al. (2006) Nucleocytoplasmic traffic disorder induced by cardioviruses. *J Virol* 80(6):2705–2717.
6. Porter FW, Palmenberg AC (2009) Leader-induced phosphorylation of nucleoporins correlates with nuclear trafficking inhibition by cardioviruses. *J Virol* 83(4):1941–1951.
7. Porter FW, Brown B, Palmenberg AC (2010) Nucleoporin phosphorylation triggered by the encephalomyocarditis virus leader protein is mediated by mitogen-activated protein kinases. *J Virol* 84(24):12538–12548.
8. Bardina MV, et al. (2009) Mengovirus-induced rearrangement of the nuclear pore complex: Hijacking cellular phosphorylation machinery. *J Virol* 83(7):3150–3161.
9. van Pesch V, van Eyll O, Michiels T (2001) The leader protein of Theiler's virus inhibits immediate-early alpha/beta interferon production. *J Virol* 75(17):7811–7817.
10. Zoll J, Melchers WJ, Galama JM, van Kuppeveld FJ (2002) The mengovirus leader protein suppresses alpha/beta interferon production by inhibition of the iron/ferritin-mediated activation of NF-kappa B. *J Virol* 76(19):9664–9672.
11. Hato SV, et al. (2007) The mengovirus leader protein blocks interferon-alpha/beta gene transcription and inhibits activation of interferon regulatory factor 3. *Cell Microbiol* 9(12):2921–2930.
12. Ricour C, et al. (2009) Inhibition of mRNA export and dimerization of interferon regulatory factor 3 by Theiler's virus leader protein. *J Gen Virol* 90(Pt 1):177–186.
13. Borghese F, Michiels T (2011) The leader protein of cardioviruses inhibits stress granule assembly. *J Virol* 85(18):9614–9622.
14. 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.
15. Bacot-Davis VR, Palmenberg AC (2013) Encephalomyocarditis virus Leader protein hinge domain is responsible for interactions with Ran GTPase. *Virology* 443(1):177–185.
16. 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.
17. Delaglio F, et al. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277–293.
18. Neuwald AF, Kannan N, Poleksic A, Hata N, Liu JS (2003) Ran's C-terminal, basic patch, and nucleotide exchange mechanisms in light of a canonical structure for Rab, Rho, Ras, and Ran GTPases. *Genome Res* 13(4):673–692.
19. Cherfils J, Zeghouf M (2013) Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiol Rev* 93(1):269–309.
20. Nilsson J, Askjaer P, Kjems J (2001) A role for the basic patch and the C terminus of RanGTP in regulating the dynamic interactions with importin beta, CRM1 and RanBP1. *J Mol Biol* 305(2):231–243.
21. Seewald MJ, Körner C, Wittinghofer A, Vetter IR (2002) RanGAP mediates GTP hydrolysis without an arginine finger. *Nature* 415(6872):662–666.
22. Tovchigrechko A, Vakser IA (2006) GRAMM-X public web server for protein-protein docking. *Nucleic Acids Res* 34(Web Server issue):W310-4.
23. de Vries SJ, van Dijk M, Bonvin AMJJ (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc* 5(5):883–897.
24. 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.
25. Görlich D, Panté N, Kutay U, Aebi U, Bischoff FR (1996) Identification of different roles for RanGDP and RanGTP in nuclear protein import. *EMBO J* 15(20):5584–5594.
26. Karalyan ZA, et al. (2013) Changes in the nuclei of infected cells at early stages of infection with EMCV. *Cell Bio* 2(3):125–130.
27. Dvorak CMT, et al. (2001) Leader protein of encephalomyocarditis virus binds zinc, is phosphorylated during viral infection, and affects the efficiency of genome translation. *Virology* 290(2):261–271.
28. Monecke T, et al. (2009) Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. *Science* 324(5930):1087–1091.
29. Askjaer P, Jensen TH, Nilsson J, Englmeier L, Kjems J (1998) The specificity of the CRM1-Rev nuclear export signal interaction is mediated by RanGTP. *J Biol Chem* 273(50):33414–33422.
30. 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.
31. 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.
32. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12–21.
33. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 160(1):65–73.
34. Tina KG, Bhadra R, Srinivasan N (2007) PIC: Protein interactions calculator. *Nucleic Acids Res* 35(web server issue):W473-6.