

The GoLoco motif: a $G\alpha_{1/0}$ binding motif and potential guanine-nucleotide-exchange factor

Studies of the desensitization of G-protein-coupled signal transduction have led to the discovery of a family of GTPase-activating proteins (GAPs) for heterotrimeric G-protein α subunits – the ‘regulator of G-protein signaling’ or RGS proteins^{1,2}. RGS12 is the largest mammalian RGS protein that is currently known² and has a multi-domain structure reminiscent of both adaptor and scaffold proteins (Fig. 1a). It consists of an N-terminal PDZ (PSD-95/Dlg/ZO-1) domain, which binds the C-termini of both RGS12 itself and the G-protein-coupled interleukin-8 receptor B or CXCR2 (Ref. 3), a phosphotyrosine-binding (PTB) domain (D. P. Siderovski, L. De Vries, and M. A. Diversé-Pierluissi, unpublished), and a central RGS domain, which is a GAP for $G_{1/0}$ -class heterotrimeric G protein α subunits³. Here, we show that the C-terminus of RGS12 contains a new 19-amino-acid motif, also present in other proteins that bind $G_{1/0}$ -class α subunits (Fig. 1b). As this motif was initially discovered by comparison of mammalian RGS proteins with *Loco*, the *Drosophila* RGS12 homologue, we have named it the $G\alpha_{1/0}$ -*Loco* or ‘GoLoco’ motif.

loco was identified in an enhancer-trap screen for genes whose expression in glial cells depends on the activity of Pointed, a member of the ets family of winged helix–turn–helix transcription factors⁴. In the same report, Granderath and colleagues performed a yeast two-hybrid screen employing *Drosophila* $G\alpha_i$ as ‘bait’ and identified four overlapping *loco* cDNAs as interacting clones: three lacking the presumptive $G\alpha$ -interacting domain (i.e. the RGS domain), but all containing the final 43 amino acids of a C-terminal region with 39% identity to RGS12. An iterative PSI-BLAST search⁵, using the corresponding 51-amino-acid region of rat RGS12 (aa 1171–1221 of SP:RGSC_RAT) and an *E*-value threshold of 1 (required to identify *Loco* in the first iteration; *E* = 4×10^{-1}), identified similar regions within other known $G\alpha$ -interacting proteins: RGS14 (*E* = 4×10^{-6} ; iteration 0), the human mosaic protein LGN (*E* = 3×10^{-1} ; iteration 0) and Purkinje-cell protein-2 (Pcp2; *E* = 8×10^{-1} ; iteration 1).

RGS12 and RGS14 are closely related RGS proteins and we previously reported a 17-amino-acid sequence of unknown function shared between RGS12 and RGS14 (Ref. 6). This 17-amino-acid region

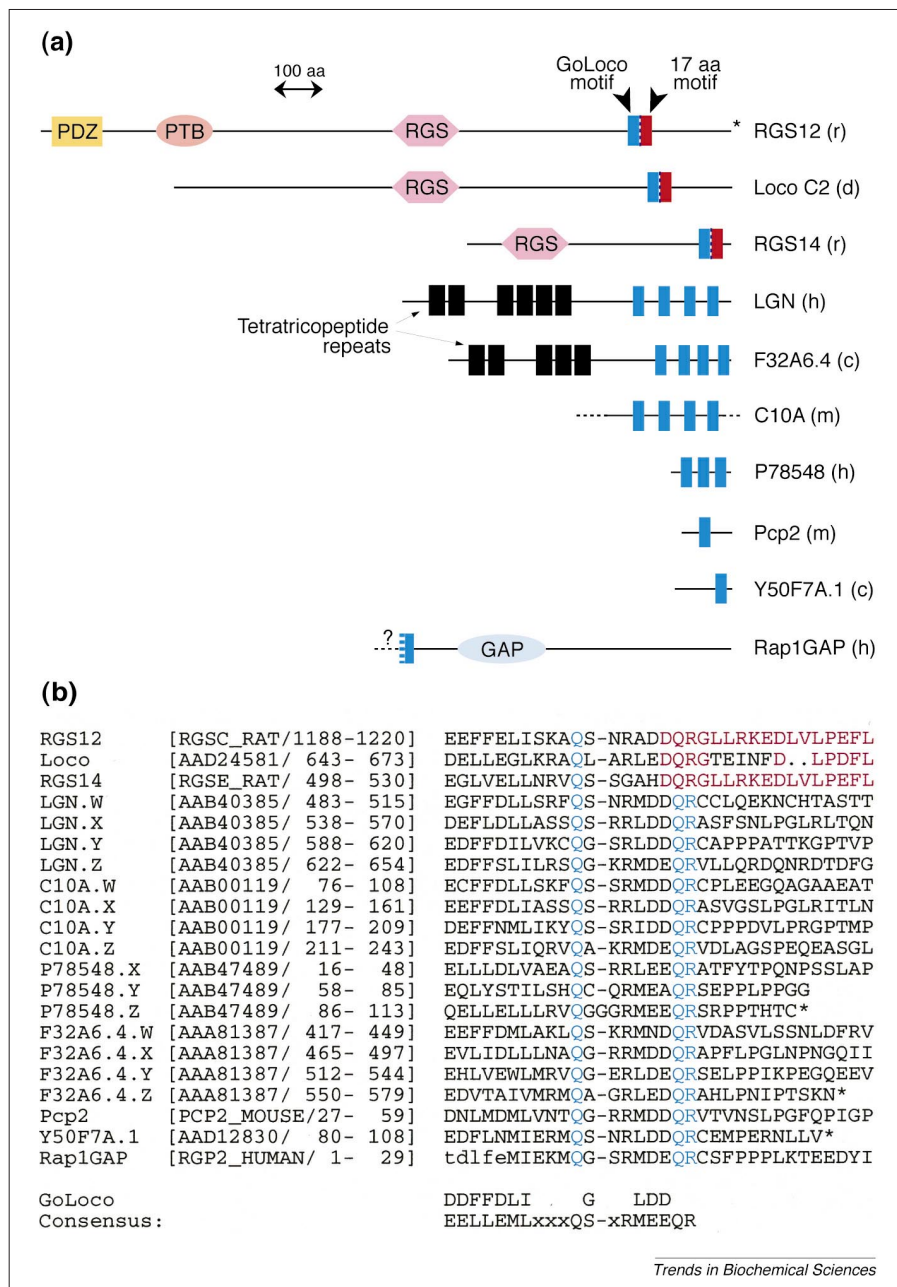


Figure 1

(a) Domain organization of proteins that contain one or more GoLoco motifs. Asterisk denotes C-terminal residues Ala-Thr-Phe-Val that bind the N-terminal PDZ (PSD-95/Dlg/ZO-1) domain of RGS12 (Ref. 3). *Loco* C2 represents the open-reading frame of the longest of two *loco* transcripts described by Granderath *et al.*⁴; while *loco* is presumed to be the *Drosophila* homologue of rat RGS12, neither of the two reported *loco* transcripts encode PDZ or phosphotyrosine-binding (PTB) domains. Dotted lines denote uncharacterized open-reading frames; the dotted line with a question mark represents translation of the Rap1GAP open-reading frame 5' of the reported start codon. Species of origin is in brackets: c, *C. elegans*; d, *Drosophila*; h, human; m, mouse; r, rat. (b) An alignment of GoLoco motifs. For each representative protein, SwissProt ID (or GenBank accession number) and location of sequence range within protein is shown in square brackets. Invariant residues are shown in blue; a 17-amino-acid motif, previously reported to be conserved between RGS12 and RGS14 (Ref. 6), is highlighted in red. Lower-case amino acids are derived from translation 5' of the reported start codon. Asterisks denote stop codons. A multiple sequence alignment of the 19-amino-acid GoLoco motif (alignment number DS38614) has been deposited with the European Bioinformatics Institute (<ftp://ftp.ebi.ac.uk/pub/databases/embl/align/>). Abbreviations: GAP, GTPase-activating protein; Pcp, Purkinje-cell protein-2; RGS, regulator of G-protein signaling.

includes the final three amino-acid residues of the GoLoco motif and might reflect another functional domain

(Fig. 1b). The human mosaic protein LGN is a $G\alpha_{12}$ -interacting protein of unknown function; the original report⁷ on LGN

foreshadowed the presence of GoLoco motifs by describing four 34-amino-acid repeats (W, X, Y and Z) in a C-terminal region distinct from six tetratricopeptide (TP) repeats⁸ that can mediate self-assembly into multiprotein complexes⁹. More recently, Luo and Denker¹⁰ found that both LGN and Pcp2 interact with $G\alpha_o$ in a yeast two-hybrid screen. Pcp2 was also shown to bind *in vitro* translated $G\alpha_o$ and $G\alpha_{12}$, but not $G\alpha_s$. Moreover, Pcp2 was reported to act as a novel guanine-nucleotide exchange factor (GEF), stimulating the release of GDP from $G\alpha_o$ (Ref. 10). This latter result raises the possibility that the GoLoco motif might also specify G-protein-coupled receptor-independent GEF activity for $G\alpha_{1/0}$ subunits.

While the PSI-BLAST search with rat RGS12 failed to assign a significant *E*-value to Rap1GAP (SP:RGP2_HUMAN), a GAP for the small G protein Rap1, we and Mochizuki *et al.*⁷ noted that half the GoLoco motif (QGxRMDEQR) is present in the N-terminus of Rap1GAP. Translation upstream of the reported ATG start codon of *rap1GAP* (GB:M64788) reveals no in-frame stop codon but, surprisingly, a similarity to the N-terminal half of the GoLoco motif (Fig. 1b). The sequence adjacent to the reported start codon (CTATTTGAGATGA) does not conform to the Kozak consensus sequence (GCCGCC[A/G]CCATGG; Ref. 11). The presence of a GoLoco motif within Rap1GAP presents the tantalizing prospect that this one protein might serve as a GEF for $G\alpha_{1/0}$ subunits and a GAP for the small G protein Rap1 – mirroring the functionality of p115RhoGEF and PDZRhoGEF, two RGS proteins each able to serve as a GAP for $G\alpha_{12/13}$ subunits and a GEF for the small G protein RhoA (Refs 12–14).

Two other proteins were also found to contain GoLoco motifs in the PSI-BLAST search using rat RGS12: one motif within Y50F7A.1, an anonymous *Caenorhabditis elegans* open-reading frame (GB:AAD12830; *E* = 4×10^{-1} ; iteration 0), and four motifs within the partial open-reading frame of the mouse *C10A* gene (GB:L23316; *E* = 5×10^{-1} ; iteration 0). An

anonymous human open-reading frame, P78548 (GB:AAB47489; *E* = 2×10^{-4} ; iteration 1), and its mouse orthologue NG1 (GB:AF03001; *E* = 4×10^{-4} ; iteration 1), each containing three GoLoco motifs, were identified in a pattern-hit initiated BLAST search (PHI-BLAST; Ref. 15) using full-length human LGN as the protein query sequence and the GoLoco motif consensus ([D/E][D/E][F/L][F/L][D/E][L/M][I/L]xxxQ[G/S]xR[L/M][D/E][D/E]QR) as the search pattern. One additional *C. elegans* open-reading frame, F32A6.4, with an overall 49% similarity to human LGN (BLAST *E*-value 5×10^{-75} for F32A6.4 versus LGN), was found to possess four GoLoco motifs (Fig. 1b). A reciprocal BLAST search using the 99-amino-acid open-reading frame reported for mouse Pcp2 (SP:PCP2_MOUSE) identified the GoLoco motif-containing proteins C10A (*E* = 3×10^{-5}), LGN (*E* = 3×10^{-3}) and F32A6.4 (*E* = 3×10^{-2}). An additional reciprocal BLAST search using the 33-amino-acid GoLoco motif region of Rap1GAP that includes predicted N-terminal amino acids (Fig. 1b) identified the proteins LGN (*E* = 3×10^{-2}), Y50F7A.1 (*E* = 2×10^{-1}), C10A (*E* = 2) and F32A6.4 (*E* = 6).

Although the ultimate function of this novel $G\alpha$ -binding motif is currently unknown, the presence of tandemly repeated GoLoco motifs in LGN, C10A, P78548/NG1 and F32A6.4 raises the possibility that these proteins aggregate GDP-bound $G\alpha$ subunits and facilitate their simultaneous activation by receptor-independent GEF activity – a possibility that would support the findings made by the late Martin Rodbell of dynamic $G\alpha$ aggregates (Ref. 16 and references therein).

Acknowledgements

We thank Harry Elsholtz and Vuk Stambolic (University of Toronto) for pointing out the work of Granderath *et al.*, David Fenstermacher (UNC-CH) for computational assistance, J. Dedrick Jordan (Mount Sinai) for insightful discussions, and reviewers' comments for refinements to our database analyses. This work was supported in part by

National Institutes of Health grant AA11605 (to D. P. S.) and by Mount Sinai Dean's Fund for Innovative Research, National Institutes of Health grant NS37443, and The NYC Speaker's Fund for Biomedical Research Award (to M. D-P.).

References

- Siderovski, D. P. *et al.* (1996) *Curr. Biol.* 6, 211–212
- De Vries, L. and Farquhar, M. G. (1999) *Trends Cell Biol.* 9, 138–144
- Snow, B. E. *et al.* (1998) *J. Biol. Chem.* 273, 17749–17755
- Granderath, S. *et al.* (1999) *Development* 126, 1781–1791
- Altschul, S. F. *et al.* (1997) *Nucleic Acids Res.* 25, 3389–3402
- Snow, B. E. *et al.* (1997) *Biochem. Biophys. Res. Commun.* 233, 770–777
- Mochizuki, N., Cho, G., Wen, B. and Insel, P. A. (1996) *Gene* 181, 39–43
- Lamb, J. R., Tugendreich, S. and Hieter, P. (1995) *Trends Biochem. Sci.* 20, 257–259
- Das, A. K., Cohen, P. W. and Barford, D. (1998) *EMBO J.* 17, 1192–1199
- Luo, Y. and Denker, B. M. (1999) *J. Biol. Chem.* 274, 10685–10688
- Kozak, M. (1996) *Mamm. Genome* 7, 563–574
- Kozasa, T. *et al.* (1998) *Science* 280, 2109–2111
- Hart, M. J. *et al.* (1998) *Science* 280, 2112–2114
- Fukuhara, S. *et al.* (1999) *J. Biol. Chem.* 274, 5868–5879
- Zhang, Z. *et al.* (1998) *Nucleic Acids Res.* 26, 3986–3990
- Jahangeer, S. and Rodbell, M. (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 8782–8786

DAVID P. SIDEROVSKI

Department of Pharmacology, UNC-Chapel Hill School of Medicine, CB#7365, Room 906B M.E. Jones Bldg, Chapel Hill, NC 27599-7365, USA.
Email: dsiderov@med.unc.edu

MARÍA A. DIVERSÉ-PIERLUISSI

Department of Pharmacology, Mount Sinai School of Medicine, New York, NY 10029, USA.

LUC DE VRIES

Division of Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA 92093, USA.

Letters to TIBS

TIBS welcomes letters on any topic of interest. Please note, however, that previously unpublished data and criticisms of work published elsewhere cannot be accepted by this journal.

Letters should be sent to:

Mary Purton, Editor, *Trends in Biochemical Sciences*
Elsevier Trends Journals
68 Hills Road, Cambridge, UK CB2 1LA