The Structure of RalF, an ADP-ribosylation Factor Guanine Nucleotide Exchange Factor from Legionella pneumophila, Reveals the Presence of a Cap over the Active Site*

Received for publication, September 20, 2004, and in revised form, October 21, 2004
Published, JBC Papers in Press, November 1, 2004, DOI 10.1074/jbc.M410820200

J. Carlos Amor‡, Jennifer Swails‡, Xinquan Zhu‡, Craig R. Roy§, Hiroki Nagai§, Alyssa Ing mundson§, Xiaodong Cheng‡, and Richard A. Kahn‡¶

From the ‡Department of Biochemistry, Emory University School of Medicine, Atlanta, Georgia 30322 and §Section of Microbial Pathogenesis, Yale University School of Medicine, Boyer Center for Molecular Medicine, New Haven, Connecticut 06536

The Legionella pneumophila protein RalF is secreted into host cytosol via the Dot/Icm type IV transporter where it acts to recruit ADP-ribosylation factor (Arf) to pathogen-containing phagosomes in the establishment of a replicative organelle. The presence in RalF of the Sec7 domain, present in all Arf guanine nucleotide exchange factors, has suggested that recruitment of Arf is an early step in pathogenesis. We have determined the crystal structure of RalF and of the isolated Sec7 domain and found that RalF is made up of two domains. The Sec7 domain is homologous to mammalian Sec7 domains. The C-terminal domain forms a cap over the active site in the Sec7 domain and contains a conserved folding motif, previously observed in adaptor subunits of vesicle coat complexes. The importance of the capping domain and of the glutamate in the "glutamic finger," conserved in all Sec7 domains, to RalF functions was examined using three different assays. These data highlight the functional importance of domains other than Sec7 in Arf guanine nucleotide exchange factors to biological activities and suggest novel mechanisms of regulation of those activities.

The subversion of mammalian proteins and functions by invading pathogens is a recurring theme in cell biology (e.g. Refs. 1 and 2). After phagocytosis by macrophages, the human pathogen Legionella pneumophila (the causative agent for Legionnaire's disease) is capable of altering the normal progression of membrane traffic in the secretory and endocytic pathways to create a stable vacuole in which the organism replicates (3–5). An essential component in this process is the formation of a type IV protein secretion apparatus and is required for the recruitment of Arf to the vacuole containing phagosomes (6–10). The translocated bacterial proteins then prevent the fusion of the phagosome with lysosomes as well as enlarge the stabilized organelle through the recruitment of vesicles and membranes from the endoplasmic reticulum through unknown mechanisms. A search for proteins that function in the secretory pathway led to the identification of a Legionella protein, termed RalF, that carries an N-terminal Sec7 domain, a ~200-residue domain found in all guanine nucleotide exchange factors (GEF) that act on ADP-ribosylation factors (Arfs) (4, 8). RalF shares 42% identity with eukaryotic Arf GEFs in the N-terminal Sec7 domain. There is essentially no sequence homology between the C-terminal half of RalF and any mammalian Arf GEF, although there is a predicted Arf GEF found by translation of the genomic sequence of Rickettsia prowazekii (40% identity with RalF in the C-terminal non-Sec7 sequence) (8).

Arf GEFs speed the exchange of GTP for GDP on Arfs, thereby promoting the activation of the GTPase, which also produces a more stable binding of the GTPase to membranes (11). Arfs are required for the budding of nascent secretory vesicles and assist in the recruitment of coat proteins or complexes as well as serve as direct activators of lipid-modifying enzymes and impact traffic at a number of steps in the secretory pathway. RalF is transported through the type IV secretion apparatus and is required for the recruitment of Arf to Legionella-containing phagosomes (8). We report here the structure of RalF at 1.4 Å of resolution, the first structure solved for a type IV channel cargo or for a full-length Arf GEF. The structure reveals that exposure of the active site requires a re-organization of protein domains not previously anticipated.

MATERIALS AND METHODS

Protein Production and Crystallization—The open reading frame of RalF or the Sec7 domain of RalF was subcloned into the pHis-1 vector (12) (a gift of Z. Derewenda, University of Virginia) at the NcoI and XhoI sites. The recombinant proteins had an additional N-terminal 28 amino acids, including a hexahistidine tag at position 24 to 19, in which the original initiating methionine was replaced by a serine. Thus, recombinant RalF contained 402 amino acids with a calculated molecular mass of 45,484.2 Da and an extinction coefficient at 280 nm of 30,580. Recombinant RalF-Sec7 contained 229 amino acids with a calculated molecular mass of 26,136.3 Da and an extinction coefficient at 280 nm of 11,520. The extinction coefficients were used to quantify purified RalF and RalF-Sec7 preparations. Recombinant RalF and RalF-Sec7 were expressed in BL21(DE3) cells

1 The abbreviations used are: GEF, guanine nucleotide exchange factor; Arf, ADP-ribosylation factor; Sfem, selenomethionine; MOPS, 3-(N-morpholino)propanesulfonic acid; GTPγS, guanosine 5′-3-D-thiotriphosphate; GFP, green fluorescent protein; PDB, Protein Data Bank; CHO, Chinese hamster ovary; r.m.s.d., root mean square deviation; GST, glutathione S-transferase.

* This work was supported in part by National Institutes of Health Grants GM61268 (to R. A. K. and J. C. A.), GM49245 and GM61355 (to X. C.), and AI41699 (to C. R. R.) and a National Science Foundation Predoctoral fellowship (to A. I.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must include this fact in any advertisement for the journal.

¶ To whom correspondence should be addressed: Dept. of Biochemistry, Emory University School of Medicine, 1510 Clifton Rd., Atlanta, GA 30322. Tel.: 404-727-3561; Fax: 404-727-3746; E-mail: rkahn@emory.edu.

The open reading frame of RalF or the Sec7 domain of RalF was subcloned into the pHis-1 vector (12) (a gift of Z. Derewenda, University of Virginia) at the NcoI and XhoI sites. The recombinant proteins had an additional N-terminal 28 amino acids, including a hexahistidine tag at position 24 to 19, in which the original initiating methionine was replaced by a serine. Thus, recombinant RalF contained 402 amino acids with a calculated molecular mass of 45,484.2 Da and an extinction coefficient at 280 nm of 30,580. Recombinant RalF-Sec7 contained 229 amino acids with a calculated molecular mass of 26,136.3 Da and an extinction coefficient at 280 nm of 11,520. The extinction coefficients were used to quantify purified RalF and RalF-Sec7 preparations. Recombinant RalF and RalF-Sec7 were expressed in BL21(DE3) cells

1 The abbreviations used are: GEF, guanine nucleotide exchange factor; Arf, ADP-ribosylation factor; Sfem, selenomethionine; MOPS, 3-(N-morpholino)propanesulfonic acid; GTPγS, guanosine 5′-3-D-thiotriphosphate; GFP, green fluorescent protein; PDB, Protein Data Bank; CHO, Chinese hamster ovary; r.m.s.d., root mean square deviation; GST, glutathione S-transferase.
and purified on Ni-IMAC (Amersham Biosciences) and Q-methylacrylate (Bio-Rad) columns. RalF was exchanged into 20 mM Tris, 20 mM NaCl, pH 7.4, and concentrated to at least 20 mg/ml. RalF crystals were obtained by vapor diffusion at 21 °C with a 1:1 mixture of RalF and 0.1 M sodium acetate-3H2O, pH 4.6, 2.0 M sodium formate over a 0.5-ml reservoir of the crystallization solution. Large crystals of the space group P31 appeared 10–14 days after initial set-up and grew for about an additional 10 days. RalF-Sec7 crystals were obtained by vapor diffusion using 35% polyethylene glycol 2000, 0.2 M sodium acetate, and 100 mM Tris, pH 8.5. Crystals of space group P212121 formed after 2 days and continued to grow for at least a week. Although crystals formed, they were not of diffraction quality. By using initially generated non-diffraction quality crystals as seeds, diffraction quality crystals were obtained. Additionally, SeMet-containing crystals were obtained by exposure of the RalF-Sec7 in BL21(DE3) cells in MOPS media, supplemented 30 min before induction of protein expression with lysine, phenylalanine, and isoleucine, as well as isoleucine, valine, and SeMet, all at 50 mg/liter. Recombinant human Arf1 was expressed and purified as previously described (13). Yields of Arf1, RalF, and RalF-Sec7 were ~40, 70, and 200 mg protein/liter bacterial culture, respectively.

Data Collection and Structure Determination—For x-ray data collection, RalF crystals were transferred to a solution containing the mother liquor saturated with SeMet and fixed after 10 min. Soaked crystals were exchanged into a mother liquor and cryo-cooled in liquid nitrogen. X-ray diffraction data were measured at a wavelength of 0.9795 Å, collected at −165 °C on an ADSC detector at the X26C beamline at the NLSL of the Brookhaven National Laboratories. RalF crystals diffracted isocratically to at least 1.4 Å of resolution. One heavy atom derivative of potassium tetrannitroplatin(II) (K2Pt(NO3)6) was obtained by soaking a fully formed crystal for 6 h in a solution of 0.1 M sodium citrate and 0.1 M sodium phosphate, pH 6.0, 1 M ammonium sulfate, and 10% glycerol. Soaked crystals were mounted on a nylon loop and cryo-cooled in liquid nitrogen. X-ray diffraction data were measured at a wavelength of 0.9600 Å. The SeMet-RalF-Sec7 crystal diffracted isocratically at a wavelength of 0.9795 Å, collected at −165 °C on an ADSC detector at the X26C beamline at the NLSL of the Brookhaven National Laboratories. RalF crystals diffracted to 2.2 Å using a synchrotron radiation source. Electron density maps of the full-length RalF were calculated in the trigonal space group P31, using isomorphous replacement of a platinum derivative with anomalous scattering (Table I). The final model of RalF, refined to 1.4 Å of resolution, consists of two monomers per asymmetric unit, with an R-factor of 0.205 and an R-free value of 0.222. The last 20 residues (355–374) of each protomer were unstructured. The average thermal factor for the refined structure was 23 Å2, and no residue was in the disallowed region of the Ramachandran plot (20).

Structure and Activities of RalF, an Arf GEF—Overview of the RalF Structure—The structure of RalF was solved by molecular replacement using the newly determined structure of the do-
cylindrical, as previously reported for isolated eukaryotic Sec7 domains (18). The isolated Sec7 domain crystallized as a monomer in space group P31, and residues (355–374) of each protomer were unstructured. The average thermal factor for the refined structure was 23 Å2, and no residue was in the disallowed region of the Ramachandran plot (20).

Nucleotide Exchange Assay—Arf GEF activity was determined by measuring the binding of 5′-[γ-32P]GTP to purified recombinant Arf1, as described previously (8). The concentration of Arf1 (1 μM) was 10-fold that of the Arf GEF in each assay. Reactions were performed in a buffer containing 50 mM HEPES, pH 7.5, 100 mM KC1, 1 mM MgCl2, 1 mM dithiothreitol, with 1.5 mg/ml azolectin vesicles. All experiments were repeated at least three times, with similar results.

Secured Alkaline Phosphatase (SEAP) Secretion Assay—RalF and RalF-E103K were ligated into the expression vector pGreenLantern (Invitrogen) such that GFP was appended to the N terminus of RalF. Plasmids encoding Mg-Tagged RALF-T31N, ARNO, and ARNO-E156K were described previously. For the SEAP assay, CHO FcRII cells were maintained in a minimal medium (21% Eagle’s medium with 10% fetal bovine serum and plated in 24-well dishes at a density of 3 × 104 cells/well. After an overnight incubation cells were cotransfected with 300 ng of a plasmid encoding a SEAP protein and 200 ng of the plasmid encoding the protein indicated. 18 h after transfection, the cells were washed, and medium was replaced. Aliquots of the supernatant were sampled 1, 4, and 7 h later, and SEAP activity was measured using the Phospha-light SEAP kit (Applied Biosystems). Intraocular-associated SEAP activity was measured at 7 h. Data are presented as the ratio of SEAP activity measured in the supernatants to the cell-associated SEAP activity.

Arf1-GFP Recruitment to Vesicles—An assay that measures the recruitment of Arf1-GFP to Legionella-containing vacuoles was used to determine the effect of Glu-103 substitutions on RalF function in vivo. The assay was conducted as described previously using bone marrow-derived macrophages from A/J mice that were transduced with a retroviral vector encoding Arf1-GFP (4, 8). Briefly, the indicated wild type and mutant RalF proteins were expressed from the plasmid pPMMB207 in a Legionella strain derived from Lp01 that has the ralF gene deleted from the chromosome (8). Macrophages were infected with L. pneumophila strain P1 ACRA501 (8) at an MOI of 50. RalF-GFP staining of Legionella-containing vacuoles was scored by fluorescence microscopy for at least 50 vacuoles, and the data are presented as the mean ± S.D. from three independent experiments. The Cya fusion assay was used to test the efficiency of Dot/Icm-mediated translocation of the RalF Glu-103 variants into host cells. In this assay, the calmodulin-dependent adenylate cyclase enzyme from the Bordetella pertussis Cya protein is fused to the N-terminal 30 residues of RalF proteins in the vector pRCS.CYA. Delivery of a Cya-RalF fusion protein into host cells results in calmodulin binding to the adenylate cyclase enzyme, resulting in the production of cAMP. Legionella strains derived from Lp01 that contained the indicated plasmids were used to infect CHO FcRII cells for 1 h, and cAMP levels were then measured using a commercially available immuno assay (Amersham Biosciences). cAMP values are the average ± S.D. from three independent assay wells.

RESULTS
domains (24–26). This shape is generated by a right-handed superhelix, formed by helices A-G and ending with the three C-terminal helices, H-I-J. The topology break between the helical bundles harbors the Arf binding cleft. The catalytically active glutamate, Glu-103 (homologous to Glu-156 in ARNO) (27), immediately N-terminal to helix G, is in the center of the interface between the two domains.

The C-terminal domain contains helices K-P and strands 1–6, forming a discrete and compact structure. The anti-parallel alignment of β1-β6 forms a concave surface over the Sec7 domain, whereas the convex side of the β-sheet is crowned by four helices K, L, O, and P. We term this the Sec7-capping domain (SCD). The presence of folding motifs or similarities in three-dimensional structures between the SCD and other proteins was sought using DEJAVU (28) and DALI (29) searches. Among the top scores found were the COP-I (Ref. 30; PDB code 1pzd), the AP-2 clathrin adaptor proteins was sought using DEJAVU (28) and DALI (29) searches. The third contact between the two domains involves the C-terminal region of the Sec7 domain interacts with a C-terminal region of the SCD in an acidic groove (Fig 3B, left). This interaction involves three glutamates, three lysines, and a tyrosine (Tyr-326) from the SCD that extends into the Sec7 domain. Specifically, loop 6/P in the Sec7 domain forms a hairpin turn, which places Tyr-326 in between four regions in the Sec7 domain, allowing charges Lys-19 in Sec7 to interact with Glu-329 (αO), Asp-327 (loop6/P), and Lys-323 (β6), respectively, in the SCD.

The central, catalytic region of the Sec7 domain interacts with the center of the β-sheet of the SCD (Fig 3B, center). The entire β-sheet of the SCD forms a domed surface over the N-terminal end of αG and the preceding loop, which includes the key catalytic residue Glu-103. Glu-103 is buried in the interface between the two domains and is critical for stabilizing a network of polar charge interactions involving both domains. This network of interactions involves Asp-60 of helix D interacting with Lys-106 of helix G, Lys-106 interacting with Glu-103, and Glu-103 interacting with Gln-297 of β4 and Gln-297 of β5.

### Table 1: Data collection and refinement statistics

| Crystal       | Native RaIF | K<PbsNO_3>2 RaIF | SeMet-RaIF-Sec7 |
|---------------|-------------|-----------------|-----------------|
| Space group   | P3 1        | P3 1            | P2 2, 2        |
| Cell parameters (Å) | a = b = 84.56 | a = b = 84.45 | a = 40.20 |
|                | c = 110.18 | c = 110.53 | c = 109.13 |
| Resolution (Å) | 40.0–1.41 | 30.0–2.15 | 50.0–2.16 |
| Measured reflections | 513,240 | 142,089 | 81,418 |
| Unique reflections | 171,080 | 47,363 | 1,821 |
| Completeness (%) | 99.5 | 99.7 | 98.7 |
| Rmerge (%)      | 27.4 | 9.6 | 91.0 |
| Rfree (%)       | 4.4 | 12.1 | 6.1 |
| Rmerge (%)      | 36.8 | 20.2 | 11.5 |

### Refinement statistics

- Rmerge = 100 × Σ |F<sub>i</sub> - F<sub>p</sub>| / Σ |F<sub>p</sub>|
- Rmerge, 100 × Σ |F<sub>i</sub> - F<sub>p</sub>| / Σ |F<sub>p</sub>|
- Rvalues were calculated for data with a 2σ cutoff.

The third contact between the two domains involves the C-terminal portion of the Sec7 domain and N-terminal portion of the SCD. The Arf binding cleft of the Sec7 domain, previously described in eukaryotic Sec7 domain structures, interacts with the SCD, principally by hydrophobic interactions (Fig 3B, right), mainly via αH of the Sec7 domain and αM of αO of the SCD. These two short helices are inserted into the loops between β2 and β3, and β4 and β5, respectively (Fig 2). The complementary interacting surfaces are similar to the interactions observed previously between the Sec7 domain of Gea2 and Arf (34).
Together these three regions of contact between domains achieve a stably folded protein that allows for a number of hydrophobic side chains to be present on the surface of the interaction that would otherwise be exposed to solvent. Even the simplest comparison between the structures of RalF and that of a Sec7 domain bound to an Arf indicates that the SCD masks the Arf binding site and that these two domains of RalF must be rearranged to allow substrate binding.

Comparison of the SCD and Substrate Binding to the Sec7 Domain—
The occupation of the Arf binding site by a domain present in the Arf GEF was unexpected. Because structures are available for other isolated Sec7 domains bound to an Arf (Gea2-apo\[316\]Arf1 (34)) we compared the interactions of the substrate, Arf1, and the SCD to the Sec7 domain.

As described in the previous section, the interaction of the SCD with the Sec7 domain is across three areas of the domain interaction face (Fig. 3). In contrast, the interactions between Gea2 and apo[i\[16]Arf1 are confined to the C-terminal portion of the Sec7 domain-Arf interface and the catalytic site (Fig. 3A, right). Thus, the interaction surface of Arf1 is shifted to the right of that seen for the SCD, as shown in Fig. 3. The Arf makes no contacts in Gea2 with the equivalent to the acidic groove in RalF. However, the Arf makes additional contacts with i\[9251\]I in Gea2. The overlap between the space occupied by the SCD and Arf1 is extensive, resulting in a similar interface area (\(2700 \AA^2\)), affecting residues in aD (Glu-56), loop F/G (99KLPGE103), G(104AQKID108 and Gln-112), and H (133–147) as well as residues 148–152 in loop I/H. The presence of the SCD in that location would occlude the binding of Arf to the active site.

The comparisons above clearly suggest that the two domains of RalF must separate before the substrate, Arf, can bind the active site. Upon separation of the Sec7 domain from its binding partner (Arf or the SCD), residues previously buried become solvent-accessible. Of note is the large preponderance of non-polar-solvated surface by the separation of the SCD from the Sec7 domain. This leads us to speculate that a hydrophobic surface, presumably a biological membrane, will play a critical role in the domain separation and activation of RalF, as has previously been shown for other Arf GEFs, some of which
contain specific lipid binding domains, e.g. a PH domain. The proposed need for both accessory proteins and membrane lipids for activation Arfs by RalF is completely consistent with roles for such agents in Arf biology in general and the activation process specifically (35–40).

**RalF Has Arf GEF Activity**—The presence of the SCD over the active site suggested the possibility that RalF may be deficient in Arf GEF activity and prompted tests of its impact on nucleotide exchange activity. Arf GEF activity of full-length RalF was compared with that of the isolated Sec7 domain and to active site (Glu-103) point mutants. Proteins were expressed in bacteria as native proteins or as N-terminal GST fusion constructs. Because the release of GDP is the rate-limiting step in nucleotide exchange, either the release of GDP or the resulting increase in the binding of GTP can be used to monitor Arf GEF activities. We employed an assay that measures the initial rate of binding of[^35S]GTPγS, a stable analog of GTP, in the presence of a constant concentration of the Arf GEF.

Full-length RalF not only had activity, it was ~10 times more active than the isolated Sec7 domain (Fig. 4; compare RalF to Sec7RalF). The presence of a GST tag at the N terminus of RalF caused a 50% decrease in specific activity, but we still observed that GST-RalF was ~4-fold more active than GST-RalF-Sec7 (see Fig. 4). Thus, the presence of the SCD appears to substantially increase the specific activity of RalF as an Arf GEF.

Perhaps even more surprising was the observation that mutation of the residue predicted to play a critical role in the catalytic mechanism, Glu-103, to either alanine or lysine had no deleterious impact on Arf GEF activity in the context of the full-length protein fused to GST (Fig. 4). In every eukaryotic Arf GEF tested to date, mutation of this conserved glutamate, known as the "glutamate finger" for its proposed role in protruding into the nucleotide binding region of the bound Arf and promoting the release of GDP, resulted in substantial (>10-fold) or complete loss of Arf GEF activity (24, 26, 27).

Based on these observations and in contrast to data from eukaryotic Arf GEFs, it appears that Glu-103 plays a less critical role in nucleotide exchange than does the conserved glutamate in mammalian Arf GEFs.

Arf GEF activity of RalF was found to be absolutely dependent on the addition of lipids in the assay (see Fig. 4, no lipid/RalF). We have not examined the specificity for lipids in the assay, as we have used an azolectin vesicles that contain a complex mixture of lipids. This leaves open the possibility that RalF may have the ability to bind specifically to one or more phospholipids or that the Arf GEF activity may require such an interaction. The lipid dependence for nucleotide exchange and presence of the SCD over the active site are also suggestive of a specific interaction between the SCD and one or more lipids, functionally analogous perhaps to the PH domain found in several Arf GEFs, e.g. ARNO.

**The Effects of Glu-103 Mutants on RalF Activity in Cell-based Assays**—The retention of full Arf GEF activity in vitro after mutation of the conserved glutamate was surprising as homologous mutations in every other Arf GEF tested have been found to dramatically decrease activity. Because in vitro assays are always subject to the criticism that they fail to properly mimic the cellular environment, we sought a cell-based assay to assess the consequences of this mutation in RalF. The overexpression of some mammalian Arf GEFs (e.g. ARNO (41) and ARNO3 (42)) has been shown to inhibit the early secretory pathway, as visualized by the inhibition of secretion of the secreted form of alkaline phosphatase. We measured the amount of SEAP found in the media of cells overexpressing either ARNO, ARNO-E156K, RalF, or RalF-E103K (Fig. 5). We included in these assays the use of a negative dominant mutant of Arf1, Arf1-T31N, as a potent inhibitor of the early secretory pathway. As described previously (41), ARNO was found to inhibit secretion of SEAP, but the E156K mutant had lost all such activity. In contrast, RalF and RalF-E103K were each found to inhibit SEAP secretion even below the level seen for wild type ARNO. These results are consistent with results from the nucleotide exchange assay and suggest that RalF-E103K retains comparable activity to RalF in impacting the secretory pathway.

RalF normally gains access to the cytosol of eukaryotic cells via the type IV secretion system after internalization of the
FIG. 3. The two domains of RalF are held together through an extensive number of interactions that map to three regions. A, ribbon diagram of RalF (left), indicating the Sec7 domain (magenta), the SCD (green), and all those residues involved in domain interactions (yellow), as determined by solvation upon domain separation. The separated RalF-Sec7 and SCD (center) and their electrostatic representation as calculated by using GRASP (51). Acidic residues/regions are in red, and basic residues/regions are in blue. RalF-Sec7 exhibits an acidic N-terminal groove containing a basic edge, a catalytic region centered around Glu-103, and a hydrophobic C-terminal area known as the Arf-interacting cleft. The SCD complements the Sec7 domain profile together forming an interacting surface of 2700 Å². For comparison (right), apoΔ16Arf1 (red) is seen bound to Gea2-Sec7 (blue) with all interacting residues, as determined by solvation upon domain separation, indicated in yellow. ApoΔ16Arf1 essentially interacts only with the catalytic area and the hydrophobic C-terminal cleft. For apoΔ16Arf1/Gea2-Sec7, the interaction area is 2600 Å². Note the overlap in SCD and substrate (Arf) binding to the Sec7 domain. B, details of side-chain positions involved in domain-domain interactions are shown for the three areas of the interface. Selected secondary structures and side chains are labeled. For the β-sheet in the central image, the numbering of the individual strands is, from left to right, 6, 1, 5, 4, 3, and 2. Key residues from the SCD are listed above and, for the Sec7, underneath the ribbon image. Carbon atoms are in black, oxygens are in red, and nitrogens are in blue. Distances between atoms, indicated by a dashed line, are all between 2.8 and 3.0 Å.

FIG. 4. The guanine nucleotide exchange activity of RalF is increased by the presence of the SCD but is not compromised by mutation of Glu-103. The Arf GEF-dependent increase in guanine nucleotide exchange was assayed by the binding of [35S]GTP to purified, recombinant Arf1 (1 μM) in 50 mM HEPES, pH 7.5, 100 mM KCl, 1 mM MgCl₂, 1 mM dithiothreitol, and 1.5 mg/ml azolectin vesicles either alone or in the presence of purified RalF constructs (0.1 μM) as described under “Experimental Procedures.” The initial rates of nucleotide binding were linear over the course of the assay and are shown for each construct, as indicated. Note the higher activity of the full-length RalF compared with the isolated Sec7 domain and the lack of impact of Glu-103 mutations on exchange activity as well as the complete dependence on lipids for activity.
Legionella organism. In this way, much lower levels of the protein are delivered to a specific membrane in the cell as compared with the use of strong viral promoters driving protein expression from plasmids. Cells L. pneumophila in which the RalF gene is deleted and derivates carrying genes directing expression of wild type or Glu-103 mutants (E103K and E103A) of full-length RalF were used to test for the ability of each protein to recruit Arf1-GFP to Legionella-containing vacuoles in bone marrow-derived macrophages (4, 8). Wild type RalF was able to restore the recruitment of Arf1-GFP to vesicles, but mutation of Glu-103 to either lysine or alanine was sufficient to completely prevent recruitment of Arf1-GFP (Fig. 5A). As a control for expression and secretion into the macrophage cytosol, each of the RalF constructs was also expressed as C-terminal fusions with adenylyl cyclase. As seen in Fig. 5B, all three RalF proteins, carrying the secretion signal for binding to the type IV channel at their C termini, promoted the elevation of intracellular cAMP levels to comparable levels. Thus, when introduced into macrophages via the normal route, mutation of Glu-103 to either lysine or alanine was sufficient to ablate the ability of RalF to recruit Arf1. These results are in contrast to the other two assays of RalF activities and the impact of Glu-103 mutations on them but are consistent with the original prediction that Glu-103 is a critical residue for the activation of Arf1 by the Legionella-encoded Arf GEF.

DISCUSSION

The structure of RalF provides insights into the structure and function of Arf GEFs that are likely to be relevant to the eukaryotic proteins as well. The presence of the SCD over the active site provides direct evidence that domain reorganization is a required step in the activation of an Arf GEF and offers a novel mechanism of regulation of Arf GEF activities. The SCD contains a conserved structural motif, the platform domain, first described in subunits of clathrin-coated vesicle adaptors as a protein interaction domain. We believe that both additional protein-protein interactions as well as binding of lipids are required for RalF to become optimally effective as an Arf GEF in cells. In addition, the lack of effect of mutation of the conserved Glu-103 in two different assays argues against it serving an essential role in catalysis of nucleotide exchange despite the evidence that it is important to recruitment of Arf1 to membranes during establishment of the replicative organelle.

The Sec7 domain of RalF has the same overall folded structure seen in eukaryotic Sec7 domains, demonstrating that the bacterial and eukaryotic proteins are highly conserved structurally. This is not surprising because RalF is predicted to have arisen as a result of a gene transfer from a eukaryotic donor (8). For these reasons we conclude that information learned from the RalF structure will also be relevant to eukaryotic Sec7 domains and Arf GEFs. Although the eukaryotic Arf GEFs do not contain sequences with high homology to the SCD, we speculate that other domains in the eukaryotic proteins serve the same or related functions.

The location of the SCD, atop the Sec7 domain, is provocative in that it clearly precludes access of substrate when bound as in the structures reported here. Clearly, separation of the SCD from the Sec7 domain is required before Arf can bind at the active site. The number of contacts between the Sec7 domain and SCD indicates a stable interaction that will require other forces to separate them. The presence of a number of hydrophobic residues on the domain interaction surfaces suggests that membranes likely play a role in promoting or stabilizing the opened conformation, although we cannot exclude a role for
a (unidentified) protein partner in RalF activation. Indeed, the presence of the platform domain in the SCD suggests that protein partners will be found and are likely to play a role in activation. Arf GEFS lack transmembrane domains and are commonly found both in cytosol and bound to specific membranes. The presence of membrane interaction domains, e.g., PH domains, on some explains their affinity for membranes, but other Arf GEFS lack such defined membrane binding domains. Also, the observation that some Arf GEFS dissociate from membranes upon exposure to brefeldin A, a direct inhibitor of some Arf GEFS, provides evidence of a linkage between the Arf GEF activity and its membrane binding. Because RalF also has been localized to membranes where it is active, we believe that the activation of RalF includes its binding to membranes and opening of the two domains. Whether the SCD interacts directly with target membranes to a helper protein or both is currently being tested.

Another role of the SCD, specifically the very C terminus, is that it contains the sequence recognized by the Dot/Icm transimport system to target proteins for secretion into host cytoplasm. This recognition sequence has been mapped to the C-terminal 20 residues of RalF. C-terminal translocation domains have been identified for other proteins that are injected into host cells by type IV secretion systems (43–45). How these domains are positioned within the substrate proteins and whether they are available for interactions with type IV secretion system components have been important unanswered questions. Here, we show that there is a long α-helix comprised of residues 331–348 that projects the C-terminal translocation domain away from the folded SCD. Although the RalF355–374 translocation signal is disordered, the structure reveals that this domain is exposed to solvent and accessible for interactions with other protein. Thus, the RalF C-terminal translocation domain is free to bind to other components of the type IV secretion machine to facilitate secretion into host cells.

We identified within the SCD a folding motif that has been seen previously in three different vesicle coat protein subunits, including two of the subunits of the clathrin-coated vesicle adaptin complex AP-2 and the γ-subunit of the non-clathrin-coated vesicle complex COP-I. ArfI binds directly to both γ and β-subunits of COP-I in protein cross-linking experiments (46, 47), although we do not yet know where. This raises the possibility that the SCD may possess some affinity for Arfs distinct from the binding site present in the Sec7 domain that could either facilitate its activity as an Arf GEF or provide a distinct role in the recruitment of Arfs and stabilization of their binding to membranes. Whether or not Arfs bind to the platform domain in the SCD it is likely that other proteins will be found to do so based upon the precedent seen with the platform domain in the coat complexes.

The finding that RalF has a higher specific activity in the in vitro Arf GEF assay than does the isolated RalF-Sec7 domain demonstrates that instead of acting as a structural impediment to Arf binding, it plays a positive role in the catalysis of nucleotide exchange. How it does so is unknown but is likely to include one or more of the following possibilities: (i) binding between the SCD and Arf to help recruit and/or orient the substrate at the membrane interface, (ii) binding of lipids by the SCD to orient the two proteins for catalysis or simply increase their effective concentrations, and (iii) increased off-rate of product due the presence of the SCD. The known role of lipids in Arf nucleotide exchange and stabilization of the product, ArfGTP, makes dissection of the different roles of lipids in RalF action difficult (37–39). We lack experimental support for the first possibility, binding between the SCD and Arf, but are currently testing this possibility. Should the SCD bind to Arf at a site away from the guanine nucleotide, it could both increase the on-rate and participate in the rolling mechanism of binding and exchange that was recently proposed by Renault et al. (48).

Binding of Arf to a Sec7 domain involves several residues in addition to the glutamic finger (Glu-103 in RalF) (11, 24–27, 34, 49). The docking of Arf on the surface of a Sec7 domain leads to rearrangements in residues that promote the dissociation of GDP even in the absence of the key glutamate, although in other proteins the Arf GEF activity is reduced by at least 10-fold (24, 26, 27). Our observation that mutation of Glu-103 to either alanine or lysine had no deleterious effect on Arf GEF activity is currently interpreted as evidence that other residues in the Sec7 domain of RalF play more important and direct roles in the weakening of the binding of the Mg2+ and β- and γ-phosphates to Arf that promote the release of GDP than they do in the mammalian Sec7 domains examined to date. It will be interesting to see if any eukaryotic Arf GEFS become less dependent on the conserved glutamate when assayed as full-length proteins.

Three different assays were employed here to test for the impact of mutations in Glu-103 on (i) in vitro Arf GEF activity, (ii) perturbation of the early secretory pathway in cultured mammalian cells, and (iii) recruitment of Arf1-GFP in macrophages infected with the Legionella organism. The first two of these assays failed to find any effect of the charge-reversing E103K mutation that has such profound effects in the context of isolated mammalian Sec7 domains, whereas the third assay did find this to be an inactivating mutation. These data serve to
highlight the fact that, whereas all are used to measure Arf GEF activities, they are clearly not all measuring the same biochemical reaction. As most Arf GEFs and Arf GAPs are large, multi-domain proteins, it is increasingly evident that they can have functions beyond the activation and inactivation of Arfs, respectively. Studies of isolated domains are critical to initial tests of protein interactions and activities, but it will only be from studies of the native proteins that the complexity and significance of those interactions will be fully understood and appreciated.

Acknowledgments—We gratefully acknowledge the assistance and input from John Horton and James Prestegard during the course of this work. We thank Zygmunt Derewenda for the gift of some plasmids used in this study.

REFERENCES
1. Kahn, R. A., Fu, H., and Roy, C. R. (2002) Trends Biochem. Sci. 27, 308–314
2. Knodler, L. A., Celli, J., and Finlay, B. B. (2001) Nat. Rev. Mol. Cell Biol. 2, 578–588
3. Horwitz, M. A. (1983) J. Exp. Med. 158, 2108–2126
4. Kagan, J. C., Stein, M. P., Pypaert, M., and Roy, C. R. (2004) J. Exp. Med. 199, 1201–1211
5. Vogel, J. P., Andrews, H. L., Wong, S. K., and Isberg, R. R. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 1699–1704
6. Terwilliger, T. C. (2000) EMBO J. 19, 4216–4227
7. Terwilliger, T. C. (2000) Acta Crystallogr. D Biol. Crystallogr. 54 (Pt 5), 905–921
8. Terwilliger, T. C., and Berendzen, J. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 7909–7914
9. Terwilliger, T. C., and Berendzen, J. (1999) Cell 92, 415–423
10. Terwilliger, T. C., and Berendzen, J. (1999) J. Biol. Chem. 274, 2525–2545
11. Bamford, M., and Sander, C. (1993) J. Mol. Biol. 233, 123–138
12. Hoffman, G. R., Rahl, P. B., Collins, R. N., and Cerione, R. A. (2003) Mol. Cell 12, 315–265
13. Shuman, H. A. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 9653–9663
14. Terui, T., Kahn, R. A., and Randazzo, P. A. (1994) Biochemistry 33, 4675–4684
15. Randazzo, P. A., Yang, Y. C., Rulka, C., and Kahn, R. A. (1993) J. Biol. Chem. 268, 9555–9563
16. Randazzo, P. A., Terui, T., Sturch, S., Fales, H. M., Ferrige, A. G., and Kahn, R. A. (1995) J. Biol. Chem. 270, 14899–14915
17. Zheng, Y., Glaven, J. A., Wu, W. J., and Cerione, R. A. (1996) J. Biol. Chem. 271, 23815–23819
18. Kwon, J. L., Miura, K., Jackson, T. R., Gruschus, J., Roller, P., Staufer, S., Clark, J., Aneja, R., and Randazzo, P. A. (2000) J. Biol. Chem. 275, 9653–9663
19. Monier, S., Chardin, P., and Goud, B. (1998) J. Cell Sci. 111, 3427–3436
20. Franco, M., Boretti, J., Robineau, S., Monier, S., Goud, B., Chardin, P., and Chavrier, P. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 9926–9931
21. Vergunst, A. C., van Lier, M. C., den Dulk-Ras, A., and Hooykaas, P. J. (2003) Plant Physiol. 133, 978–988
22. Vergunst, A. C., Schrammjejer, B., den Dulk-Ras, A., and Hooykaas, P. J. (2003) Science 299, 979–982
23. Simione, M., McCullen, C. A., Stahl, L. E., and Binns, A. N. (2001) Mol. Microbiol. 41, 1283–1293
24. Zhao, L., Helms, J. B., Brunner, J., and Wieland, F. T. (1999) J. Biol. Chem. 274, 14209–14216
25. Austin, C., Hinners, I., and Tooze, S. A. (2000) J. Biol. Chem. 275, 21862–21869
26. Renaut, L., Guibert, B., and Cherfils, J. (2003) Nature 426, 525–530
27. Fuchs, R., Rodriguez, G., Meacci, E., Vitale, N., Moss, J., and Vaughan, M. (1998) J. Biol. Chem. 273, 28543–28548
28. Carson, M., and Bugg, C. E. (1996) J. Mol. Graph. 4, 121–122
29. Nicholls, A., Sharp, K. A., and Honig, B. (1993) Proteins 11, 281–296