J Mol Med (1999) 77:695–698

Digital Object Identifier (DOI)
10.1007/s001099900054

Bioinformatics:
Bits and Bytes

Chris P. Ponting

Raf-like Ras/Rap-binding domains in RGS12- and still-life-like signalling proteins

Received: 25 August 1999 / Accepted: 31 August 1999 / Published online: 2 October 1999

Abstract Ras proteins play critical roles in regulating cell growth and differentiation, and mutated Ras genes are expressed in a variety of human cancers. Consequently, much interest has centered on the binding partners of Ras, including the Ras-binding domain (RBD) of Raf kinase. Here evidence is presented that domains homologous to the Raf RBD are present in tandem in RGS12, RGS14 and LOCO, and singly in molecules similar to mouse Tiam-1. In addition, RGS12, RGS14 and LOCO are shown to contain single "LGN motifs" that are guanine nucleotide exchange factors specific for the α-subunit of G proteins. These findings indicate "cross-talk" interactions between signalling pathways involving Ras and Rap and pathways involving Rho, Rac and Gα GTPases.

Key words Regulator of G protein signalling · LGN motifs · $G_i\alpha$ GTPase · Tiam-1 · Signalling pathway

Abbreviations *GAP* GTPase-activating protein · *GEF* Guanine nucleotide exchange factor · *RBD* Ras-binding domain

C.P. Ponting
National Center
for Biotechnology Information,
National Library of Medicine,
National Institutes of Health,
Bethesda MD 20814, USA
e-mail: Ponting@ncbi.nlm.nih.gov
Tel.: +1-301-4355916
Fax: +1-301-4809241

Present address:
C.P. Ponting
MRC Funtional Genetics Unit,
Department of Human Anatomy
and Genetics, University of Oxford,
South Parks Road,
Oxford OX1 3QX, UK,
e-mail: Ponting@molbiol.ox.ac.uk

Ras and heterotrimeric G proteins' α subunits are GTPases that play critical roles in the initiation of eukaryotic intracellular signalling pathways. These enzymes cycle between inactive GDP-bound forms and active GTP-bound forms. The latter target numerous effector molecules, thereby stimulating the generation of second messenger molecules. The activities of these GTPases are regulated in part by GTPase-activating proteins (GAPs) that stimulate hydrolysis of GTP, and guanine nucleotide exchange factors (GEFs) that stimulate GDP release [1]. Biomedical interest in these pathways stems mainly from the finding of activated mutant *Ras* genes in human tumours [2].

Identification of numerous GAPs, GEFs and effector molecules that bind Ras or Gα GTPases has illuminated many of the pathways that radiate from these prolific signalling molecules. Of particular interest are signalling molecules that interact with more than one GTPase since these mediate "cross-talk" interactions between different signalling pathways. In a previous study [3] we have shown that several GEFs specific for the GTPase Ral p24 contain a Ras^{GTP}-binding "RA domain". These domains were predicted to occur in several distinct signalling molecules contexts. In addition, RA domains were predicted to adopt a ubiquitin-like fold, similar to that known for the Ras-binding domain (RBD) of the Ser/Thr-specific protein kinase Raf1 [4]. Although RA and RBD domains share no significant similarities in sequence, this prediction was borne out by subsequent crystallographic structure determinations [5, 6, 7].

Here I present revised domain assignments for RGS12 and RGS14 (Fig. 1) that are GAPs specific for $G_i\alpha$ subunits [8, 9]. Evidence suggests that these molecules contain tandem domains that are likely to bind, in a GTP-dependent manner, one or more of Ras-like molecules such as K-Ras, Rap1A, Rap2A, R-Ras and TC21. RGS12 and RGS14 contain "regulators of G protein signalling" (RGS) domains that have been shown to specifically stimulate the GTPase activities of Gα subunits, thereby down-regulating G protein coupled receptor-mediated signalling pathways [10, 11, 12]. RGS12 isoforms have been shown to contain an N-terminal PDZ domain [13] that interacts with transmembrane receptors and with the C-terminus of an alternatively spliced RGS12 variant [9].

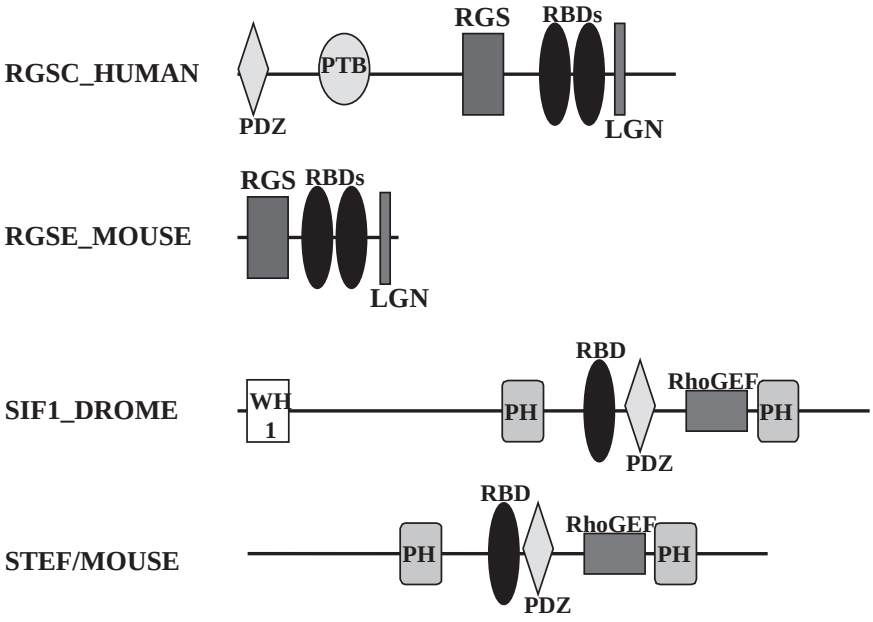

Fig. 1 Schematic representation of the domain architectures of human RGS12 (*RGSC_HUMAN*), mouse RGS14 (*RGSE_MOUSE*), *Drosophila* Still life 1 (*SIF1_DROME*) and mouse STEF (*STEF/MOUSE*), approximately to scale. *LGN* motif in Leu-Gly-Asn repeat containing protein motif [21]; *PDZ* PSD-95, Dlg1, ZO-1/2 domains [13]; *PH* pleckstrin homology domains [27]; *PTB* phosphotyrosine-binding/interaction domains [24]; *RGS* "regulator of G protein signalling" domains [10, 11, 12]; *RhoGEF* Dbl-like GEFs specific for Rho/Rac-type small GTPase domains [28]; *WH1* WASp-homology 1 domains [29]

```
KRAF_HUMAN          NTIRVFLPNKQRTVVNVRNGMSLHDCLMKALKVR-GLQPECCAVFRLLHE-HKGKKARLDWNTDAASLIG-EELQVDFL 4506401   56- 131
KRAA_PIG            GTVKVYLPNKQRTVVTVRDGMSVYDSLDKALKVR-GLNQDCCVVYRLIK----GRKTVTAWDTAIAPLDG-EELIVEVL 3024070   19-  91
RMIL_COTJA          PIVRVFLPNKQRTVVPARCGVTVRDSLKKALMMR-GLIPECCAVYRIQD----GEKKPIGWDTDISWLTG-EELHVEVL  464647  155- 227
KRAF_DROME          ILLRAHLPNQQRTSVEVISGVRLCDALMKALKLR-QLTPDMCEVSTTHS-----GRHIIPWHTDIGTLHV-EEIFVRLL  266434  183- 254
KRAF_CAEEL          KMIMVHLPFDQHSRVEVRPGETARDAISKLLKKR-NITPQLCHVNASSDPKQESIELSLTMEEIASRLPG-NELWVHSE  585373   85- 161
RGSE_HUMAN      1   KYCCVYLPDGTASLALARPGLTIRDMLAGICEKR-GSLYLTSSLPGGN------EQKALVLDQDCTVLAD-QEVRLENR 3914626  149- 219
RGSC_HUMAN      1   KHCCIHLPDGTSCVVAVKAGFSIKDILSGLCERH-GINGAAADLFLVG------GDKPLVLHQDSSILES-RDLRLEKR 3914623  962-1032
LOCO-C1/DROME   1   SLCRVILTDGATTIVQTRPGETVGELVERLLEKRNLVYPYYDIVFQG-------STKSIDVQQPSQILAG-KEVVIERR 4581018  360- 430
RGSE_HUMAN      2   TFELELTALERVVRISAKPTKRLQEALQPILEKH-GLSPLEVVLHRPG------EKQPLDLGKLVSSVAA-QRLVLDTL 3914626  221- 291
RGSC_HUMAN      2   LFRLDLVPINRSVGLKAKPTKPVTEVLRPVVARY-GLDLSGLLVRLSG------EKEPLDLGAPISSLDG-QRVVLEEK 3914623 1034-1104
LOCO-C1/DROME   2   VAFKLDLPDPKVISVKSKPKKQLHEVIRPILSKY-NYKMEQVQVIMRD------TQVPIDLNQPVTMADG-QRLRIVMV 4581018  431- 501
TIAM_MOUSE          TPSWFCLPNNQPALTVVRPGDTARDTLELICKTH-QLDHSAHYLR----------LKFLMENRVQFYIPQPEEDIYELL 2494864  765- 832
STL_DROME           KTFKVAMPDNAYSTVYLRDAMSVEEFLASACARR-NLNPMEHFVR---------VKKRRDMEDHNYFVPH-RNDLIENY 1813378 1121-1188
STEF_MOUSE          VQTYVHFQDNEGVTVTIKPEHRVEDILALACKMR-QLEPTHYGLQLRKV-----VDKSVEWCVPALYEYM-QEQVYDEI 5264503  831- 902
consensus/80%       .h.bhhbs..p.s.l.s+ss.plp-hl..hh..+..l....h.l..........p..lsbpp....l...pc..lc.b
2-structure/1GUA EEEEEE   EEEEEE      HHHHHHHHHH      HHHEEEEEE        EEEE    HHHH    EEEEEE
```

Fig. 2 Multiple alignment of representative Raf1 kinase RBD sequences with homologous domain sequences in RGS12 (*RGSC_HUMAN*), RGS14 (*RGSE_HUMAN*), LOCO, Tiam-1, Still life (*STL*) and STEF. *Beneath the alignment* the known secondary structure of human Raf1 kinase RBD [4]. The alignment is annotated according to an 80% consensus (calculated using http://www.bork.embl-heidelberg.de/Alignment/consensus.html; N. Brown and J. Lai, unpublished). *Gray* hydrophobic (*h*) residues (A, C, F, I, L, M, V, W, Y), big (*b*) residues (E, F, I, K, L, M, Q, R, W, Y) and aliphatic (*l*) residues (I, L, V); *boldface* charged (*c*) residues (D, E, H,K, R), positively charged (+) residues (H, K, R), negatively charged (–) residues (D, E), polar (*p*) residues (D, E, H, K, N, Q, R, S, T) and small (*s*) residues (A, C, S, T, D, N, V, G, P). GenBank identifier (*gi*) accession codes and residue limits are shown following the alignment

Granderath et al. [14] recently described *Drosophila melanogaster* LOCO-c1 and LOCO-c2. These are RGS domain containing proteins that are required for glial cell differentiation. LOCO proteins have been shown to be similar to RGS12 and RGS14 throughout their sequences, including their RGS domains and C-terminal conserved regions B, C and D [14]. Database searches were undertaken to investigate whether regions B, C and D are significantly similar to other proteins in the non-redundant protein sequence database held at the NCBI (ftp://ncbi.nlm.nih.gov/blast/db). These searches employed the position-specific and iterative version of BLAST (PSI-BLAST) [15] and an *E*-value inclusion threshold of 10^{-2}.

A PSI-BLAST search with LOCO-c1 region B (amino acid residues 335–495) revealed, as expected, significant sequence similarity to mammalian RGS12 and RGS14 ($10^{-13} > E > 10^{-17}$) in the first round of searching. By round 2, this search also identified a *second* sub-optimal alignment of the C-terminal half of mouse RGS14 region B (amino acid residues 390–444) with the N-terminal half of fly LOCO region B (amino acid residues 374–428). This implies the presence of tandem repeats within LOCO region B. Although the *E*-value estimate for this second alignment (E=7.6) would be insufficient as evidence for a single repeat, it is strongly suggestive of tandem repeats within the same sequence.

Round 2 of this PSI-BLAST search also suggested that the putative tandem repeats in LOCO, and by extension in RGS12 and RGS14, are homologues of the Ras- and Rap-binding domains (RBDs) in Raf1 kinases (Fig. 1). Part of LOCO-c1 region B was aligned with *Caenorhabditis elegans* lin-45 Raf ki-

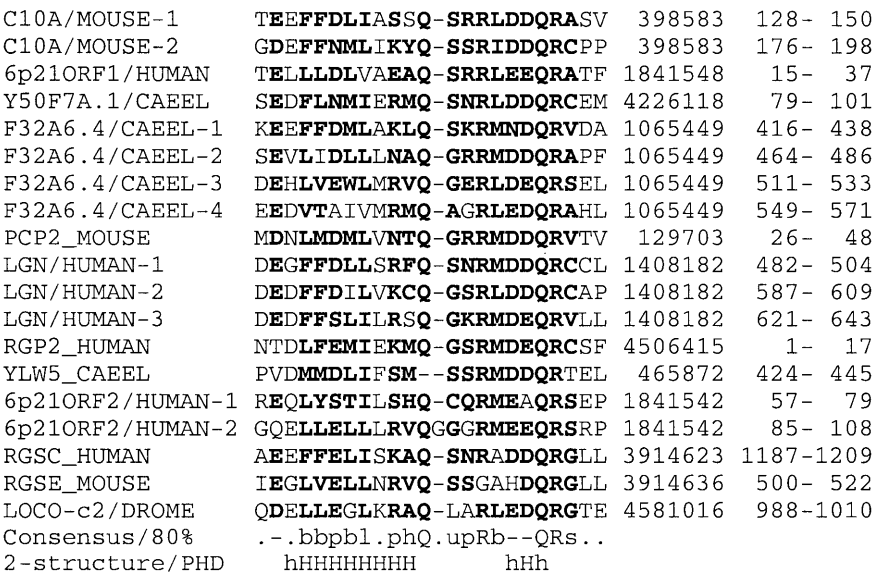

Fig. 3 Multiple alignment of representative LGN sequences, including those in RGS12 (*RGSC_HUMAN*), RGS14 (*RGSE_HUMAN*) and LOCO. The alignment has been annotated in a similar manner to Fig. 1. Predicted [26] secondary structures are shown beneath the alignment. RGP2_HUMAN has been extended at its N-terminus using a human EST sequence (e.g. GenBank identifier 4393914)

nase (amino acid residues 59–191) with an *E*-value of 7×10^{-3}. Support for the prediction of RBD homologous domains in RGS12, RGS14 and LOCO is provided by the finding that mouse RGS14 has already been described as a Rap1/Rap2-interacting protein (Janoueix-Lerosey et al. 1997, unpublished; GenBank identifier 1814396).

A similar search using of mouse RGS14 region B (amino acid residues 264–440) revealed significant similarity, by round 4, to *Drosophila* Still life (type 1; amino acid residues 1100–1158; $E=2\times10^{-3}$) [16] in a region between its PH and PDZ domains. Molecules with similar domain architectures to Still life, namely mouse Tiam-1 [17] and STEF [18], are also likely to contain single RBD homologues (Fig. 2). A PSI-BLAST search with the region intervening between the N-terminal PH and PDZ domains of STEF showed significant similarity ($E=6\times10^{-4}$) to rat RGS12 by round 2.

A multiple alignment (Fig. 2) shows that human Raf1 Arg89 is conserved, or else substituted with positively charged Lys or His residues, in Still life, Tiam1 and N-terminal RGS12-like RBDs. Arg-89 lies at the centre of the Rap1A-Raf1 RBD binding interface [4] and is substituted for leucine in *Drosophila* Raf, resulting in a rough eye phenotype [19]. Many of the other Raf1 RBD residues that interact with Ras-like GTPases [4] are not conserved in the newly identified RBD homologues. This suggests that some RBD homologues may not bind Ras-like GTPases, in a similar manner to RA domain homologues that appear to lack this function [20]. The finding that mouse RGS14 binds Rap1 and Rap2 (Janoueix-Lerosey et al. 1997, unpublished; GenBank identifier 1814396), however, argues that at least one of the two RGS14 RBD homologues does indeed bind Ras-like GTPases.

Region D of RGS12, RGS14 and LOCO (Region D) was shown to be significantly similar to LGN motifs [21]. These are known to be GEFs that are specific for $G\alpha_i$ subunits [22]. A 22 amino acid alignment block of known LGN motifs was used to query current sequences using MoST [23]. Iteration 1 revealed a significantly similar sequence (amino acid residues 1188–1209) in rat RGS12 ($E=4.2\times10^{-6}$) and similar sequences in RGS14 and LOCO (Fig. 3). RGS12 also contains a phosphotyrosine-binding/interaction domain [24] that appears not to have been noted previously, although it is annotated as such by SMART ([25] and unpublished results) and by the SwissProt database (accession code: RGSC_HUMAN).

These newly identified RBD-containing proteins are complex multidomain molecules (Fig. 1). RGS12 and LOCO interact with $G\alpha_i$ subunits [9, 14] and RGS12 acts as a $G\alpha_i$-specific GAP [9]. The finding of LGN motifs in RGS12, RGS14 and LOCO is a surprise since this presumed Gα GEF motif [22] might be thought to antagonise the function of their RGS Gα GAP domains. However, it is suggested that these LGN motifs target Gα subunits other than $G\alpha_i$, such as $G\alpha_s$, $G\alpha_q$ and $G\alpha_{12}$.

The finding of Raf1-like RBDs in proteins thought to be Rho/Rac GEFs (Still life, Tiam1 and STEF) and other proteins harbouring Gα GAP and GEF sequences (RGS12, RGS14 and LOCO) suggests hitherto unforeseen "cross-talk" interactions between Ras and Rap signalling pathways and pathways involving Rho, Rac and Gα GTPases. Further investigations are required to determine whether these RBD homologues bind Ras-like GTPases in a GTP-dependent manner and, if so, their relative specificities for Ras and similar molecules.

Addendum from author LGN motifs have recently been documented as "Goloco" motifs (Siderovski DP, Diversé-Pierluissi MA, de Vries L (1999) The Goloco motif: a $G\alpha_{i/o}$ binding motif and potential guanine-nucleodide exchange factor. Trends Biochem Sci 24:340–341

Acknowledgement CPP is a National Research Council (USA) Senior Associate.

References

1. Boguski MS, McCormick F (1993) Proteins regulating Ras and its relatives. Nature 366:643–654
2. Barbacid M (1987) Ras genes. Annu Rev Biochem 56:779–827
3. Ponting CP, Benjamin DR (1996) A novel family of Ras-binding domains. Trends Biochem Sci 21:422–425
4. Nassar N, Horn G, Herrmann C, Scherer A, McCormick F, Wittinghofer A (1995) The 2.2Å crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1 A and a GTP analogue. Nature 375:554–560
5. Geyer M, Herrmann C, Wohlgemuth S, Wittinghofer A, Kalbitzer HR (1997) Structure of the Ras-binding domain of RalGEF and implications for Ras binding and signalling. Nat Struct Biol 4:694–699
6. Huang L, Hofer F, Martin GS Kim SH (1998) Structural basis for the interaction of Ras with RalGDS. Nat Struct Biol 5:422–426
7. Vetter IR, Linnemann T, Wohlgemuth S, Geyer M, Kalbitzer HR, Herrmann C, Wittinghofer A (1999) Structural and biochemical analysis of Ras-effector signaling via RalGDS. FEBS Lett 451:175–180
8. Snow BE, Antonio L, Suggs S, Gutstein HB, Siderovski DP (1997) Molecular cloning and expression analysis of rat Rgs12 and Rgs14. Biochem Biophys Res Commun 233:770–777
9. Snow BE, Hall RA, Krumins AM, Brothers GM, Bouchard D, Brothers CA, Chung S, Mangion J, Gilman AG, Lefkowitz RJ, Siderovski DP (1998) GTPase activating specificity of RGS12 and binding specificity of an alternatively spliced PDZ (PSD-95/Dlg/ZO-1) domain. J Biol Chem 273:17749–17755
10. De Vries L, Mousli M, Wurmser A, Farquhar MG (1995) GAIP, a protein that specifically interacts with the trimeric G protein G alpha i3, is a member of a protein family with a highly conserved core domain. Proc Natl Acad Sci USA 92:11916–11920
11. Watson N, Linder ME, Druey KM, Kehrl JH, Blumer KJ (1996) RGS family members: GTPase-activating proteins for heterotrimeric G-protein alpha-subunits. Nature 383:172–175
12. Berman DM, Wilkie TM, Gilman AG (1996) GAIP and RGS4 are GTPase-activating proteins for the Gi subfamily of G protein alpha subunits. Cell 86:445–452
13. Ponting CP, Phillips C, Davies KE, Blake DJ (1997) PDZ domains: targeting signalling molecules to sub-membranous sites. Bioessays 19:469–479
14. Granderath S, Stollewerk A, Greig S, Goodman CS, O'Kane CJ, Klämbt C (1999) *loco* encodes an RGS protein required for *Drosophila* glial differentiation. Development 126:1781–1791
15. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402
16. Sone M, Hoshino M, Suzuki E, Kuroda S, Kaibuchi K, Nakagoshi H, Saigo K, Nabeshima Y, Hama C (1997) Still life, a protein in synaptic terminals of *Drosophila* homologous to GDP-GTP exchangers. Science 275:543–547
17. Habets GG, Scholtes EH, Zuydgeest D, van der Kammen RA, Stam JC, Berns A, Collard JG (1994) Identification of an invasion-inducing gene, Tiam-1, that encodes a protein with homology to GDP-GTP exchangers for Rho-like proteins. Cell 77:537–549
18. Hoshino M, Sone M, Fukata M, Kuroda S, Kaibuchi K, Nabeshima Y, Hama C (1999) Identification of the stef gene that encodes a novel guanine nucleotide exchange factor specific for rac1. J Biol Chem 274:17837–17844
19. Melnick MB, Perkins LA, Lee M, Ambrosio L, Perrimon N (1993) Developmental and molecular characterization of mutations in the *Drosophila*-raf serine/threonine protein kinase. Development 118:127–138
20. Kalhammer G, Bahler M, Schmitz F, Jockel J, Block C (1997) Ras-binding domains: predicting function versus folding. FEBS Lett 414:599–602
21. Mochizuki N, Cho G, Wen B, Insel PA (1996) Identification and cDNA cloning of a novel human mosaic protein, LGN, based on interaction with $G_{\alpha i2}$. Gene 181:39–43
22. Luo Y, Denker BM (1999) Interaction of heterotrimeric G protein $G\alpha_o$ with Purkinje cell protein-2. Evidence for a novel nucleotide exchange factor. J Biol Chem 274:10685–10688
23. Tatusov RL, Altschul SF, Koonin EV (1994) Detection of conserved segments in proteins: iterative scanning of sequence databases with alignment blocks. Proc Natl Acad Sci U S A 91:12091–12095
24. Bork P, Margolis B (1995) A phosphotyrosine interaction domain. Cell 80:693–694
25. Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci USA 95:5874–5864
26. Rost B, Sander C (1993) Prediction of protein secondary structure at better than 70% accuracy. J Mol Biol 232:584–599
27. Gibson TJ, Hyvönen M, Musacchio A, Saraste M, Birney E (1994) PH domain: the first anniversary. Trends Biochem Sci 19:349–353
28. Cerione RA, Zheng Y (1996) The Dbl family of oncogenes. Curr Opin Cell Biol 8:216–222
29. Callebaut I, Cossart P, Dehoux P (1998) EVH1/WH1 domains of VASP and WASP proteins belong to a large family including Ran-binding domains of the RanBP1 family. FEBS Lett 441:181–185