Cloning and Characterization of a rhoGAP Homolog from Dictyostelium discoideum*

(Received for publication, January 29, 1997)

Steven B. Ludbrook‡, John F. Eccleston, and Molly Strom§

From the National Institute for Medical Research, Department of Physical Biochemistry, The Ridgeway, Mill Hill, NW7 1AA, London, United Kingdom

Small GTPases interact with a variety of proteins that affect nucleotide binding and cleavage. GTPase activating proteins (GAPs) are one class of these proteins that act by accelerating the intrinsic GTPase rate resulting in the formation of the biologically inactive GDP-bound form of the GTPase. For the Rho subfamily of GTPases, there is a growing number of proteins with rhoGAP activity that are identifiable by a homologous region of about 150 amino acids. We have exploited this homology using the polymerase chain reaction to clone the first rhoGAP homolog, called DdRacGAP, from the slime mold Dictyostelium discoideum. The GAP domain of DdRacGAP (amino acids 1–212), when expressed and purified from Escherichia coli, is active on both Dictyostelium and human Rho family GTPases but not human Ras. The full-length protein is 1356 amino acids in length and has several interesting homologies in addition to the GAP domain, including an SH3 domain, a dbl homology domain, and a pleckstrin homology domain.

Members of the Rho family of small GTPases have been shown to be key components in the transduction of extracellular signals to the actin cytoskeleton in mammalian cells resulting in focal adhesion and stress fiber formation (Rho-like proteins), membrane ruffling (Rac-like proteins), and filopodia formation (Cdc42-like proteins) (1–3). In addition to these effects on the cytoskeleton, the Rho GTPases are involved in various other cellular processes. Activation of Rho type GTPases leads to activation of MAP kinase pathways and progression of the cell cycle through G1 (4). Rac is also an essential component for the induction of phagocytosis in neutrophils (5, 6), and in fibroblasts, it has been suggested to regulate pinocytosis (2) while Rho has been shown to be involved in endocytosis (7, 8).

All small GTPases of the Ras superfamily cycle between a GTP and GDP bound state, and this cycling is a key feature of their biological activity (9). A variety of proteins that regulate this cycling has been characterized and include GTPase activating proteins (GAPs), nucleotide exchange factors (GEFs), and nucleotide dissociation inhibitors (GDIs) (10). One of the best characterized classes of interacting/regulatory proteins for small GTPases are the GAPs, which accelerate the intrinsic GTPase rate. RhoGAP was the first GAP for the Rho family of GTPases to be described biochemically and cloned (11). Based on sequence homology, two other proteins were immediately identified as potential rhoGAPs (bcr and n-chimaerin) and were later shown to have GAP activity (12). This highly conserved region of sequence homology has been the key for identifying potential new rhoGAP proteins, and with only one exception (the p85 subunit of phosphatidylinositol 3-kinase), these proteins have all been shown to have GAP activity (13, 14).

In general, rhoGAPs are unspecific in that they have GAP activity toward more than one member of the Rho family; though they are generally inactive toward Ras or members of other small GTPase families (with the exception of CeGAP and RLIP76, both of which have been suggested to bridge the Ras pathway (15, 16)). The rhoGAP domain is approximately 150 amino acids in length and is most often situated toward the C terminus of these proteins. RhoGAPs range in size from 34 kDa to >150 kDa and contain additional regions that are involved in protein-protein interactions, including SH2 and SH3 domains, proline-rich regions (SH3-binding motifs), or lipid interaction (pleckstrin homology (PH) domain). Two rhoGAPs (bcr and abr) also contain a region of homology (known as the DH domain) to the dbl oncogene, a GEF, and this domain has exchange activity toward Rho GTPases (17).

The role of GAP as an upstream negative regulator is evident in studies done both in vitro and in vivo. Biochemical assays have established that GAP accelerates the intrinsic GTPase rate, thereby converting the Rho protein from the active GTP bound form to its inactive GDP-bound form. Microinjection of different rhoGAP proteins into serum-starved 3T3 fibroblasts showed that they were able to block induction of stress fibers (p190GAP and rhoGAP) and membrane ruffles (bcr) (18). In the case of Rac-stimulated NADPH oxidase, the p190GAP has been shown to inhibit superoxide formation and subsequently phagocytosis in neutrophils (19). Other potential functions of GAPs are well defined. The fact that GAP proteins have a variety of domains that are involved in protein-protein interactions that are not necessary for their GAP activity implies that they do at least interact with other proteins, and whether this is part of their function as negative regulator or part of some effector-like role is not yet clear.

The slime mold Dictyostelium discoideum is a professional phagocyte (23) that feeds by ingestion of bacteria and other particles in a fashion similar to mammalian leukocytes or neutrophils. D. discoideum is amenable to molecular and genetic techniques and has a well characterized cytoskeleton in addition to phagocytic and pinocytic pathways. These features
Characterization of a D. discoideum rhoGAP Homolog

Fig. 1. Genomic organization and amino acid sequence of DdRacGAP. A, genomic organization of DdRacGAP locus. The coding sequence extends over 4 kb and is interrupted only twice by introns, designated I1 and I2. Start and stop codons are indicated. The three genomic DNA clones are also indicated. Restriction enzyme sites are abbreviated: X, XhoI; E, EcoRI; S, SpeI; C, ClaI. B, top, a schematic diagram showing the arrangement of the domains found in DdRacGAP based on sequence homology. B, bottom, the primary amino acid sequence of DdRacGAP is shown. The amino acids encoded by the Agt11 cDNA clone are underlined. The DNA sequence has been deposited in the GenBankTM/EBI Data Bank (accession number Y10159).

make it a good model system in which to study Rac-related pathways. Three Rac1/2 homologs (Rac1A, Rac1B, and Rac1C) have been described in D. discoideum (20). These proteins have over 90% identity to each other and over 80% identity to the human Rac1/2 GTPases (20). In addition, five more GTPases have been described, RacA, B, C, D, and E (20, 21). These proteins have the most homology to human Rac, especially in the N-terminal 100–120 amino acids but are not close homologs to human Rho or Cdc42. Although the exact functions of the D. discoideum Rac proteins are unknown, RacE has been shown to be involved in cytokinesis (21), indicating that these proteins will also be important in regulating cytoskeletal-dependent events.

We report here the cloning of a rhoGAP homolog from D. discoideum, which we have called DdRacGAP since all the homologous regions of the domains found in DdRacGAP based on sequence homology to human Rac proteins. The putative protein product is quite large at approximately 150 kDa. The GAP domain shows 20–25% identity to other rhoGAPs and is situated at the N terminus of the protein, which is an unusual location for this domain. The protein also has an SH3 domain situated at the N terminus of the protein, which is an unusual sequence homology to the human Rac proteins. The putative amino acid sequence of DdRacGAP.

EXPERIMENTAL PROCEDURES

Materials—All restriction enzymes and T4 DNA ligase were from Boehringer Mannheim. Taq polymerase and buffers were from Cetus. All radiochemicals were from Amersham Int. Sequencing was done using the Sequenase version 2.0 kit (U.S. Biochemical Corp.). All restriction enzymes and T4 DNA ligase were from Boehringer Mannheim. Taq polymerase and buffers were from Cetus. All radiochemicals were from Amersham Int.

Preparation of pGEX2T Constructs—For rac1A and racC from D. discoideum, oligonucleotides were prepared corresponding to the published cDNA sequences (20), which also created a 5′ BamHI and a 3′ EcoRI restriction enzyme site. These oligonucleotides were used in the polymerase chain reaction (PCR) to amplify the cDNAs from a D. discoideum Agt11 cDNA library (CLONTECH), which were subsequently cloned into BamHI-EcoRI-digested pGEX2T (Pharmacia Biotech Inc.) and sequenced to ensure no errors had occurred. The Q61L mutation in Rac1A was introduced by 2-step PCR as described (22). For DdRacGAP expression, two oligonucleotides, designed to the 5′ and 3′ ends of the Agt11 cDNA clone which also created a BamHI site at each end, were used in PCR to amplify the insert DNA. The DNA fragments were cloned into the BamHI site of pGEX2T. Initially, it was not clear that the ATG at bp 35 did encode the start codon so the GST fusion protein contains additional amino acids (GPDFFDIK) before the initiation methionine.

Expression and Purification of GST-fusion Proteins—Escherichia coli cells harboring the various pGEX2T plasmid constructs were grown at 37 °C to an A600 of 0.6–0.8 cm−1 and then induced by addition of IPTG to 0.1 mM for 16–18 h at 23 °C. GST-tagged proteins were affinity purified and thrombin cleaved at 4 °C according to the following procedure. E. coli cells were lysed in buffer A (50 mM Tris, 50 mM NaCl, 10 mM MgCl2, 1 mM DTT, pH 7.5) and centrifuged at 100,000 × g for 1 h, and the soluble protein fraction was passed over a 10-mL GST-agarose column and then washed with buffer A until the A280 was ~0.05 cm−1. Protein was eluted by cleavage with thrombin (10 units/ml) (Sigma) for 10 h at a flow rate of 0.08 ml/min in wash buffer containing 2.5 mM CaCl2 and the thrombin was removed by passage of the eluate over a 3-m-phenanthroline-agarose (Sigma) column. Peak fractions were concentrated using a Centricon-10 (Amicon), and protein concentration was measured using the Bio-Rad protein assay (Bio-Rad). The purity of the small GTPases was greater than 90%, and the purity of DdRacGAP was about 30% (see “Results and Discussion”), as judged by Coomassie Blue staining after SDS-polyacrylamide gel electrophoresis.

Cloning and Sequencing of DdRacGAP cDNA—Genomic DNA—4′ and 3′ oligonucleotides (5′-AAAAACCTTCTTGACGAAGA-3′ and 5′-GAGGAAAGATTCCACTTTG-3′), corresponding to the KLF1REL and NKMNTNTLAV, conserved amino acids in blocks 2 and 3, respectively (see Fig. 2A), were used in PCR amplification to isolate GAP homologous DNA fragments from a D. discoideum Agt11 cDNA library (CLONTECH). The amplification protocol was as follows: 1 cycle of 92 °C for 2 min, 43 °C for 3 min, and 72 °C for 1 min; 10 cycles of 92 °C for 30 s, 53 °C for 30 s, and 72 °C for 1 min; 10 cycles of 92 °C for 30 s, 53 °C for 30 s, and 72 °C for 1 min; and finally 10 cycles of 92 °C for 30 s, 57 °C for 30 s, and 72 °C for 1 min. The DNA fragments were cloned into EcoRV-digested Bluescript vector (Stratagene) and sequenced. The DNA from a positive clone was ATP3(dCtP)-labeled with...
Characterization of a D. discoideum rhoGAP Homolog

A

| Block | Location | Sequence 1 | Sequence 2 |
|-------|----------|------------|------------|
| 1     | 1061     | VFGVXIAVYXKX | ENKX |
|       | 381      | VFGVXIAVYXKX | ENKX |
|       | 241      | VFGVXIAVYXKX | ENKX |
|       | 6        | VFGVXIAVYXKX | ENKX |

The Ready to Go kit (Pharmacia Biotech Inc.) and used as a probe to screen the GT11 cDNA library. The cDNA clone insert was 677 bp and included the initial methionine to amino acid 212. For genomic DNA cloning in each case, DNA of the appropriate size was gel-purified after digestion with restriction enzymes and separation on a 0.8% agarose gel. The DNA fragments were ligated to gel-purified pBluescript after digestion with the same restriction enzymes and treatment with shrimp alkaline phosphatase (U. S. Biochemical Corp.). The ligations were transformed into Ultracompetent XL1-BLUE cells (Stratagene), and the entire library was screened using either radiolabeled EcoRI-SpeI DNA (for genomic DNA clones 1 and 2) or SpeI-ClaI DNA (for genomic DNA clone 3) as a probe. The inserts were sequenced on both strands from approximately 500 bp upstream of the start codon to 200 bp downstream of the stop codon. The sequences of all PCR-amplified DNA fragments were checked against this genomic DNA sequence to ensure that no errors had occurred.

B

GAP Assays—Complexes of Rho proteins with \( [\alpha-32P]GTP \) were prepared according to the following procedure. 1 mM Rho GDP (about 1.0 \( \mu \)g of protein) was incubated with 0.25 \( \mu \)M \( [\alpha-32P]GTP \) (3000 Ci/mmol) in 40 \( \mu \)l of 50 mM Tris-HCL, 5 mM NaCl, 1 mM DTT, 4 mM EDTA-Na2, pH 7.5, at 25 °C for 5 min, after which MgCl2 and unlabeled Rho GDP as a carrier were added to 30 mM and 5 mg/ml, respectively. Labeled protein was separated from free nucleotide using a PD-10 gel filtration column (Pharmacia Biotech Inc.) in assay buffer (50 mM Tris-HCl, 2 mM MgCl2, 1 mM DTT, 4 mM EDTA-Na2, pH 7.5). For GAP assays, 100-\( \mu \)l reaction mixtures containing 2 nM \( [\alpha-32P]GTP \) labeled protein, 1 mM ATP, and either assay buffer for control reactions or dilutions of GAP protein were incubated at 15 °C. Samples (10 \( \mu \)l) were removed at intervals, and the nucleotide was separated from the protein by the addition of perchloric acid to 0.5%. The samples were brought to approximately pH 4 by adding sodium acetate to 140 mM, centrifuged briefly to remove the protein, and the nucleotide solution (representing 2–4 \( \mu \)l of the original 100-\( \mu \)l reaction volume) was spotted onto polyethyleneimine-cellulose thin layer chromatography plates (Sigma). Nucleotides were separated by eluting the plate with 0.75M KH2PO4, pH 3.4. The samples were analyzed on a Beckman LS7000 scintillation counter to determine the percentage of Rho \( [\alpha-32P]GTP \) that was hydrolyzed. The percentage of Rho \( [\alpha-32P]GTP \) hydrolyzed was calculated as the difference between the percentages of Rho \( [\alpha-32P]GTP \) that were hydrolyzed in the presence and absence of GAP protein.

C

Alignment of GAP, DH, PH, and SH3 domains. In all cases, the sequences were first aligned using the BLITZ program and then adjusted by eye to make any improvements necessary. Figure 2, A, alignment of the GAP domain showing the three highly conserved blocks as in Ref. 11. The accession numbers of the genes used in the line-up are: human bcr, P11274; human n-chimaerin, X51408; mouse 3BP-1, X87671; and human rhoGAP, Z23024. B, alignment of the DH domain showing the three structurally conserved regions (SCRs) (10). The accession numbers of the genes used are: human dbl, P10911; human Tiam-1, U16296; human FGD-1, U11690; and human bcr as in panel A. C, alignment of the PH domain. Conserved amino acids, which are a majority at that site, are shown in bold, and the overall organization (six homologous blocks) as described in (28) is maintained. The accession numbers for the genes used are: human pleckstrin, P08567; human \( \beta \)-adrenergic receptor kinase 1, denoted as ARK, P25098; and human dbl as in panel B, D, alignment of the SH3 domain. Amino acids conserved in at least 4/6 of the sequences are shown in bold. The accession numbers for the genes used are: human src, P12831; human crk, D10656; and \( D. \) discoideum myosin ID heavy chain, L16509; human vav, P15498; and human grb2, M96995.
Characterization of a D. discoideum rhoGAP Homolog

RESULTS AND DISCUSSION

Cloning of DdRacGAP from D. discoideum—To clone GAP homologs from D. discoideum, we used a PCR-based approach. Two oligonucleotides, corresponding to conserved sequences KYLRELP in block 2 and NKMTNTNLAV in block 3, in rhoGAPs (see Fig. 2A) were prepared and used as primers to amplify from DNA isolated from a D. discoideum cDNA library. DNA products of approximately 200 bp were cloned and sequenced, and one isolate showed homology to rhoGAP. This PCR fragment was then used as a probe to isolate a cDNA clone from the agT11 library, which was approximately 700 bp and included the start codon (Fig. 1B, the underlined amino acids correspond to those encoded by the cDNA clone). As introns in D. discoideum are generally short and easily identifiable, the remaining 5′ promoter region and 3′ coding sequences were isolated in three overlapping genomic DNA clones: Clone 1, the 2.0-kb Xho-I-SpeI fragment containing the promoter and first 180 amino acids; Clone 2, the 2.0-kb EcoRI-ClaI fragment; and Clone 3, the 4.5-kb SpeI fragment containing the stop codon (Fig. 1A). Fig. 1A shows the organization of the DdracGAP gene and Fig. 1B shows a schematic diagram of putative protein product and the amino acid sequence. DdRacGAP has several interesting protein homologies in addition to the GAP domain (Fig. 1B), including an SH3 domain and a DH-PH domain homology. The presence of an SH3 domain is unusual in the rhoGAP family of proteins as it is only seen in two members known to date (p85, the 85-kDa subunit of phosophatidylinositol 3-kinase (13), and the hematopoietic C1 protein (24)). The DH-PH combination is seen in all RhoGEFs and in many oncogene products (36), several of which have been shown to have exchange activity for Rho GTPases, but only two other GAPs (bcr and abr) have this homology in addition to the GAP activity although several rhoGAPs have the PH domain alone (13). Another feature is the long stretches of asparagine, threonine, and serine (amino acids 330–440, see Fig. 1B), which is commonly seen in D. discoideum proteins and is of unknown function.

The location of the GAP domain (abutting the N terminus) is unique compared with all other rhoGAPs, but the significance, if any, of this is unknown. Fig. 2A shows the alignment of the GAP homology domain. The highest degree of homology is to bcr (27% identity) and n-chimaerin (23% identity), both active on Rac but not Rho. Recently, the structure of the rhoGAP has been solved (25), and the structure implicates the potential importance of one amino acid in particular, Asn-391 in rhoGAP, which is present in all active GAPs but absent in p85. This amino acid is also present in the DdRacGAP (Asn-162).

The DH domain alignment is shown in Fig. 2B. The highest homology was to the FGD-1 protein (21% identity) and to the RhoGEF domain of RasGRF (20% identity), and while the identity to dbl, bcr, and vav is only around 10–12%, DdRacGAP does have almost all of the highly conserved amino acids in the three structurally conserved regions (SCRs) indicative of this domain (see Fig. 2B) (10). We have not yet been able to test whether this domain is active as an exchange factor as shown with the bcr and abr proteins (17) since this assay requires post-translationally processed Rho protein. All of the Rho exchange factors have a PH domain located C-terminal to the DH domain, and this is also the case with the DdRacGAP as seen in Figs. 1B and 2C. The PH domain homology was very low (<10% identity), but the Blitz search did identify several proteins containing PH domains. The alignment shown in Fig. 2C indicates the six main homology blocks based on an alignment of 45 PH domains (26), and many of the key conserved amino acids are present in DdRacGAP. PH domains are thought to be involved in intracellular localization of proteins via their varying affinity for different types of lipids (27). In keeping with this, it was shown that the PH domain of dbl was not necessary for interaction with GTPases but was important for its transforming activity (28).

DdRacGAP also contains an SH3 domain that is just C-terminal to the GAP homology domain (Fig. 2D). DdRacGAP showed 20–25% identity to the SH3 domains of a variety of other proteins, and again the amino acids most highly conserved are for the most part conserved in DdRacGAP (Fig. 2D). SH3 domains are known to bind proline-rich sequences approximately 10 amino acids in length. Finally, there is a region rich in proline (amino acids 450–770, see Fig. 1B), suggesting either SH3 binding motifs (present in several other rhoGAPs) or target sequences for proline-directed Ser/Thr kinases. The presence of SH3 and PH domains and a proline-rich region, all commonly found in molecules involved in signal transduction, in addition to the DH domain, suggest that DdRacGAP interacts with other proteins and may have other functions, possibly as part of a larger protein complex as is becoming a recurring theme in the biological activity of many small GTPases (10).

Specificity of DdRacGAP—To determine whether the GAP domain can in fact accelerate the intrinsic GTPase rate of any of the D. discoideum Rac proteins, we expressed and purified the N-terminal fragment containing the GAP domain (amino acids 1–212 of DdRacGAP) from E. coli (see “Experimental Procedures”). Over 95% of the expressed protein was insoluble even when the cells were grown at 23 °C, but the remaining soluble protein was isolated to about 30% purity. The major

| Protein     | GAP activitya | mg/ml | b |
|-------------|---------------|-------|---|
| DdRac1A     | 31            |       |   |
| DdRacC      | 13            |       |   |
| HsRhoA      | 113           |       |   |
| HsRac1      | 6.2           |       |   |
| Ha-N-ras    | <0.5b         |       |   |

a GAP activity is expressed as 1/[DdRacGAP] preparation required to double the intrinsic initial rate of GTP cleavage.

b For ras, there was no acceleration of GTPase at the maximal concentration of GAP that could be used in the assay (1.840 mg/ml).
contaminants were at 60 and 70 kDa and are most likely the E. coli proteins groEL and dnaK, respectively (29). This partially purified protein was used to characterize the DdRacGAP activity. Fig. 3 shows the results of incubation of DdRacGAP with [γ-32P]GTP-labeled Rac1A and RacC. The acceleration of the initial rate of GTP cleavage by DdRacGAP was approximately 20-fold faster for the Rac1A protein than for RacC (Fig. 3 and Table I). DdRacGAP accelerates the cleavage of Rho·GTP to Rho·GDP and is not active on a Rac1A Q61L mutant (data not shown), showing that the GAP activity was not due to nonspecific phosphatase activity or any nonspecific effect of the partially purified protein or contaminating proteins.

We further characterized the substrate specificity of the GAP domain of DdRacGAP by testing its ability to act on the human RhoA, Rac1, and Ras GTPases as shown in Table I. DdRacGAP did have GAP activity for human RhoA and Rac1 proteins but had no detectable activity on human Ras. In fact, it was about 3-fold more active on the human RhoA protein compared with the Rac1A protein, raising the possibility that Rac1A may not be the primary target in D. discoideum. Of the reported D. discoideum Rac proteins, RacC and RacD have the least homology to the human Rac1 and D. discoideum Rac1A, 1B, and 1C proteins (about 60% identity) while RacA and RacB have about 70% identity. Since RacC and RacB were complete sequences (RacA and RacD were missing both N- and C-terminal amino acids), we intended to test both RacC and RacB in the GAP assay, but we were not able to isolate enough of RacB protein for our assay. RacC, however, was not as good a substrate as Rac1A as seen in Fig. 3 and in Table I, indicating that the DdRacGAP does exhibit some substrate specificity among the D. discoideum Racs.

Comparison of the primary sequences of the D. discoideum Rac proteins indicates that, as is the case with other Rho proteins, amino acids outside of the effector loop (amino acids 32–42 of Rac1A) are important for interaction with GAP (30, 31). This is supported by the observation that the effector domains of the reported D. discoideum Racs are highly conserved. In fact, the effector domains of Rac1A and RacC are identical, and yet DdRacGAP shows a strong preference for Rac1A over RacC as a substrate. In the case of rhoGAP, the use of Rho/Ras chimeras indicated that the region around amino acids 74–90 of human RhoA (loop 7 of Ras) are important for activation (30). It was also found that the same region of Rac is required for activation by bcr (32). Diekmann et al. (32) reported that the effector domain and flanking regions of Rac (amino acids 22–46) were important for activation by bcr although the effect was much less than for the region between amino acids 74–90. The regions directly flanking the effector domain of Rac1A and RacC, in particular the amino acid C-terminal to the effector domain, are not conserved between the two proteins, and this may be a region to investigate for interaction with DdRacGAP. Rac1A and RacC also have several amino acid differences in the region corresponding to 74–90 of human RhoA in addition to a highly divergent C terminus (from approximately amino acid 120 to the end) (20), suggesting that these are also likely to be candidate regions for interaction with DdRacGAP as is the case with several other rhoGAPs (30–32).

General Discussion—We have identified the first rhoGAP for the Rho small GTPases from D. discoideum. This protein has an intriguing primary structure exhibiting homology to both GAP and GEF, which is a seemingly contrary juxtaposition of functional domains, and is only seen in two other GTPases (bcr and abr). The significance of this is not yet clear but a more in-depth analysis of the biochemical activity of these two domains, both together and independently, and cellular localization of DdRacGAP, will hopefully provide more information on the possible function of such a protein. Indeed, experiments analyzing the DH domain of several oncogene products have indicated that this domain itself may act as a bridge for signaling between two pathways since there are DH-containing proteins that bind to Rac without effecting nucleotide exchange but do have exchange activity for either Rho or Cdc42 (33–35). One could envisage a similar scenario where DdRacGAP could bridge two pathways via its two different activities by the possibility of the GAP domain acting to turn-off a GTPase in one pathway while the GEF could activate (or at least bind) the GTPase of another. Further experiments looking at the substrates for both binding and enzymatic activities should indicate if this is a possibility. In any case, it is very important to establish what the primary targets of the GAP activity and of the DH domain are to begin to analyze this experimentally.