HIV Rev response element (RRE) directs assembly of the Rev homooligomer into discrete asymmetric complexes

Matthew D. Daugherty[a,c], David S. Booth[b,c], Bhargavi Jayaraman[c], Yifan Cheng[c], and Alan D. Frankel[c,1]

[a]Chemistry and Chemical Biology Graduate Program, [b]Graduate Group in Biophysics, and [c]Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94158

Communicated by Keith R. Yamamoto, University of California, San Francisco, CA, June 2, 2010 (received for review February 25, 2010)

RNA is a crucial structural component of many ribonucleoprotein (RNP) complexes, including the ribosome, spliceosome, and signal recognition particle, but the role of RNA in guiding complex formation is only beginning to be explored. In the case of HIV, viral replication requires assembly of an RNP composed of the Rev protein homooligomer and the Rev response element (RRE) RNA to mediate nuclear export of unspliced viral mRNAs. Assembly of the functional Rev-RRE complex proceeds by cooperative oligomerization of Rev on the RRE scaffold and utilizes both protein-protein and protein-RNA interactions to organize complexes with high specificity. The structures of the Rev protein and a peptide-RNA complex are known, but the complete RNP is not, making it unclear to what extent RNA defines the composition and architecture of Rev-RNA complexes. Here we show that the RRE controls the oligomeric state and solubility of Rev and guides its assembly into discrete Rev-RNA complexes. SAXS and EM data were used to derive a structural model of a Rev dimer bound to an essential RRE hairpin and to visualize the complete Rev-RRE RNP, demonstrating that RRE binding drives assembly of Rev homooligomers into asymmetric particles, reminiscent of the role of RNA in organizing more complex RNP machines, such as the ribosome, composed of many different protein subunits. Thus, the RRE is not simply a passive scaffold onto which proteins bind but instead actively defines the protein composition and organization of the RNP.

nuclear export | ribonucleoprotein assembly | RNA-protein recognition

Complex retroviruses encode essential regulatory proteins that direct nuclear export of the viral RNA genome at late stages in the viral life cycle (1). In the case of HIV, Rev binds to the Rev response element (RRE), a ~350-nt structured RNA found in the introns of unspliced viral mRNAs, and interacts with the Crm1 nuclear export receptor to facilitate transport to the cytoplasm before splicing is completed (1, 2). In this way, assembly of the Rev-RRE ribonucleoprotein (RNP) complex is coupled to expression of the virion structural proteins and packaging of the genomic RNA.

In addition to RRE binding, oligomerization of Rev along the RNA is required for export (3). Early studies defined one specific site in the RRE, known as stem IIB, as necessary, but not sufficient, for in vivo function (2, 4–6). Rev binds to stem IIB using a 17-amino acid α-helical arginine-rich motif (ARM), whose interaction is well understood at the biochemical and structural levels (7, 8). However, full RNA export activity requires more than 230-nt of the ~350-nt structured RRE, as well as two Rev oligomerization domains (Fig. 1) (2, 3, 5, 6). The RRE drives assembly of a highly cooperative complex with the Rev homooligomer, with an affinity ~500-fold higher than Rev binding to stem IIB alone (9). The oligomerization domains and large RRE structure both are required for tight complex assembly in vitro, correlating with their requirement for RNA export activity in vivo and demonstrating that proper RNP formation is essential for its biological function (9). Kinetic analyses suggest an ordered assembly of Rev monomers on the RRE, initiated at stem IIB and propagated at additional sites by protein-protein interactions (10). It is not known if an oligomeric assembly also is important for other steps in the export pathway.

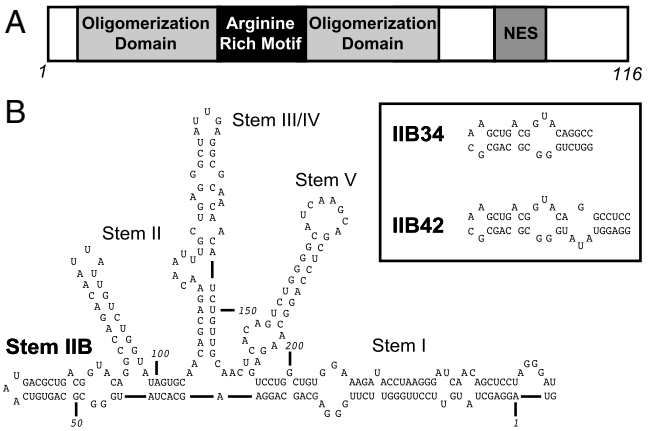

Fig. 1. Rev and the RRE. (*A*) Domain structure of Rev used in this study. (*B*) Sequence and predicted secondary structure of the 242-nt RRE and stem IIB fragments used.

The structure of the Rev ARM bound to stem IIB (PDB code: 1ETF) (7) and recent crystal structures of Rev dimers (PDB codes: 3LPH and 2X7L) (11) illuminate essential protein-RNA and protein-protein interactions that mediate Rev-RRE assembly. These results, and models generated from genetic and biochemical mapping (12), provide an initial representation of Rev-RRE structure, but a complete understanding of how the RRE organizes the functional RNP has been hampered by low protein solubility, typically limited to 1–5 μM (13, 14) and an inability to define the Rev oligomerization state (14, 15). These limitations, and the observation that Rev-RNA complexes form long filaments when prepared by denaturation-renaturation in vitro (13, 16), have called into question the existence of discrete Rev-RRE complexes. However, recent biochemical analyses with Rev purified under native conditions indicate that RNA plays a crucial role in cooperatively forming high-affinity, functional RNPs (9) and suggest that the method of Rev preparation has important consequences for assembly. We surmised that natively prepared Rev, combined with specific RNAs derived from the RRE, might form stable, defined RNPs amenable to biophysical study.

Author contributions: M.D.D. and A.D.F. designed research; M.D.D., D.S.B., and B.J. performed research; M.D.D., D.S.B., B.J., Y.C., and A.D.F. analyzed data; and M.D.D., D.S.B., and A.D.F. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: frankel@cgl.ucsf.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1007022107/-/DCSupplemental.

Here we show that RNA binding indeed drives assembly of discrete Rev-RRE complexes and provides initial structural characterizations of two complexes. We find that RNA, or RNA surrogates such as a negatively charged protein fusion or oxyanionic counterions, greatly enhance Rev solubility, and that binding to specific RRE RNAs prevents uncontrolled growth of Rev filaments and determines its precise oligomerization state. Small-angle X-ray scattering (SAXS) of a dimeric assembly intermediate and electron microscopy of the full Rev-RRE complex show that each forms defined RNP particles with distinct lobed features that lack the obvious symmetry observed in Rev dimer structures and expected for complexes assembled with a homo-oligomeric protein. These asymmetric structures are reminiscent of RNPs composed of heterooligomeric subunits, such as the ribosome, and indicate that the RRE plays an active role in defining the overall architecture of the RNP.

Results

RNA Controls the Solubility and Oligomerization State of Rev. The importance of Rev oligomerization and RNA binding for HIV RNA export is well known, but attempts to understand the molecular basis of Rev-RRE assembly have been hampered by low Rev solubility and uncontrolled oligomerization (14, 15). We recently described how Rev purified under nondenaturing conditions assembles into a cooperative, high-affinity oligomeric complex with the RRE that correlates with its activity in vivo (9). Given that these binding properties had not been described previously, we reasoned that preparing Rev-RNA complexes under similar native conditions might generate samples suitable for structural characterization.

We considered two important differences between our protocol (9) and others (3, 4, 12–15). First, we expressed Rev as a fusion to GB1, a commonly used expression tag derived from the well-folded B1 domain of streptococcal protein G (17). Second, we did not remove RNA during the purification, even after cleaving Rev from the GB1 fusion, and thus the preparation contained endogenous *Escherichia coli* RNAs. To assess the influence of each factor, we purified GB1-fused Rev (termed GB1-Rev) and removed RNA with RNases and high salt washes, typically obtaining GB1-Rev at substantially higher concentrations (50–200 μM) (Fig. S1, lane 1) than previously observed with Rev (1–5 μM). When the GB1 solubility tag was removed by proteolysis, a visible Rev precipitate rapidly formed (Fig. S1, lane 5). However, when cleaved in the presence of equimolar nonspecific RNA, no precipitate was observed (Fig. S1, lane 17), indicating that Rev solubility is enhanced greatly by adding RNA *in trans*. Interestingly, GB1 attached *in cis* prevents protein aggregation by interacting with Rev (Fig. S4*A*; see *Discussion*) and not just serving as an "inert" solubility-enhancing domain (17). We further found that potassium phosphate or sodium sulfate at 100 mM also maintains Rev in a soluble state without the appended GB1 domain (Fig. S1, lanes 9 and 13), perhaps mimicking the RNA phosphate backbone. These data suggest that solubility of the highly basic Rev protein relies on charge neutralization by binding to RNA or an RNA surrogate, such as a fused acidic protein domain or oxyanionic counterions. Each of these conditions can contribute to enhanced solubility and, importantly, provide the means to prepare samples at concentrations suitable to examine Rev-RNA structure.

We next examined how RNA affects Rev oligomerization, particularly because previous attempts to define the oligomer, either with or without RNA, resulted in differing conclusions (14, 15). GB1-Rev in the absence of RNA has an expected mass of 22 kDa but behaves as a heterogeneous, concentration-dependent mixture by size exclusion chromatography (SEC) (Fig. 2*A*). At 500 μM (11 mg/ml), GB1-Rev eluted as a wide peak, indicative of a broad distribution of oligomeric states with an average mass of 210 kDa (9–10 GB1-Rev monomers), as estimated from

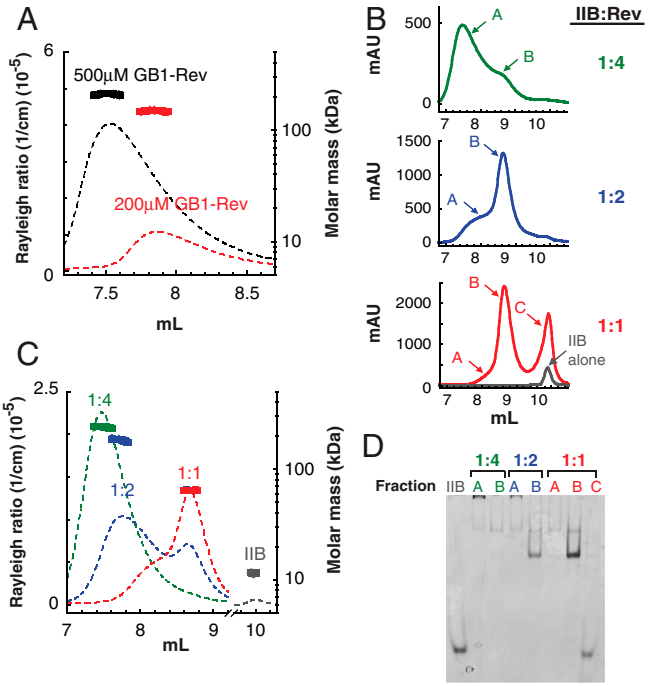

Fig. 2. RNA controls the oligomerization state of Rev. (*A*) MALS measurement from SEC of 200 μM (red) and 500 μM (black) GB1-Rev in the absence of RNA. Light scattering is shown as a function of elution volume (dashed lines, left axis). Calculated molar masses are shown for each peak (solid lines, right axis). (*B*) SEC of 200 μM GB1-Rev in the presence of 50 μM (*Top*), 100 μM (*Middle*), or 200 μM (*Bottom*) IIB34 RNA, with absorbance monitored at 260 nm. The profile of 10 μM IIB34 alone is shown in the bottom panel. Arrows indicate fractions analyzed by native PAGE (*D*). (*C*) MALS (dashed lines, left axis) and calculated molar masses (solid lines, right axis) determined from the SEC peaks in *B*, with 50 μM unbound IIB34 shown for reference. (*D*) Native gel analyses of SEC fractions (*B*), with unbound IIB34 RNA in the left lane.

coupled multiangle light scattering (MALS) and refractive index measurements (Fig. 2*A* and Table S1). At 200 μM, GB1-Rev still eluted broadly but with a slightly smaller peak mass of 150 kDa (Fig. 2 *A* and *B*, and Table S1). Thus, Rev forms heterogeneous oligomers in the absence of RNA, consistent with previous observations that Rev can aggregate into filamentous structures of varying size (13, 16, 18, 19).

Isolated hairpins of the RRE bind only Rev monomers or dimers (6, 9) and ordered Rev filaments only form in the absence of RNA (13, 16), prompting us to examine whether RNA might control the oligomer even at high protein concentrations. Indeed, adding a minimal stem IIB RNA (IIB34; expected mass of 11 kDa) to 200 μM GB1-Rev, ranging from 1:4 to 1:1 RNA: protein stoichiometry, drastically sharpened the SEC profile to a highly symmetric peak (Fig. 2*B*) and shifted it to a smaller apparent size (from 240 kDa to 63 kDa; Fig. 2*C*) despite being bound to RNA. The complexes also became increasingly distinct by native gel analyses (Fig. 2*D*). The measured mass of 63 kDa, and the observation that free RNA appears in the SEC profile at a 1:1 stoichiometry, is consistent with 2–3 GB1-Rev monomers bound to a single IIB RNA (Table S1).

Assembly of Discrete Monomeric and Dimeric Rev-RNA Complexes. To gain further insight into the coupling between Rev oligomerization and specific RNA binding, and to generate highly defined complexes for structural investigation, we wished to understand the behavior of two mutants that helped define a model in which Rev possesses two separable oligomerization surfaces (12). This study found that the L60R mutation generated largely dimeric complexes on the RRE while the L18Q mutation generated

largely monomeric complexes. We reasoned that these mutants might yield discrete Rev-RNA complexes with the stem IIB hairpin.

In the absence of RNA, mutating both oligomerization surfaces (L18Q/L60R) did not generate a homogeneous GB1-Rev monomer as might have been expected. Instead, two main species were observed (Fig. 3*A* and Fig. S2*A*), with masses of 100 kDa (4–5 monomers) and 45 kDa (2–3 monomers) (Fig. 3*A* and Table S1). Removing the GB1 domain from wild-type Rev (expected mass of 13 kDa) and binding to stem IIB RNA generated a mixture of species, with a peak corresponding to 2–3 Rev monomers per RNA, but the L18Q/L60R mutant formed a single species of 24.5 kDa (Fig. 3 *A* and *B* and Table S1), consistent with a single Rev monomer bound to IIB RNA.

We next reasoned that it might be possible to use RNA to form a discrete dimeric complex with the L60R mutation alone, which does not hinder dimer formation (12). Similar to the results above, Rev L60R formed a single species (Fig. 3 *C* and *D* and Fig. S2*B*) when bound to a 42-nucleotide stem IIB hairpin (IIB42) previously shown to cooperatively bind a Rev dimer (9). The Rev L60R-IIB42 RNA complex has a molecular mass of 39.8 kDa, consistent with a Rev dimer bound to a single RNA (Fig. 3*C* and Table S1). Thus, targeted disruption of oligomerization surfaces combined with binding to specific RNA elements can generate discrete complexes of defined size and stoichiometry.

Molecular Envelope of a Dimeric Rev-RNA Complex. To gain structural insight into the dimeric Rev-IIB complex, which may represent an early intermediate in the kinetic assembly pathway (10), we utilized SAXS to generate a low-resolution molecular envelope of the Rev L60R-IIB42 RNA complex (20). The scattering curves at three concentrations had similar shapes (Fig. 4*A*) and the Guinier plot was linear (Fig. S3), indicating that the complexes did not aggregate. The interatomic distance probability curve $P(r)$ displayed a wide maximum (Fig. 4*B*), consistent with an elongated or lobed structure, and ab initio models suggest that the complex adopts a three-lobed structure (Fig. 4*C*). The crystal structure of the Rev dimer (PDB code: 3LPH) and the NMR structure of the bound IIB34 RNA (PDB code: 1ETF) (7) fit well into the SAXS density (Fig. 4*D*), after aligning the ARM of the Rev dimer with the ARM in the NMR structure to place the RNA and extending the RNA by four base pairs. Thus, the dimeric Rev-RNA complex appears to form a lobed RNP in solution consistent with the assembled structural model.

The C Terminus of Rev Is Disordered. To generate a more complete view of the Rev complex, we used NMR to examine the C terminus of Rev, which was not present in the crystal structures or fit into the SAXS envelope. Circular dichroism and computational studies indicated the C terminus is disordered (21), but solid state NMR studies suggested some residues are helical in Rev filaments (16). In solution, the ^{1}H-^{15}N HSQC NMR spectra of the 40 kDa Rev L60R-RNA complex displayed only about 50 amide resonances of the 116 Rev residues (Fig. 5*A*), which were sharp and had poor chemical shift dispersion indicative of disordered protein regions (22). When truncated at residue 70, most of these peaks were lost (Fig. 5*A*), indicating that the resonances

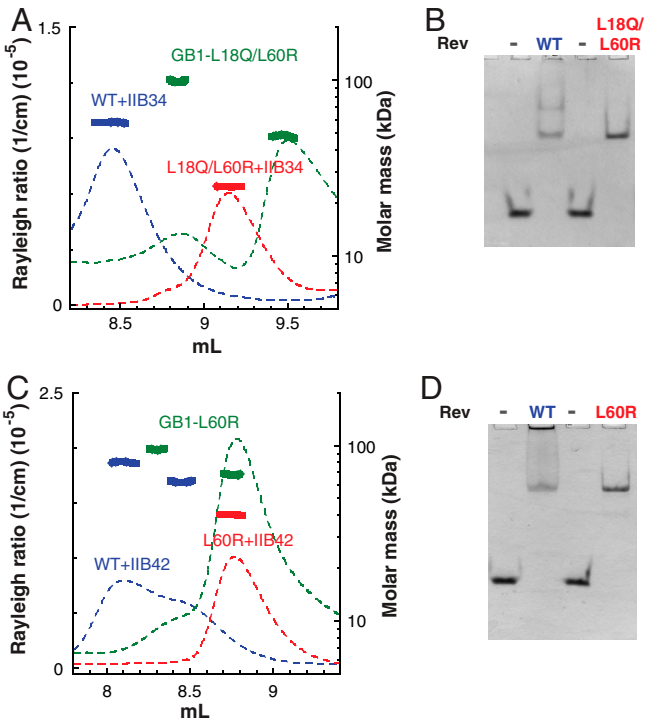

Fig. 3. Assembly of defined monomeric and dimeric Rev-RNA complexes. (*A*) Measured MALS (dashed lines, left axis) and calculated masses (solid lines, right axis) determined from SEC peaks (Fig. S2*A*) of 500 μM GB1-Rev L18Q/L60R in the absence of RNA (green), 200 μM cleaved wild-type Rev with 240 μM IIB34 RNA (blue), and 200 μM Rev L18Q/L60R with 240 μM IIB 34 RNA (red). (*B*) Native gel of unbound RNA and SEC fractions from *A*. (*C*) MALS (dashed lines, left axis) and calculated masses (solid lines, right axis) determined from SEC peaks (Fig. S2*B*) of 500 μM GB1-RevL60R in the absence of RNA (green), 200 μM cleaved wild-type Rev with 240 μM IIB42 RNA (blue) and 200 μM Rev L60R with 240 μM IIB42 RNA (red). (*D*) Native gel of unbound RNA and SEC fractions from *C*.

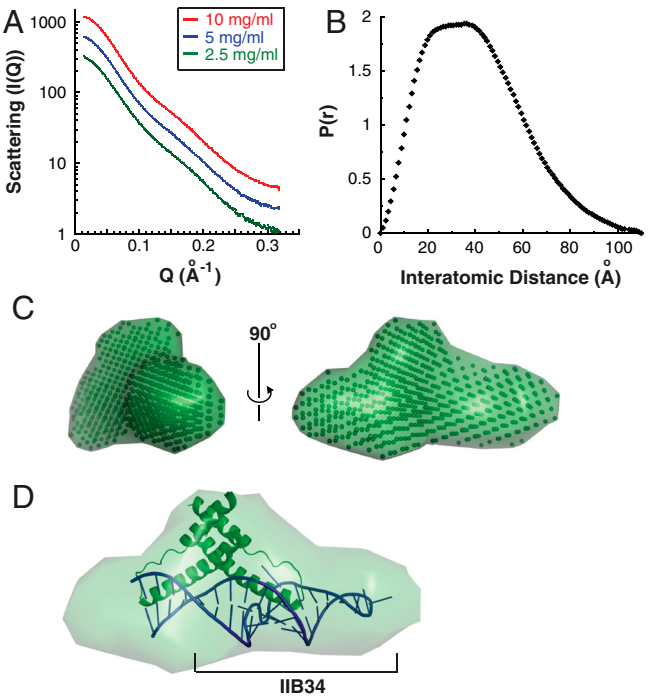

Fig. 4. SAXS-derived model of a Rev dimer bound to RNA. (*A*) SAXS data of the Rev dimer with IIB42 RNA at the three concentrations indicated. (*B*) The interatomic distance probability function calculated from the average scattering curve at all three concentrations. (*C*) DAMMAVER average dummy atom model of 20 independent DAMMIN calculated models. Dummy atoms are shown as spheres, surrounded by a molecular surface. (*D*) Fit of the crystal structure of a Rev dimer (green model) (PDB code: 3LPH) and the NMR structure of the IIB34 RNA (blue model) (7) into the SAXS-derived molecular envelope (green surface). The NMR structure contains only 34 of the 42 nucleotides used for SAXS (bracketed), so an additional four base pairs of ideal A-form RNA was appended to the model, thereby accounting for the additional observed SAXS density. The contribution of the disordered Rev C termini is not accounted for but would be expected to be small.

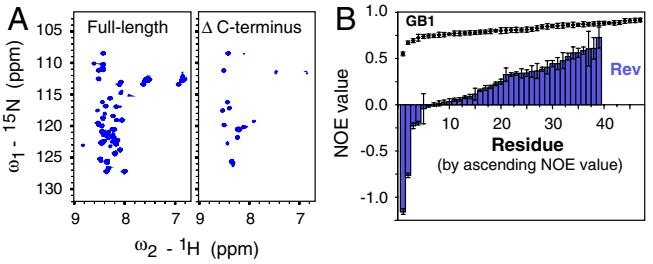

Fig. 5. NMR of Rev-RNA complexes reveals a disordered C terminus. (*A*) ^{1}H-^{15}N HSQC of full-length and C-terminally truncated Rev L60R-RNA complexes. (*B*) ^{1}H-^{15}N heteronuclear NOE values sorted by ascending value, derived from the spectrum of the full-length protein-RNA complex in *A*. Values for GB1 were derived from additional spectra (Fig. S4*B*) and illustrate the large difference in NOE values between a structured protein (GB1) and the disordered regions of Rev (23). Error bars shown were based on background noise values for both spectra.

belong to the C terminus. The ^{15}N-^{1}H NOE values of these resonances were largely below 0.5 (mean value: 0.20, median value: 0.25) (Fig. 5*B*), conclusively establishing that these residues are poorly structured (23).

Rev Forms a Discrete Hexameric Complex with Full-Length RRE. Rev-mediated RNA export in vivo and high-affinity assembly of the functional RNP in vitro require more than 230 nucleotides of the RRE (Fig. 1*B*) (6, 9). We asked if the RRE is able to drive formation of discrete Rev-RRE complexes under conditions similar to those for the monomeric and dimeric complexes. SEC of the RRE alone showed an 86 kDa species, and upon adding eightfold excess of purified, cleaved Rev, the peak mass shifted to 160 kDa, corresponding to the RRE with six Rev monomers bound (Fig. 6*A*, Fig. S5, and Table S1). The stoichiometry is consistent with previous estimates of 6–8 monomers bound to this length of RRE (6). Similarly, adding eightfold excess of GB1-Rev produced a narrow SEC peak with a mass of 216 kDa, also corresponding to six Rev monomers bound (Fig. 6*A*, Fig. S5, and Table S1). Native gel analyses further indicated that these

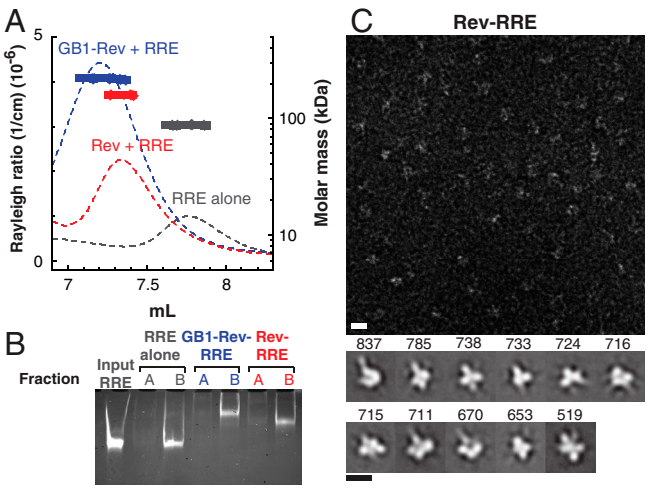

Fig. 6. Assembly and electron microscopy of discrete Rev-RRE particles. (*A*) Measured MALS (dashed lines, left axis) and calculated molar masses (solid lines, right axis) determined from SEC peaks (Fig. S5) of the 5 μM 242-nucleotide RRE alone (gray), and 40 μM Rev with 5 μM RRE (red) and 40 μM GB1-Rev with 5 μM RRE (blue). (*B*) Native gel of unbound RRE and SEC peaks. The predominant peak shown in *A* is indicated as fraction B, while fraction A represents a small amount of dimeric RRE generated during annealing (Fig. S5). (*C*) A representative negative stain field of Rev-RRE particles (*Top*). (Scale bar, 10 nm.) Below are 11 classes of particle averages with the number of individual particles in each class indicated.

SEC species migrate as distinct bands (Fig. 6*B*), consistent with assembly of discrete Rev-RRE complexes.

If Rev indeed forms specific, discrete hexameric complexes on the RRE, we predicted that previously characterized oligomerization mutations (12) would alter the observed complexes. Upon adding eightfold excess Rev L18Q/L60R to the RRE, the predominant SEC peak eluted later and was broader than that observed with wild-type Rev and had a calculated mass consistent with 3–4 Rev monomers bound (Fig. S6*A*). Furthermore, distinct bands were no longer seen on native gels (Fig. S6*B*), consistent with formation of heterogeneous complexes and further confirming that Rev oligomerization is needed to form discrete Rev-RRE complexes.

Electron Microscopy of Rev-RRE Complexes. We next wished to visualize the overall architecture of the RNP complexes by EM, particularly since previous studies generated filamentous particles ~14 nm wide and 30–1500 nm long (13, 16), much larger and more heterogeneous than expected from the binding (9), SEC, native gel, and MALS data. Indeed, Rev alone forms protein filaments (Fig. S7), but negative stain EM images of our soluble Rev-RRE complexes revealed distinct globular particles of ~10 nm diameter (Fig. 6*C*), a dimension consistent with the measured mass of 160 kDa (Fig. 6*A*). Lobed features were clearly seen within individual particles that were defined further by placing 7,801 selected particles into 11 distinct class averages (Fig. 6*C*). Small differences in size were observed between the class averages that may represent different views of the same oligomeric state or slight heterogeneities between particles. Interestingly, the Rev-RRE particles lack obvious internal symmetry, consistent with a role for the asymmetric RRE in defining the architecture of the Rev-RRE complex. EM of GB1-Rev-RRE complexes also revealed distinct particles, slightly larger than the Rev-RRE complexes (Fig. S8) as expected with its six appended GB1 domains. Similar asymmetry and lobed features were prominent in the fusion protein complexes (Fig. S8), further indicating that the Rev-RRE complexes are structurally well defined.

Discussion

The structure of the HIV Rev-RRE complex has eluded efforts for nearly two decades, in part due to low Rev solubility and uncontrolled oligomerization. Such problems are not uncommon for large molecular assemblies, particularly RNPs, but recent successes with the spliceosome and viral nucleoproteins illustrate that these problems are surmountable (24–26). By maintaining natively purified Rev in the presence of RNA or an RNA surrogate, such as a negatively charged fusion domain or oxyanionic salts, we obtain Rev preparations 100–1,000-fold more soluble than previously reported, suggesting that Rev aggregation is driven largely by charged interactions and not interactions between its hydrophobic oligomerization domains. Interestingly, although fused solubility-enhancing domains are generally valued for being inert (17), GB1 solubilizes Rev by serving as a binding partner when its natural RNA partner is absent (Fig. S4*A*), suggesting that GB1 may be useful for stabilizing other proteins prone to charge-mediated aggregation. This property of GB1 may mimic the role of trigger factor as a chaperone that solubilizes ribosomal proteins during ribosome subunit biogenesis (27).

Even under conditions where Rev does not aggregate to the point of precipitation, the protein forms large, heterogeneous oligomers in solution (Fig. 2*A*). Experiments presented here show that RNA plays an active role in controlling Rev oligomerization by forming defined RNPs rather than allowing formation of large protein oligomers, consistent with the essential role of the RRE in forming export-competent RNPs (6, 9). This influence of RNA is emerging as a common theme in viral RNA-binding homooligomers; other recent nucleoprotein structural studies

demonstrate that the length of bound RNA also helps specify the oligomeric state of those complexes (24, 26).

To gain structural insight into how RNA drives formation of the Rev-RRE complex, we set out to isolate stable Rev-RNA species that would likely represent the first intermediates in the kinetic assembly pathway (10). Consistent with previous studies demonstrating the consequences of single residue changes in the oligomerization domains (12), we now are able to purify discrete dimeric or monomeric assemblies at nearly millimolar concentrations (>10 mg/ml) using these mutants together with an appropriate dimeric or monomeric RRE site (Fig. 3). The complexes generated are suitable for structural methods: SAXS data suggest a plausible orientation in which two Rev monomers bind cooperatively to a single RNA (Fig. 4), and NMR experiments (Fig. 5) indicate that the C-terminal region of Rev is disordered. These data are consistent with a structural model for stabilization of the Rev dimer in which the oligomerization and RNA-binding domains cooperatively bind to the stem IIB hairpin to form a distinct lobed structure (Fig. 4*D*), with the appended disordered C terminus available to bind Crm1.

In addition to these well-defined subcomplexes, we are able to purify discrete complexes of functional, wild-type Rev bound to the full-length RRE (Fig. 6). Unlike the heterogeneous state of wild-type Rev bound to the single IIB site (Fig. 2), when the entire RRE is bound, a defined hexameric species is observed, consistent with the cooperativity and high affinity of Rev-RRE assembly. The binding properties of this complex strongly correlate with RNA export function (9), and its assembly is disrupted by oligomerization mutations (Fig. S6), demonstrating its functional relevance. EM images provide initial structural views of these Rev-RRE complexes (Fig. 6), confirming that they are discrete in nature and distinct from the heterogeneous filamentous structures or disordered aggregates previously observed (13, 16, 18, 19). The formation of Rev filaments can be blocked either by binding to RNA (Fig. 6) or to an antibody directed to one of the hydrophobic oligomerization surfaces (28), suggesting filaments are stabilized by ionic interactions between oppositely charged regions of the protein, by hydrophobic interactions, or both. It will be interesting to determine at a structural level whether the RRE prevents filament growth by capping the ends of the Rev oligomer or altering the orientation of Rev subunits in the RNP.

Both the dimeric Rev-RNA and full Rev-RRE complexes display some obvious structural features, like those of other ordered asymmetric RNPs such as the ribosome and spliceosome (29, 30). Especially interesting is the observation that Rev-RNA particles lack the symmetry seen in the structures of Rev dimers without RNA (PDB codes: 3LPH and 2X7L) (11) and in structures of other homooligomeric RNPs such as respiratory syncytial virus nucleoprotein-RNA (26), archaeal Sm-RNA (31), and Hfq-RNA (32) complexes that all bind nonstructured RNAs. We infer that the crucial difference is the RRE structure, which provides a framework to organize the Rev-RRE RNP into a discrete, asymmetric particle, analogous to the role that rRNA plays in cooperative assembly of heterooligomeric protein subunits of the ribosome (33). The overall architecture of the Rev-RRE RNP, as specified by the RRE, likely has critical functional ramifications, including exposure of the NES regions for Crm1 recognition and organizing the RNP to bind other proteins that assist in passage through the nuclear pore or disassemble particles in the cytoplasm. Thus, our characterization of discrete Rev-RRE particles not only lays the groundwork for further structural studies on this essential viral complex but also extends our understanding of how RNA can drive assembly of ordered RNPs.

Materials and Methods

Purification of Native Rev. Rev protein was expressed with an N-terminal His-GB1 tag as previously described (9) in *E. coli* strain BL21/DE3 in LB medium or, for NMR experiments, in M9 minimal media supplemented with trace minerals, thiamine, and $^{15}NH_4Cl$ as the sole nitrogen source. Purification was performed as described (9) with the following modifications to remove endogenous *E. coli* RNA. The cell lysate was supplemented with RNase A (50 μg/ml) and T1 (50 U/ml) (Roche) and NaCl to 2 M prior to centrifugation. Cleared lysate was bound to Ni-NTA resin (Qiagen) and washed thoroughly in the presence of 2 M NaCl and rinsed and eluted using buffers containing 250 mM NaCl. Fractions were analyzed by SDS/PAGE, pooled, and dialyzed against buffer B [40 mM Tris pH 8.0, 200 mM NaCl, 2 mM β-ME] at 4 °C. Specific RNA or 100 mM Na_2SO and 400 mM $(NH_4)_2SO_4$ were added to GB1-Rev prior to TEV proteolysis to prevent Rev aggregation. Rev or Rev-RNA complexes were collected in the flow-through of Ni-NTA resin, stored at 4 °C and used soon after purification to minimize aggregation. Complete details can be found in *SI Text*.

SEC, MALS, and Native Gel Analyses. Analytical SEC was performed using an Ettan LC system (GE Life Sciences) with a silica gel KW803 column (Shodex) equilibrated in buffer B at a flow rate of 0.35 ml/min. The system was coupled on-line to an 18-angle MALS detector (DAWN HELEOS II, Wyatt Technology) and a differential refractometer (Optilab rEX, Wyatt Technology). Molar masses were determined using ASTRA 5.3.1.5 software.

SEC fractions were loaded onto 6% (for full-length RRE) or 10% (for IIB hairpins) polyacrylamide (37.5:1 mono:bis, 0.5x TBE) gels, run for 1–2 hr, and stained with ethidium bromide or toluidine blue.

SAXS and NMR. Preparative SEC was performed on Rev L60R in 2:1 stoichiometry with IIB42 on an AKTApurifier system (GE Life Sciences) with a Superdex 200 10/300 GL column equilibrated in buffer B at a flow rate of 0.5 ml/min. Only the peak corresponding to 2:1 Rev-RNA complexes was collected and concentrated for SAXS and NMR experiments. SAXS data were acquired, processed and used to generate ab initio models based on previously described methods (34) (see *SI Text* for details). The Rev protein dimer structure (PDB code: 3LPH) was aligned with the Rev ARM bound to IIB34 RNA determined by NMR (PDB code 1ETF) (7) (backbone RMSD = 1.0). Four base pairs of ideal A-form RNA were appended to total 42 nt, and the complete model was fit into the SAXS density and visualized using PyMOL (http://www.pymol.org). Detailed methods for NMR data acquisition and processing are described in *SI Text*.

Electron Microscopy. Rev protein alone or SEC-purified Rev-RRE or GB1-Rev-RRE complexes were negatively stained as previously described (35). Images were acquired on a Tecnai T12 electron microscope (FEI Company) equipped with a LaB_6 filament and recorded on a Gatan 4096 × 4096 UltraScan (Gatan, Inc., Pleasanton, CA) CCD camera. Particles were interactively selected from micrographs using Ximdisp (36). Class averages were generated using multivariate statistical analysis as described (37). Complete details can be found in *SI Text*.

ACKNOWLEDGMENTS. We thank John Gross, Mark Kelly, and Greg Lee for assistance with NMR, Dan Southworth for assistance with MALS, and Kristin Krukenburg, Stephen Floor, and ALS staff for assistance with SAXS. We thank John Gross, Geeta Narlikar, J. J. Miranda, and Miles Pufall for helpful suggestions and critical reading of the manuscript. M.D.D. was supported by a Howard Hughes Medical Institute predoctoral fellowship. D.S.B is a NIGMS-IMSD fellow. This work was supported by the HARC Center National Institutes of Health Grant P50GM82250 (to A.D.F and Y.C.).

1. Cullen BR (2003) Nuclear mRNA export: Insights from virology. *Trends Biochem Sci* 28:419–424.
2. Pollard VW, Malim MH (1998) The HIV-1 Rev Protein. *Annu Rev Microbiol* 52:491–532.
3. Malim MH, Cullen BR (1991) HIV-1 structural gene expression requires the binding of multiple Rev monomers to the viral RRE: Implications for HIV-1 latency. *Cell* 65:241–248.
4. Malim MH, et al. (1990) HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. *Cell* 60:675–683.
5. Huang X, et al. (1991) Minimal Rev-response element for Type 1 human immunodeficiency virus. *J Virol* 65:2131–2134.
6. Mann DA, et al. (1994) A molecular rheostat: Co-operative rev binding to stem I of the Rev-response element modulates human immunodeficiency virus type-1 late gene expression. *J Mol Biol* 241:193–207.
7. Battiste JL, et al. (1996) Alpha helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science* 273:1547–1551.
8. Tan R, Chen L, Buettner JA, Hudson D, Frankel AD (1993) RNA recognition by an isolated alpha helix. *Cell* 73:1031–1040.
9. Daugherty MD, D'Orso I, Frankel AD (2008) A solution to limited genomic capacity: Using adaptable binding surfaces to assemble the functional HIV Rev oligomer on RNA. *Mol Cell* 31:824–834.

10. Pond SJ, Ridgeway WK, Robertson R, Wang J, Millar DP (2009) HIV-1 Rev protein assembles on viral RNA one molecule at a time. *Proc Natl Acad Sci USA* 106:1404–1408.
11. Dimattia MA, et al. (2010) Implications of the HIV-1 Rev dimer structure at 3.2 A resolution for multimeric binding to the Rev response element. *Proc Natl Acad Sci USA* 107:5810–5814.
12. Jain C, Belasco JG (2001) Structural model for the cooperative assembly of HIV-1 Rev multimers on the RRE as deduced from analysis of assembly-defective mutants. *Mol Cell* 7:603–614.
13. Heaphy S, Finch JT, Gait MJ, Karn J, Singh M (1991) Human immunodeficiency virus type 1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich "bubble" located within the rev-responsive region of viral mRNAs. *Proc Natl Acad Sci USA* 88:7366–7370.
14. Wingfield PT, et al. (1991) HIV-1 Rev expressed in recombinant Escherichia coli: Purification, polymerization, and conformational properties. *Biochemistry* 30:7527–7534.
15. Cole JL, Gehman JD, Shafer JA, Kuo LC (1993) Solution oligomerization of the rev protein of HIV-1: Implications for function. *Biochemistry* 32:11769–11775.
16. Havlin RH, Blanco FJ, Tycko R (2007) Constraints on protein structure in HIV-1 Rev and Rev-RNA supramolecular assemblies from two-dimensional solid state nuclear magnetic resonance. *Biochemistry* 46:3586–3593.
17. Zhou P, Wagner G (2009) Overcoming the solubility limit with solubility-enhancement tags: Successful applications in biomolecular NMR studies. *J Biomol NMR* 46:23–31.
18. Watts NR, et al. (1998) Three-dimensional structure of HIV-1 Rev protein filaments. *J Struct Biol* 121:41–52.
19. Blanco FJ, Hess S, Pannell LK, Rizzo NW, Tycko R (2001) Solid-state NMR data support a helix-loop-helix structural model for the N-terminal half of HIV-1 Rev in fibrillar form. *J Mol Biol* 313:845–859.
20. Koch MH, 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.
21. Auer M, et al. (1994) Helix-loop-helix motif in HIV-1 Rev. *Biochemistry* 33:2988–2996.
22. Wang Y, Jardetzky O (2002) Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Sci* 11:852–861.
23. Kay LE, Torchia DA, Bax A (1989) Backbone dynamics of proteins as studied by 15 N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry* 28:8972–8979.
24. Albertini AA, et al. (2006) Crystal structure of the rabies virus nucleoprotein-RNA complex. *Science* 313:360–363.
25. Pomeranz Krummel DA, Oubridge C, Leung AK, Li J, Nagai K (2009) Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. *Nature* 458:475–480.
26. Tawar RG, et al. (2009) Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. *Science* 326:1279–1283.
27. Martinez-Hackert E, Hendrickson WA (2009) Promiscuous substrate recognition in folding and assembly activities of the trigger factor chaperone. *Cell* 138:923–934.
28. Stahl SJ, et al. (2010) Generation and characterization of a chimeric rabbit/human Fab for co-crystallization of HIV-1 Rev. *J Mol Biol* 397:697–708.
29. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* 289:905–920.
30. Ritchie DB, Schellenberg MJ, MacMillan AM (2009) Spliceosome structure: Piece by piece. *Biochim Biophys Acta* 1789:624–633.
31. Toro I, et al. (2001) RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex. *EMBO J* 20:2293–2303.
32. Schumacher MA, Pearson RF, Moller T, Valentin-Hansen P, Brennan RG (2002) Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: A bacterial Sm-like protein. *EMBO J* 21:3546–3556.
33. Williamson JR (2008) Cooperativity in macromolecular assembly. *Nat Chem Biol* 4:458–465.
34. Krukenberg KA, Bottcher UM, Southworth DR, Agard DA (2009) Grp94, the endoplasmic reticulum Hsp90, has a similar solution conformation to cytosolic Hsp90 in the absence of nucleotide. *Protein Sci* 18:1815–1827.
35. Ohi M, Li Y, Cheng Y, Walz T (2004) Negative staining and image classification—Powerful tools in modern electron microscopy. *Biol Proced Online* 6:23–34.
36. Smith JM (1999) Ximdisp—A visualization tool to aid structure determination from electron microscope images. *J Struct Biol* 125:223–228.
37. Shaikh TR, et al. (2008) SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nat Protoc* 3:1941–1974.