A Conserved Binding Motif Defines Numerous Candidate Target Proteins for Both Cdc42 and Rac GTPases

(Received for publication, August 3, 1995)

Peter D. Burbelo, David Drechsel, and Alan Hall‡

From the Medical Research Council Laboratory for Molecular Cell Biology, CRC Oncogene and Signal Transduction Group, and the Department of Biochemistry, University College London, London WC1E 6BT, United Kingdom

Rho, Rac, and Cdc42 are small GTPases that regulate the formation of a variety of actin structures and the assembly of associated integrin complexes, but little is known about the target proteins that mediate their effects. Here we have used a motif-based search method to identify putative effector proteins for Rac and Cdc42. A search of the GenBank™ data base for similarity with the minimum Cdc42/Rac interactive binding (CRIB) region of a potential effector protein p65PAK has identified over 25 proteins containing a similar motif from a range of different species. These candidate Cdc42/Rac-binding proteins include family members of the mixed lineage kinases (MLK), a novel tyrosine kinase from Drosophila melanogaster (DPR2), a human protein MSE55, and several novel yeast and Caenorhabditis elegans proteins. Two murine p65PAK isoforms and a candidate protein from C. elegans, F09F7.5, interact strongly with the GTP form of both Cdc42 and Rac, but not Rho in a filter binding assay. Three additional candidate proteins, DPR2, MSE55, and MLK3 showed binding to the GTP form of Cdc42 and weaker binding with Rac, and again no interaction with Rho. These results indicate that proteins containing the CRIB motif bind to Cdc42 and/or Rac in a GTP-dependent manner, and they may, therefore, participate in downstream signaling.

Members of the Ras superfamily of small GTPases play a wide variety of cellular signaling roles that mediate proliferation and differentiation, cytoskeletal organization, protein transport, and secretion. The Ras GTPases have been studied most thoroughly, and now several components of the Ras signaling pathway have been identified using a combination of biochemical and genetic approaches (1, 2). A related family of GTPases, the Rho subfamily, consists of three Rho genes, two Rac genes, Cdc42 and its close homologue G25K, rhoG, and TC10 (3). Early work in Saccharomyces cerevisiae identified CDC42Sc as a protein required for bud emergence (4, 5). In mammalian cells, the Rho subfamily members control the polymerization of actin and the assembly of focal complexes at the plasma membrane in response to extracellular signals (3, 6). For example, microinjection of Rho into serum-starved Swiss 3T3 cells rapidly stimulates stress fiber and focal adhesion formation (7), while Rac induces membrane ruffles (8) and Cdc42 induces the formation of filopodia (9). In addition to their effects on the actin cytoskeleton, Rho GTPases also have a role in regulating kinase signaling pathways. For example, Rho, Rac, and Cdc42 stimulate a novel nuclear signalling pathway leading to transcriptional activation of the serum response element (10). Rac and Cdc42, but not Rho, have also been shown to activate the c-jun amino-terminal kinase (JNK) signaling pathway leading to c-jun transcriptional activation (11, 12). The mechanisms by which the Rho subfamily of GTPases regulate these apparently diverse biological processes is still not clear.

A large number of mammalian nucleotide exchange factors (GEFs), related to the yeast exchange factor Cdc24 have been identified including dbl, vav, ost, ect2, lbc, and Tiam1, and they may provide tissue specificity or receptor-specific activation of Rho family members (2, 3). In addition, over 10 mammalian GTPase-activating proteins (GAPs)1 for Rho family members have been described including BCR, p50rhoGAP, chimaerin, ABR, p190-A, p122, and myr-5 (13). The ability of these GAP proteins to interact with GTPases in a GTP-dependent manner suggests that in addition to being negative regulators, they may also act as effector proteins. The multidomain nature of many of these GAPs further supports this notion; for example, in addition to the GAP domain, p190 has a GAPase domain (14), p122 interacts with phospholipase δ (15), and the myosin family member, myr5, binds to actin (16).

The identification of effector proteins for Rho-related GTPases is the first step toward defining their biological activity, and a number of candidate proteins have already been reported. The target protein involved in Rac-mediated activation of the NADPH oxidase complex in phagocytic cells has been identified as p67phox (17). Although the Rac interactive site has been mapped to the amino-terminal 199 amino acids of the protein, no significant similarities have so far been found between this sequence and other sequences in the data base. A tyrosine kinase containing a SH3 domain, p120ACK, has been shown to bind specifically to Cdc42 in a GTP-dependent manner, although its biological role is unclear (18). Another protein, the serine/threonine kinase p65PAK, binds to both Cdc42 and Rac (but not Rho) in a GTP-dependent manner, but again it is not known whether p65PAK mediates any of the described biological effects of these GTPases (19, 20). Interestingly, p120ACK and p65PAK do have some sequence similarities in a region outside the kinase domain that represents an interactive site for Cdc42/Rac GTPases (19).

In this report, we have localized the Cdc42 binding site of a murine p65PAK isoform to a minimal conserved region of 16 amino acids. Using this small protein motif, we have searched the GenBank™ data bases and identified 25 potential Cdc42- and/or Rac-binding proteins. In vitro binding assays confirm that several of these proteins bind to Cdc42 and/or Rac in a

1 The abbreviations used are: GAP, GTPase-activating protein; GST, glutathione S-transferase; MLK, mixed-lineage kinase; PCR, polymerase chain reaction.
GTP-dependent fashion. Not all of these newly identified proteins are kinases, suggesting a role for other types of proteins in downstream signaling events mediated by the Rho family proteins.

**MATERIALS AND METHODS**

Recombinant Protein Production—Two murine p65PAK isoforms were obtained by screening a 14-day embryonic library with a PCR product from the kinase domain (19) of the rat p65PAK.2. One of the isoforms, p65PAK-α, is the mouse homologue of rat p65PAK (98% amino acid identity), while the second, p65PAK-β, is a distinct isoform (81% identity to rat p65PAK). DPR2 (21) was a kind gift from Dr. T. Matsui (Kobe University School of Medicine, Kobe, Japan). The MSE55 cDNA has been described previously (22) and was a gift from Dr. W. Bahou (State University of New York). A Caenorhabditis elegans cDNA, cml2 g10, coding for the predicted F09F7.5 gene product, was a kind gift from Dr. L. Naborocki (University College London, London, United Kingdom). MLK3 was cloned using the polymerase chain reaction (PCR) from HT-1080 mRNA using the published nucleotide sequence (23). PLC-1 cDNA was described previously (24) and was a gift from Dr. P. Parker (ICRF, London, UK). Native major sperm protein (MSP) protein was purified from Ascariis suum (25) and was a gift from Dr. M. Stewart (MRC at Cambridge University).

cDNA fragments were generated by restriction enzymes or PCR and subcloned into the pGEX-4T-3 bacterial expression vector. All constructs were confirmed by DNA sequence analysis. Fusion proteins were made as glutathione S-transferase (GST) fusion proteins in bacteria by isolopropyl-1-thio-d-galactopyranoside treatment, purified on a glutathione-affinity column as described by the manufacturer (Pharmacia Biotech Inc.), and checked for protein integrity by SDS-polyacrylamide gel electrophoresis. The following constructs were used: p65PAK-α (full-length, residues 1–545), p65PAK-α−1 (residues 29–546), p65PAK-β−1 (residues 118–546), p65PAK-β−3 (residues 29–90), DPR2 (residues 460–541), MSE55 (residues 11–120), F09F7.5 (residues 12–59), MLK3 (residues 454–538), and PLC-1 (residues 560–728). Syn-1 was derived using two complementary oligonucleotides containing the amino acid sequence EISLARREFHLNHVGLE and subcloned in-frame into the EcoRI-XhoI site of pGEX-4T-3.

Filter Binding Assay for p21 Proteins—Protein-protein interactions were visualized using a dot-blot assay. GST-fusion proteins were spotted onto a nitrocellulose filter which was subsequently incubated with proteins immobilized by glutathione-Sepharose. The nucleotide requirement for binding to each of the beads was expressed as a percentage of input binding.

Identification of Potential Cdc42 and Rac GTPases

Identification of Targets Proteins for Cdc42 and Rac GTPases

Two murine p65PAK isoforms, p65PAK-α and p65PAK-β, were expressed as GST-fusion proteins and, using a filter binding assay, were shown to bind to Rac and Cdc42, but not to Rho as expected (Fig. 1). A truncated p65PAK-β construct, GST-p65PAK-β−Δ3, continued to show strong Cdc42 binding, although binding to Rac was reduced (Fig. 1). The ability of p65PAK-β−Δ3 to bind to Cdc42/Rac and its limited amino acid homology with the two other Cdc42/Rac-binding proteins p120ACK and STE20 suggested that a region of approximately 16 rather than 40 amino acid residues (corresponding to amino acid residues 74–89 within rat p65PAK-α) might be sufficient for binding to Cdc42/Rac.

Identification of Potential Cdc42 and Rac Binding Motifs—In order to identify potential new binding proteins for Cdc42 and Rac, we performed an iterative search of the GenBank™ protein data base using the BLASTp algorithm (26). The initial query sequence with the conserved Cdc42/Rac binding site in rat p65PAK-α (residues 74–89) revealed significant matches to numerous other proteins shown in Fig. 2. Sites within these proteins along with the corresponding sequences in p120ACK and STE20 were then used to research the data base to identify additional proteins containing this motif. Finally, a FASTA (27) search of GenBank™ with each of the potential proteins turned up partial cDNAs for additional proteins. In total, over 25 distinct proteins from a variety of organisms were identified that contain a potential Cdc42 binding motif (Fig. 2). We propose that this motif be termed CCRIB for Cdc42/Rac interacting binding motif.

The CCRIB motif occurs in both kinases and non-kinase proteins. For example, one of the CCRIB-containing proteins, DPR2, is a Drosophila tyrosine kinase (21), which shows similarity in the kinase domain to pp125FAK, c-Abl, and p120ACK. Several serine/threonine kinase including at least 3 isoforms of p65PAK and two human and three yeast kinases also contain a potential CCRIB motif. In addition, several members of the MLK family of serine/threonine kinases also contain a protein sequence resembling a CCRIB motif. Two different yeast kinases encoded by X82499 and Z48149, contain identical CCRIB motifs. Non-kinase CCRIB-containing proteins include a human sequence MSE55 (and two related genes, accession numbers T06431 and T75138/F12871), WASP, a human gene responsible for Wiscott-Aldrich syndrome (28), three different C. elegans genes (FO9F7.5, T2365.3, and T75138/F12871) and two potential yeast genes (P38785 and D9740.18).

CRIB Proteins Bind to Cdc42 and/or Rac—Several of the proteins containing the CRIB motif were subcloned into the pGEX-4T-3 bacterial expression vector to test the prediction that these proteins would interact with Cdc42 and/or Rac. Fusion proteins were purified by glutathione affinity chroma-
Sequence alignment of proteins containing the CRIB motif. The accession number and/or name of each of the proteins is shown on the left. The species origin is denoted by: R, rat; M, mouse; M, human; S, S. cerevisiae; Sp, S. pombe; C, C. elegans; B, bovine; D, Drosophila; and A, Ascaris. Amino acid sequence comparison is shown between the different CRIB proteins. The number of residues matching the strict consensus sequence of eight amino acid residues is also shown. Sequence analysis suggests that C09B8.7 is a C. elegans homologue of p65ACK. Binding of Cdc42 and Rac determined experimentally is shown as positive (+) or negative (−). P21 binding to p65ACK-Rho, p65ACK-γY, p120ACK, and -Rho is shown as positive (+) or negative (−). The accession number and/or name of each of the proteins is shown on the left. The species origin is denoted by: R, rat; M, mouse; H, human; S, S. cerevisiae; Sp, S. pombe; C, C. elegans; B, bovine; D, Drosophila; and A, Ascaris. Amino acid sequence comparison is shown between the different CRIB proteins. The number of residues matching the strict consensus sequence of eight amino acid residues is also shown. Sequence analysis suggests that C09B8.7 is a C. elegans homologue of p65ACK. Binding of Cdc42 and Rac determined experimentally is shown as positive (+) or negative (−). P21 binding to p65ACK-Rho, p65ACK-γY, p120ACK, and -Rho is shown as positive (+) or negative (−).

In conclusion, we have delineated a short motif (the CRIB motif) in over 10 distinct proteins that confers binding to the Cdc42 and/or Rac GTPases. At least another 15 proteins have been identified by a database search that contain a potential CRIB site. The length of the CRIB motif is approximately 16 amino acids containing a region of variable length between the two halves of the binding motif. The CRIB motif contains eight core amino acids with the sequence I-S-X-P-(X)2-4-F-X-H-X-H-V-G. It is interesting to note that proteins with one or two differences within the core sequence can still show binding to Cdc42/Rac.

Of the proteins shown to contain a functional CRIB motif, only the kinases have defined biochemical activity. In addition to the SH3-containing kinase p120ACK, a new Drosophila tyrosine kinase, DPR2, was found to contain the CRIB motif. Although the CRIB domains are found at a relatively similar distance after the kinase domains, DPR2 is unlikely to be the Drosophila homologue of p120ACK, since it lacks a SH3 domain and the overall amino acid similarity is low. In addition to the p65ACK family of kinases, a new serine/threonine kinase family, the MLK kinases, were found to contain a potential CRIB motif. To date, there appears to be at least four members of the MLK family including MLK1 (30), MLK2/MAST (30, 31), MLK3/SPRK/PTK1 (32, 32, 33), and DLK (34). All four of these kinases share a characteristic hybrid kinase domain between serine/threonine and tyrosine kinases, although SPRK/MLK3 (23) and DLK (34) have now been shown to possess serine/threonine catalytic activity. In addition, all four MLK members contain putative α-helical leucine zipper motifs COOH-terminal to the kinase domain, although little is known of the function of this region. One might expect that the activity of at least MLK2 and MLK3 kinases will be stimulated following Cdc42 or Rac binding as seen with p65ACK (19, 20), but this has yet to be tested.

The biochemical function of the non-kinases containing the CRIB motif is not clear at all. Some of these proteins also contain proline-rich regions (e.g. F097.5 and M55555) raising the possibility that they might act as adapters and interact with SH3-containing protein(s) or other polyproline-binding proteins such as profilin.
FIG. 4. **Proteins containing the CRIB motif bind to Cdc42 in a GTP-dependent fashion.** Approximately 5 μg (or 10 μg of MLK3 and Syn-1) of each GST protein was immobilized to the glutathione-Sepharose. V12 Cdc42 was loaded with either [3H]GDP or [3H]GTP by the nucleotide exchange reaction and added to the proteins immobilized on glutathione beads. Following washing, the resin was subjected to liquid scintillation counting. The binding to each of the beads was expressed as a percentage of input counts, typically around 2.5 × 10^6 cpm for both GDP- and GTP-bound forms.

The identification of so many potential effectors for Cdc42 and Rac is quite surprising. However, numerous activities have already been ascribed to Rac and Cdc42, and each of these activities may require a distinct effector(s). Rac regulates the assembly of focal complexes and the polymerization of actin in lamellipodia, it has an essential role in Ras-induced cellular transformation and can act as an oncogene in its own right (35), and it regulates the activity of cPLA2 (36). Rac and Cdc42 can also activate the N K kinase cascade (11, 12) and in Drosophila both Rac and Cdc42 have been shown to be involved in the extension of neuronal growth cones (37). The identification of so many potential effectors for Cdc42 and Rac GTPases regulate the molecular mechanisms by which Cdc42/Rac GTPases regulate these processes.

Acknowledgments—We thank our colleagues (mentioned under "Materials and Methods") for their kind gifts of cDNAs.

REFERENCES
1. Greenwald, I., and Rubin, G. M. (1992) Cell 68, 271–281
2. Boguski, M. S., and McCormick, F. (1993) Nature 366, 643–654
3. Hall, A. (1994) Annu. Rev. Cell Biol. 10, 31–54
4. Johnson, D., and Pringle, J. R. (1990) J. Cell Biol. 111, 143–152
5. Drubin, D. G. (1991) Cell 65, 1093–1096
6. Chant, J., and Stowers, L. (1995) Cell 81, 1–4
7. Ridley, A. J., and Hall, A. (1992) Cell 70, 389–399
8. Ridley, A. J., Paterson, N. F., Johnston, C. L., Diekmann, D., and Hall, A. (1992) Cell 70, 401–410
9. Nokes, K., and Hall, A. (1995) Cell 81, 53–62
10. Hill, C. S., Wynne, J., and Treisman, R. (1995) Cell 81, 1159–1170
11. Coso, O. A., Chiariello, M., Yu, J. C., Teramoto, H., Crespo, P., Xu, N., Miki, T., and Gutkind, J. S. (1995) Cell 81, 1137–1146
12. Minden, A., Lin, A., Clarot, F. X., Abo, A., and Karin, M. (1995) Cell 81, 1147–1157
13. Lamsa, N., and Hall, A. (1994) Trends Genet. 10, 436–440
14. Settleman, J., Narasimhan, V., Foster, L. C., and Weinberg, R. A. (1992) Cell 69, 539–549
15. Homma, Y., and Emori, Y. (1995) EMBO J. 14, 286–291
16. Reinhard, J., Schmid, A. A., Diekmann, D., Hall, A., Ruppert, C., and Bahler, M. (1996) EMBO J. 15, 697–704
17. Diekmann, D., Abo, A., Johnston, C., Segal, A. W., and Hall, A. (1994) Science 263, 531–533
18. Manser, E., Leung, T., Salihuddin, H., Tan, L., and Lim, L. (1993) Nature 363, 364–367
19. Manser, E., Leung, T., Salihuddin, H., Zhao, Z., and Lim, L. (1994) Nature 367, 40–46
20. Martin, G. S., Bollag, G., McCormick, F., and Abo, A. (1995) EMBO J. 14, 1970–1978
21. Ito, M., Matsu, T., Taniguchi, and Chihara, K. (1994) Gene (Amst.) 139, 215–218
22. Bahou, W. F., Campbell, A. D., and Wicha, M. S. (1992) J. Biol. Chem. 267, 13988–13992
23. Ing, Y. L., Leung, I. W., Heng, H. H. Q., Tsui, L. C., and Lassam, N. J. (1994) Oncogene 9, 1745–1750
24. Katan, M., Kris, R. W., Tidy, N., Philipp, R., Meldrum, E., Aldape, R. A., Knopf, J. L., and Parker, P. (1988) Cell 54, 171–177
25. Sepsenwol, S., Ris, H., and Roberts, T. M. (1989) J. Cell Biol. 108, 55–66
26. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410
27. Pearson, W. R., and Lipman, D. J. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 2444–2448
28. Derry, J. M. J., Ochs, H. D., and Francke, U. (1994) Cell 78, 635–644
29. Lancaster, C. A., Taylor-Harris, P. M., Self, A. J., Brill, S., van Erp, H. E., and Hall, A. (1994) J. Biol. Chem. 269, 1137–1142
30. Darrow, D. S., Deveraux L., Kistchitz, E., and De Kreter, T. (1993) Eur. J. Biochem. 213, 701–710
31. Katoch, M., Hirai, M., Sugimura, T., and Terada, M. (1995) Oncogene 10, 1447–1451
32. Gallo, K. A., Mark, M. R., Scadden, D. T., Wang, Z., Gu, Q., and Godowski, P. J. (1994) J. Biol. Chem. 269, 15092–15100
33. Ezoe, K. E., Lee, S. T., Strunk, K. M., and Spritz, R. A. (1994) Oncogene 9, 935–938
34. Holzman, L. B., Merritt, S. E., and Fan, G. (1994) J. Biol. Chem. 269, 30808–30817
35. Qu, R. G., Chen, J., Kirk, D., McCormick, F., and Symons, F. (1995) Nature 378, 454–459
36. Peppelenbosch, M. P., Oli, R. G., de Vries-Smits, A. M. M., Tertoelen, L. J. G., de Laut, S. W., McCormick, F., Hall, A., Symons, M. H., and Bos, J. L. (1995) Cell 81, 849–856
37. Luo, L., Liao, Y. J., J. A., Oxford, Y. L., and Than, Y. N. (1994) Genes & Dev. 8, 1787–1802
38. Marcus, S., Polverino, A., Chang, E., Robbins, D., Cobb, M. H., and Wigler, M. (1990) Proc. Natl. Acad. Sci. U. S. A. 92, 6180–6184
39. Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, D., Durbin, R., Faveille, A., Fraser, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., J., Johnston, L., Jones, M., Ker, M., Kirsten, J., Lai, K., Latreille, P., Lightning, J., Lloyd, C., Lot, J., O’Callaghan, M., Parsons, J., Percy, C., Rifkin, L., Roper, A., Saunders, D., Shewen, D., Sims, M., Smaldon, N., Smith, A., Smith, M., Sonnhammer, E., Staden, R., Sulston, J., Thiery-Mieg, J., Thomas, K., Vau, M., Vaughan, K., Waters, R., Watson, A., Weinstock, L., Wilkin, S., and Wholdman, P. (1994) Nature 368, 32–38
A Conserved Binding Motif Defines Numerous Candidate Target Proteins for Both Cdc42 and Rac GTPases
Peter D. Burbelo, David Drechsel and Alan Hall

J. Biol. Chem. 1995, 270:29071-29074.
doi: 10.1074/jbc.270.49.29071

Access the most updated version of this article at http://www.jbc.org/content/270/49/29071

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 39 references, 9 of which can be accessed free at http://www.jbc.org/content/270/49/29071.full.html#ref-list-1