A modulator of rho family G proteins, rhoGDI, binds these G proteins via an immunoglobulin-like domain and a flexible N-terminal arm

Nicholas H Keep[1†], Maria Barnes[2,3], Igor Barsukov[2,3], Ramin Badii[2,3], Lu-Yun Lian[2,3], Anthony W Segal[1], Peter CE Moody[2] and Gordon CK Roberts[2,3*]

Background: The rho family of small G proteins, including rho, rac and cdc42, are involved in many cellular processes, including cell transformation by ras and the organization of the actin cytoskeleton. Additionally, rac has a role in the regulation of phagocyte NADPH oxidase. Guanine nucleotide dissociation inhibitors (GDIs) of the rhoGDI family bind to these G proteins and regulate their activity by preventing nucleotide dissociation and by controlling their interaction with membranes.

Results: We report the structure of rhoGDI, determined by a combination of X-ray crystallography and NMR spectroscopy. NMR spectroscopy and selective proteolysis show that the N-terminal 50–60 residues of rhoGDI are flexible and unstructured in solution. The 2.5 Å crystal structure of the folded core of rhoGDI, comprising residues 59–204, shows it to have an immunoglobulin-like fold, with an unprecedented insertion of two short β strands and a 3_{10} helix. There is an unusual pocket between the β sheets of the immunoglobulin fold which may bind the C-terminal isoprenyl group of rac. NMR spectroscopy shows that the N-terminal arm is necessary for binding rac, although it remains largely flexible even in the complex.

Conclusions: The rhoGDI structure is notable for the existence of both a structured and a highly flexible domain, both of which appear to be required for the interaction with rac. The immunoglobulin-like fold of the structured domain is unusual for a cytoplasmic protein. The presence of equivalent cleavage sites in rhoGDI and the closely related D4/Ly-GDI (rhoGDI-2) suggest that proteolytic cleavage between the flexible and structured regions of rhoGDI may have a role in the regulation of the activity of members of this family. There is no detectable similarity between the structure of rhoGDI and the recently reported structure of rabGDI, which performs the same function as rhoGDI for the rab family of small G proteins.

Addresses: [1]Department of Medicine, University College London, Rayne Institute, 5 University Street, London WC1E 6JJ, UK, [2]Department of Biochemistry, University of Leicester, University Road, Leicester LE1 7RH, UK and [3]Biological NMR Centre, University of Leicester, University Road, Leicester LE1 7RH, UK.

[†]Present address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

*Corresponding author.
E-mail: gcr@le.ac.uk

Key Words: dissociation inhibitor, G proteins, GTPase, rac, rhoGDI

Received: 29 January 1997
Revisions requested: 26 February 1997
Revisions received: 21 March 1997
Accepted: 2 April 1997

Structure 15 May 1997, 5:623–633
http://biomednet.com/elecref/0969212600500623

Introduction

The rho family of ras-related G proteins (GTPases) includes the isoforms of rho, rac, cdc42 and TC10 [1]. The proteins rac, rho and cdc42 are all involved in the reorganization of the actin cytoskeleton in response to external stimuli [2–4]. These proteins also have roles in cell transformation by ras, in cytokinesis, in focal adhesion formation and in the stimulation of stress-activated kinase [4–7]. The immediate downstream target of these G proteins is usually a serine/threonine kinase [7]. Additionally, rac has a specific role in the activation of the phagocytic NADPH oxidase [8,9].

Like ras, members of the rho family of GTPases act as molecular switches by virtue of a conformational change between the GTP- and GDP-bound forms. The inherent rates of GTP hydrolysis and nucleotide dissociation are slow, and proteins have been described that accelerate hydrolysis (GTPase-activating proteins; GAPs) and catalyze exchange of GTP for GDP (guanine nucleotide exchange factors; GEFs) [1,4]. There are also two known groups of guanine nucleotide dissociation inhibitors (GDIs) that inhibit the release of guanine nucleotides from members of the rab family (rabGDI and its homologues) and the rho family (rhoGDI and its homologues). There is no significant sequence homology between proteins of the two groups of GDIs, and these proteins are of very different sizes — rabGDIs are proteins of at least 450 residues, whereas rhoGDIs are only about 200 residues [1]. Three human homologues of rhoGDI have been reported. The form studied here, rhoGDI (rhoGDI-1) [10], is found in a broad range of cells and is active towards

all forms of rho, rac and cdc42. D4-GDI or Ly-GDI (rhoGDI-2) [11,12] is particularly abundant in haematopoetic cells and also has a broad range of activity toward the rho family of proteins. The most recently identified of these proteins, rhoGDI-3 [13], is specifically active towards rhoB and rhoG and, unlike the other two forms which are cytoplasmic, it is associated with the membrane or possibly the cytoskeleton [13].

Most small G proteins undergo post-translational modification by transfer of farnesyl or geranylgeranyl groups to cysteines at or near their C termini. In the case of the rho family, the enzyme geranylgeranyl transferase I recognizes the CXXL sequence and transfers a geranylgeranyl group to the cysteine residue. The G protein is further modified by cleavage of the XXL sequence and the formation of a methyl ester at the now isoprenylated C-terminal cysteine. The rab family of G proteins contains the sequence CC or CXC at the C terminus, and both cysteines are modified by addition of a geranylgeranyl group by geranylgeranyl transferase II [14,15]. Interaction of both rhoGDI [16] and rabGDI [17] with their cognate G proteins is dependent on the geranylgeranylation of the C-terminal cysteine(s) of the G protein. In the cell, many of the targets of the G proteins are in the membrane and G protein isoprenyl modification appears to mediate membrane binding. Cytosolic G proteins of the rho and rab families normally exist in the cell bound to their cognate GDIs, and the GDIs can actively remove their G proteins from the membrane [18–20]. This role of GDI in partitioning of the G proteins between membrane and cytosol may be physiologically more important than the inhibition of nucleotide dissociation.

In this paper, we describe the structure of rhoGDI. This protein is notable for the existence of both a structured and a highly flexible domain; the structured domain comprises an immunoglobulin-like fold, which is unusual for a cytoplasmic protein. There is no detectable similarity between the structure of rhoGDI and the recently reported structure of rabGDI, which performs the same function for the rab family of small G proteins. Both the structured and flexible domains of rhoGDI appear to be required for its interaction with rac, and comparison with the closely related D4/Ly-GDI (rhoGDI-2) suggests that proteolytic cleavage between the flexible and structured regions may have a role in the regulation of the activity of members of the rhoGDI family.

Results and discussion

RhoGDI consists of a flexible N-terminal arm and an ordered domain

The ^{15}N–^{1}H HSQC spectrum of rhoGDI (Fig. 1a) shows that the cross-peaks can be divided into two groups, according to their linewidths. Between 50 and 60 of the backbone amide cross-peaks (predominantly in the region ^{1}H 7.7–8.6 ppm, ^{15}N 115–129 ppm) are much narrower than the remainder, as are many of the sidechain amide cross-peaks (^{1}H 6.7–7.7 ppm, ^{15}N 111–114 ppm). These narrow peaks can be selectively observed by plotting the spectrum with a higher threshold (Fig. 1b). These sharp resonances must arise from residues that have substantial local mobility. Selective proteolysis of rhoGDI with trypsin gave a stable protease-resistant domain of about 17 kDa, identified by N-terminal sequencing as starting at residue 59 of rhoGDI. The ^{15}N–^{1}H HSQC spectrum of this fragment (Fig. 1c) lacks most of the sharp cross-peaks seen with the full-length protein; the remaining cross-peaks appear in the same positions as in the intact protein, indicating that the structure of the 17 kDa fragment is essentially the same as that of the corresponding portion of the intact molecule. This demonstrates that rhoGDI has two structural domains, a flexible domain comprising the N-terminal ~60 residues of the protein and an ordered or folded domain comprising the remaining ~140 residues.

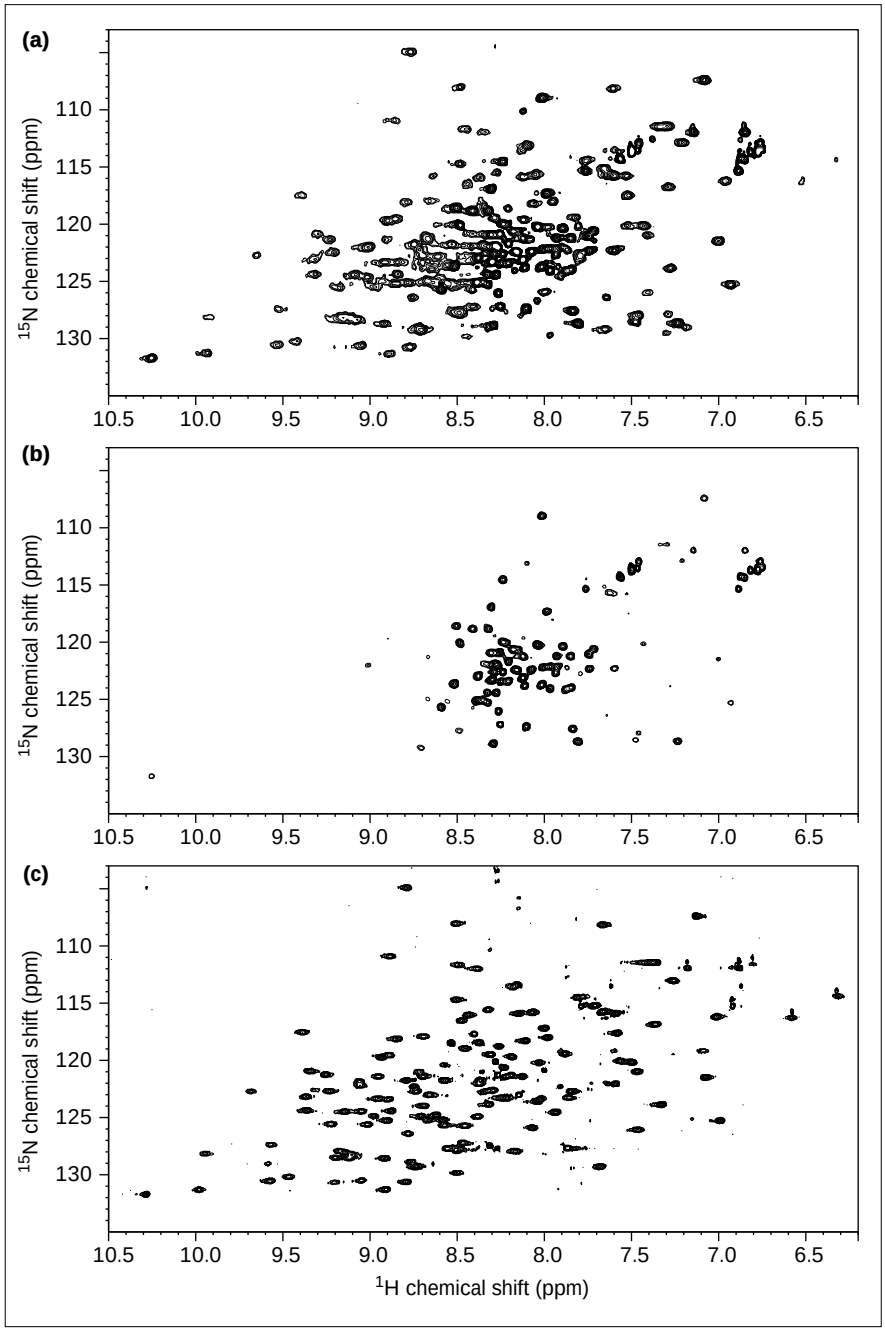
Crystal structure of the folded domain of rhoGDI

The folded domain produced by proteolysis of rhoGDI with trypsin (comprising residues 59–204 of the intact protein) was crystallized from ammonium sulphate, and the structure was solved to a resolution of 2.5 Å by multiple isomorphous replacement with anomalous scattering (see Materials and methods; Table 1). A portion of the electron-density map of the rhoGDI structure is shown in Figure 2. The asymmetric unit contains three copies (A, B, C) of the rhoGDI molecule; these each have essentially identical structures except for the first nine amino acids (Fig. 3a). The rms differences (calculated using LSQMAN [21]) between all atoms in residues 68–203 are: B to A, 0.68 Å; C to A, 0.71 Å; C to B, 0.58 Å. The first nine residues are in different conformations in molecules A and B (Fig. 3a; the rms difference for residues 59–67 is 12.35 Å when the molecules are superposed using residues 68–203) and are not ordered in molecule C. The model contains all residues from Val59 of rhoGDI, except for residues 59–67 in molecule C and for the C-terminal residue (Asp204) in all three molecules, which appears to be disordered. There are also 191 waters and 9 sulphates.

The structure of rhoGDI consists of nine β strands in two antiparallel sheets, forming a β sandwich, and one two-turn 3_{10} helix (Figs 3a,b). In the orientation shown in Figure 3, the top sheet consists of strands β2 (residues 87–89), β1 (70–78), β4 (111–118) and β7 (156–159), and the bottom sheet of strands β3 (102–105), β9 (190–199), β8 (173–183), β5 (125–134) and β6 (137–144). The short 3_{10} helix is packed against the edge of the bottom β sheet. The structure of rhoGDI was compared with known protein structures using the program DALI [22]. The 52 proteins that gave statistically significant structural alignments nearly all had immunoglobulin-like folds; these consist of a core of seven β strands in two sheets forming

Figure 1

^{15}N–^{1}H HSQC spectra of intact and truncated rhoGDI. (a) Spectrum of intact rhoGDI. Between 50 and 60 of the backbone amide cross-peaks and many of the sidechain amide cross-peaks have linewidths much narrower than the remainder. (b) The same spectrum plotted with a higher threshold, and thus showing only the sharp cross-peaks. (c) Spectrum of truncated rhoGDI containing residues 59–204, showing the loss of most of the sharp cross-peaks, while the remainder of the cross-peaks appear in the same positions.



an antiparallel β sandwich with a Greek key topology [23]. In rhoGDI, the β strands that correspond to those in the immunoglobulin fold are β1, β4, and β7 in the first sheet and β9, β8, β5 and β6 in the second, shown in green and red, respectively, in Figure 3b. The topology of the immunoglobulin-like core of rhoGDI corresponds to the s-type of immunoglobulin-like folds as defined by Bork *et al*. [24]; this is a variation of the classical constant domain fold in which the d strand is swapped to the other sheet and renamed c′. Proteins that also show this topology include fibronectin repeats, CD4 domains 2 and 4, CD2 domains and two domains of *Escherichia coli* β-galactosidase; these were among the most significant matches given by DALI. The classical immunoglobulin fold has a disulphide bridge between strands b and f (β4 and β7 in the case of rhoGDI). As expected for a cytosolic protein, there is no disulphide bridge in rhoGDI. Some known s-type immunoglobulin domains do have the b–f strand

Table 1

Phasing and refinement statistics for rhoGDI.

	Data Set		
	Native	Mercury	Selenium
Wavelength (Å)	0.977	0.832	0.977
Resolution (Å)	2.7	3.2	2.5
Observations	96 325	66 276	16 1161
Unique observations	17 218	9729	20 345
Completeness (%)	99.7	92.8	92.9
Redundancy	5.6	6.8	7.9
I/σ	22.5	10.2	25.5
R_{sym}*	0.081	0.094	0.071
R_{anom}†		0.066	0.036
Number of sites		3	9
R_{cullis}‡		0.69	0.89
$R_{cullis\ (anom)}$‡		0.78	0.94
Phasing power§		1.72	0.79
Highest resolution shells			
Resolution range	2.8–2.7	3.31–3.2	2.59–2.5
Completeness (%)	100	95.6	95.3
Median redundancy	5	7	11
R_{sym}*	33.6	18.1	26.4
I/σ	6.6	10.3	8.7
Refinement statistics¶			
Resolution (Å)	8–2.7		8–2.5
Number of atoms	3663		3663
R_{cryst} (%)#	24.0		22.5
R_{free} (%)#	31.0		29.8
Rms bond lengths (Å)			0.012
Rms bond angles (°)			2.0

Cell parameters are a = b = 156.57, c = 132.50, in space group R32 (155) hexagonal setting. *$R_{sym} = \Sigma\Sigma|\langle I\rangle - I_j|/\Sigma\langle I\rangle$, where I_j is the observed intensity and $\langle I\rangle$ is the mean intensity. †$R_{anom} = \Sigma(|\langle I\rangle - \langle I+\rangle| + |\langle I\rangle - \langle I-\rangle|)/\Sigma\langle I\rangle$. ‡$R_{cullis}$ = (lack of closure)/(isomorphous difference); $R_{cullis\ (anom)}$ = (lack of closure)/(anomalous difference) for acentric terms. §Phasing power $= \Sigma F_h^c / \Sigma[|F_p^o \exp(i\theta_p) + F_h^c \exp(i\theta_c)| - F_{ph}^o|$, where F_h^c and θ_c are the calculated heavy-atom amplitude and phase, F_p^o and F_{ph}^o are the observed amplitudes for the protein and the heavy-atom derivative and θ_p is the MIR phase and the sum is over all observations. ¶Refinement was carried out using the selenium data set; statistics for the native data set are the result of using the native data set and changing the seleniums to sulphurs, not the result of an independent refinement. #$R_x = \Sigma|F_p^o - F_p^c|$; for x = cryst the sum is over the 95% of data used for refinement, for x = free this sum is over the remaining 5% of data.

disulphide bond (e.g. CD4 domain 2), whereas others have an unusual a–g strand disulphide bond (e.g. neuroglian domain 2), and some have both (CD2 domain 2). S-type domains that, like rhoGDI, lack a disulphide bond include growth hormone receptor domain 2 and fibronectin [24]. There is only one cysteine in the truncated rhoGDI, which lies at the end of strand a. The structure of rhoGDI contains an unprecedented insertion (residues 80–110 shown in blue in Fig. 3b), replacing a hairpin loop between strands a and b (equivalent to β1 and β4 in rhoGDI) in the classical immunoglobulin fold. This insertion consists of two short β strands (β2 and β3) and a 3_{10} helix; each of these β strands extends one of the two sheets.

Figure 3c shows the structure of the ordered domain coloured by the average mainchain B factors. (The sidechain and overall B values follow a similar pattern, but with more variation due to surface sidechains.) Most of the β strands have low B values, including the two short inserted strands β2 and β3. Strand 6 is somewhat more mobile, as is the 3_{10} helix; both these elements lie on the edge of the structure. The N and C termini are highly mobile, and there is a particularly disordered, charged loop from Asp183 to Lys186.

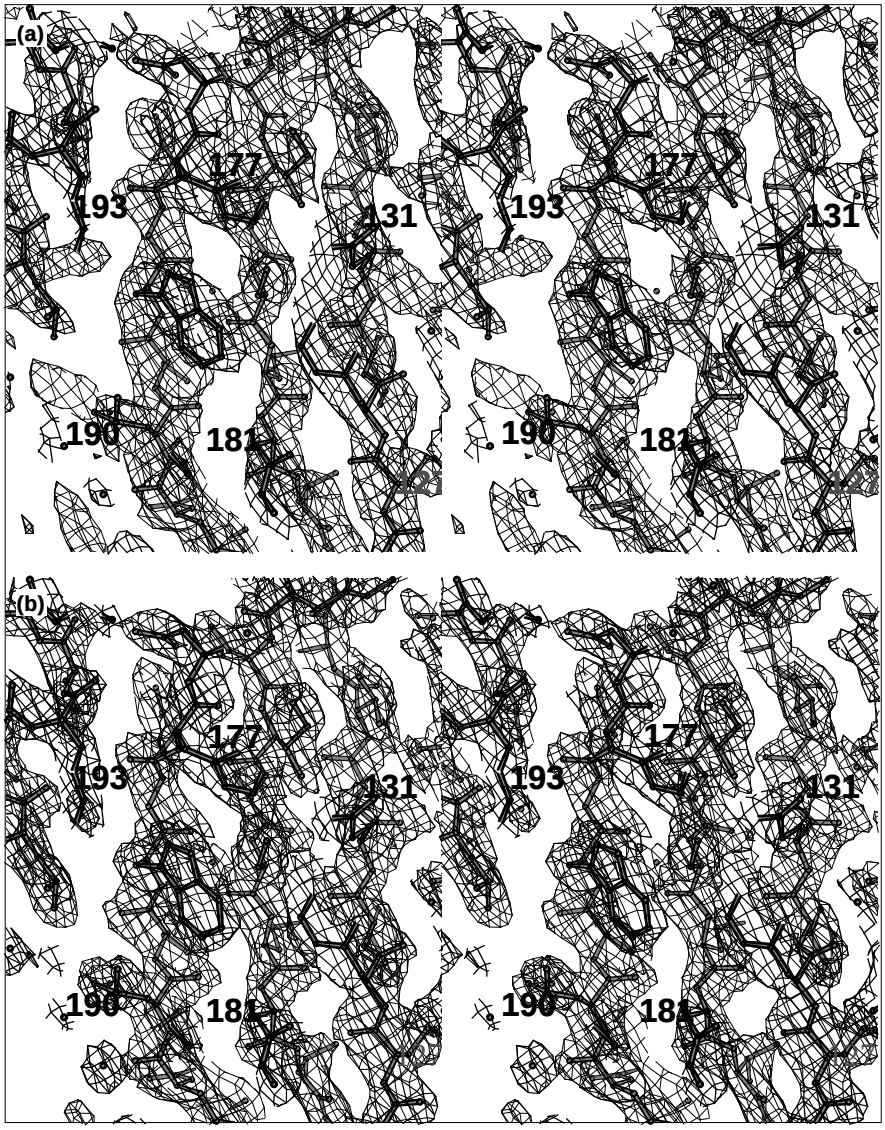
Figure 4 indicates the buried residues that form the hydrophobic core. Mutagenesis studies in which the C terminus of rhoGDI was progressively truncated have shown that one or more of the residues between 197 and 200, but not between 201 and 204, are essential for rhoGDI binding to cdc42 (Fig. 5; [25]). This can be understood from the structure of rhoGDI, which shows that Ile198, in β strand 9, forms part of the hydrophobic core of the folded domain; in the truncated rhoGDI (residues 1–197), the structured domain is thus unlikely to be fully folded, explaining its inability to bind cdc42. Residues in the insertion (residues 80–110) clearly form part of the hydrophobic core. In particular, the two tryptophans 192 and 194 in β strand 9 make contact with several of the buried residues in the insertion (Leu88, Phe97 and Phe102). It is clear from the B factors and from consideration of the residues involved in the hydrophobic core that the unusual insertion into the immunoglobulin fold is an integral and stable part of the structure of rhoGDI.

There is no detectable similarity between the structure of rhoGDI reported here and the recently reported structure of rabGDI [26], which performs the same function for G proteins of the rab family. The latter is a much larger protein of 450 residues, which has structural similarities to FAD-dependent flavoproteins, although there is no known role for FAD in its function.

A pocket in the rhoGDI molecule may bind the C-terminal isoprenyl group of G proteins

Inspection of the accessible surface of the folded domain of rhoGDI (Fig. 5a) reveals a large cavity (12×8×10 Å) between the two β sheets of the immunoglobulin fold. Although this pocket is largely hydrophobic, it contains a region of negative potential. The T and B cell homologue of rhoGDI, rhoGDI-2 (D4/Ly-GDI), has a sequence very similar to that of rhoGDI-1 (Fig. 4), but it has a 10–20-fold lower affinity than rhoGDI-1 for isoprenylated cdc42. A single mutation in rhoGDI-2, Asn177→Ile (rhoGDI numbering), substantially explains this difference in affinity [25]. Residue 177 lies on the surface of the protein, in the pocket (Fig. 5b), and the effect of substitution at this position on the affinity for cdc42 indicates that the pocket plays an important role in G-protein binding.

Figure 2



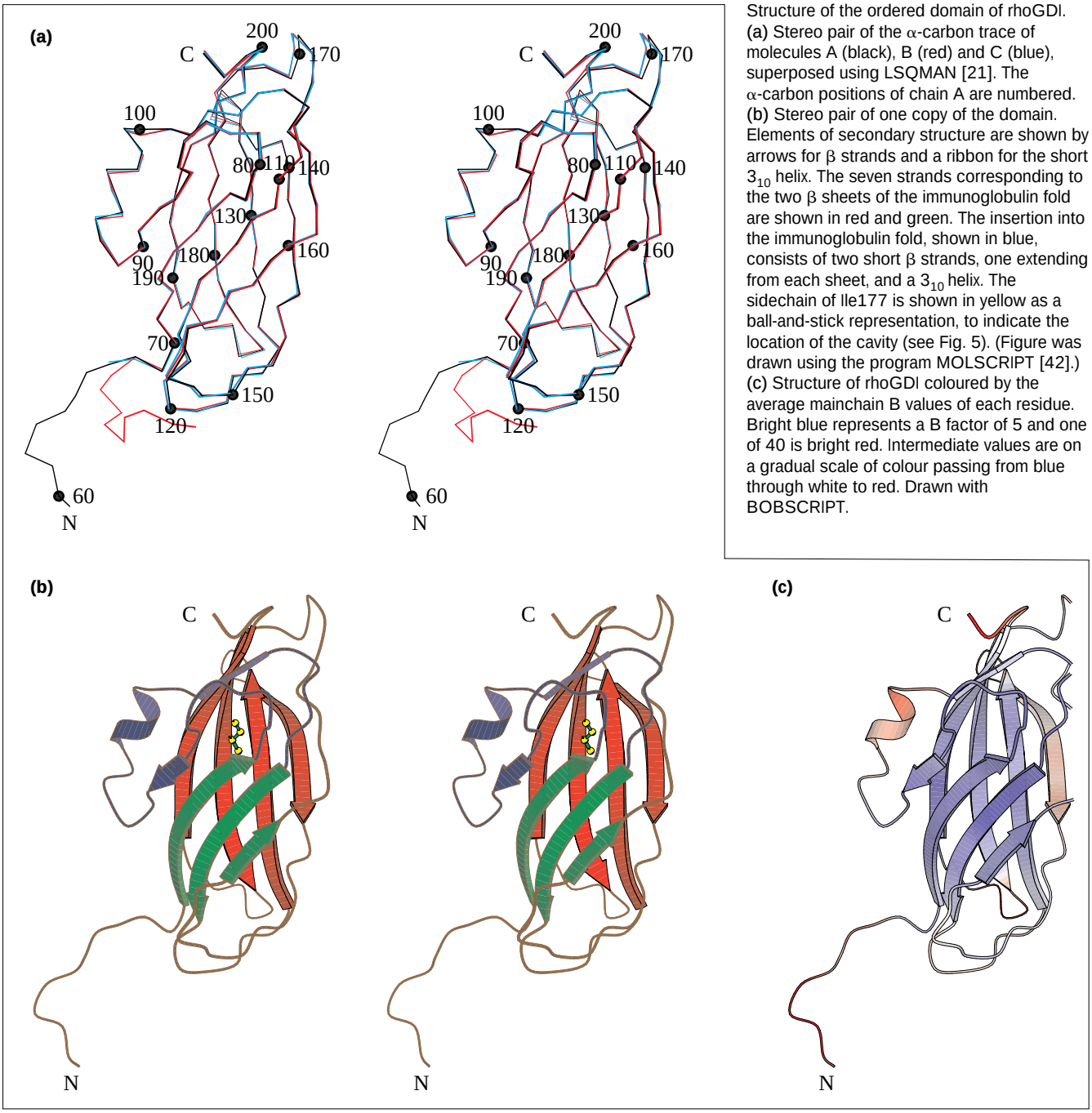
A region of the electron-density map of crystal structure of rhoGDI containing parts of three of the β strands. Residues 131, 177, 181, 190 and 193 are labelled to assist orientation. (a) Stereo pair of the original multiple isomorphous replacement solvent-flattened electron-density map (MIR figure of merit = 0.39; and after solvent flattening with the program dm is 0.84). (b) Stereo pair of the final $2F_o–F_c$ map, with the final atomic model superimposed in thick black lines. Drawn with BOBSCRIPT (R Esnouf's modification of MOLSCRIPT [42]).

Geranylgeranylation of the G protein is essential for tight binding to rhoGDI [16,27], so recognition of the isoprenyl group is likely to be part of the function of rhoGDI. We suggest that this cavity in the surface of rhoGDI may be involved in binding the isoprenyl group of post-translationally modified rho, rac and cdc42, and that the ordered domain as a whole interacts with the C-terminal region of the G proteins. This C-terminal region was shown to be disordered in the recently determined structure of rac1a (M Hirshberg and G Dodson, personal communication) and in general has been removed by proteolysis or engineering from the G-protein structures, determined so far (e.g. for rac1, [28]). Since the N-terminal structure is not known and is probably very flexible, it is difficult to model the complex of rhoGDI with a G protein from the currently available structures. The cavity in rhoGDI is the right size to accommodate a geranylgeranyl group (Fig. 4a), and the regions of negative electrostatic potential in and near the cavity might interact with the stretch of basic amino acids that immediately precede the isoprenylated cysteine in the G-protein sequence. Electron-density maps from crystals soaked in isoprenyl derivatives of cysteine or from crystals grown in the presence of such derivatives showed no density for an isoprenyl group. However, in the crystal form reported here (which was the only one was obtained), Leu170 from an adjacent copy of the molecule is inserted into the cavity (Fig. 5c), preventing the cavity being occupied by an isoprenyl group in this crystal form. Addition of *N*-acetyl-*S*-farnesylcysteine to either full length or trypsin-truncated rhoGDI did, however, produce a limited number of specific changes in the HSQC spectrum of the protein. These changes all

involved resonances from the ordered domain, showing that the isoprenyl group does interact with the structured region of rhoGDI in solution (LYL, RB, IB, NK, GCKR, unpublished results).

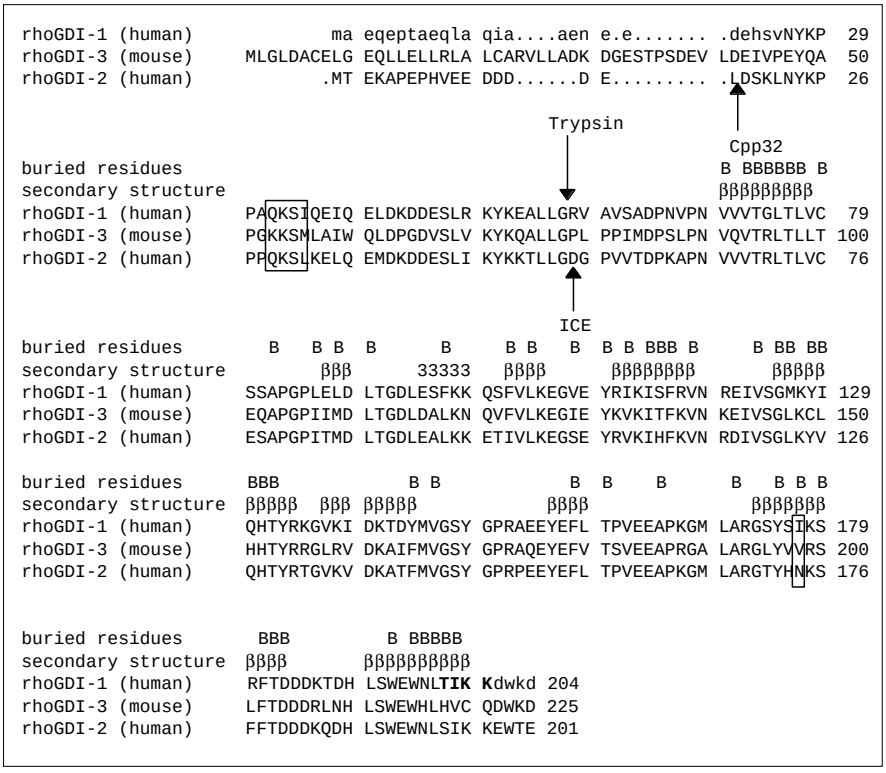
The calycin family (lipocalins, avidins and fatty acid binding proteins) also bind largely hydrophobic ligands in a cavity between their β sheets [29]. Although the calycins all have a 3_{10} helix, rhoGDI does not share the structural features common to this family of proteins, which consist of two β sheets of at least four strands each, at roughly 90° to each other. Four strands of one sheet and two of the other superimpose very closely throughout the family [30]. There is some similarity in the size and amino acid composition of the cavities in the two protein types, for example both contain several aromatic residues. The more

Figure 3



Structure of the ordered domain of rhoGDI. (a) Stereo pair of the α-carbon trace of molecules A (black), B (red) and C (blue), superposed using LSQMAN [21]. The α-carbon positions of chain A are numbered. (b) Stereo pair of one copy of the domain. Elements of secondary structure are shown by arrows for β strands and a ribbon for the short 3_{10} helix. The seven strands corresponding to the two β sheets of the immunoglobulin fold are shown in red and green. The insertion into the immunoglobulin fold, shown in blue, consists of two short β strands, one extending from each sheet, and a 3_{10} helix. The sidechain of Ile177 is shown in yellow as a ball-and-stick representation, to indicate the location of the cavity (see Fig. 5). (Figure was drawn using the program MOLSCRIPT [42].) (c) Structure of rhoGDI coloured by the average mainchain B values of each residue. Bright blue represents a B factor of 5 and one of 40 is bright red. Intermediate values are on a gradual scale of colour passing from blue through white to red. Drawn with BOBSCRIPT.

Figure 4



Alignment of the amino acid sequences of rhoGDI-1 (rhoGDI), rhoGDI-2 (D4/Ly-GDI) and rhoGDI-3, carried out using the CLUSTAL W program [43], summarizing some of the known biochemical information. The secondary structure is shown as β for a β strand and 3 for the 3_{10} helix. Buried residues are indicated by B; the criterion used was that the solvent accessible area, using a probe sphere of 1.4 Å in the CCP4 programs AREAIMOL and RESAREA, should be less than 30 $Å^2$ (20 $Å^2$ for glycines) in all three copies. The trypsin cleavage site in rhoGDI reported in this paper and the Cpp32 and ICE cleavage sites in Ly(D4)-GDI [33,34] are marked. The residues that can be deleted from rhoGDI-1 without loss of binding activity are in lower case, and those whose deletion causes loss of activity are in bold. The single residue difference between rhoGDI-1 and rhoGDI-2, which accounts for most of the difference in their affinities for cdc42 [25], is in the smaller box. The possible effector-binding region discussed in the text is in the larger box.

open rhoGDI cavity may be related to a clear functional difference between rhoGDI and the calycins—the hydrophobic ligand of rhoGDI, unlike those of the calycins, is attached to a protein, which could provide amino acids to close the open face of the cavity.

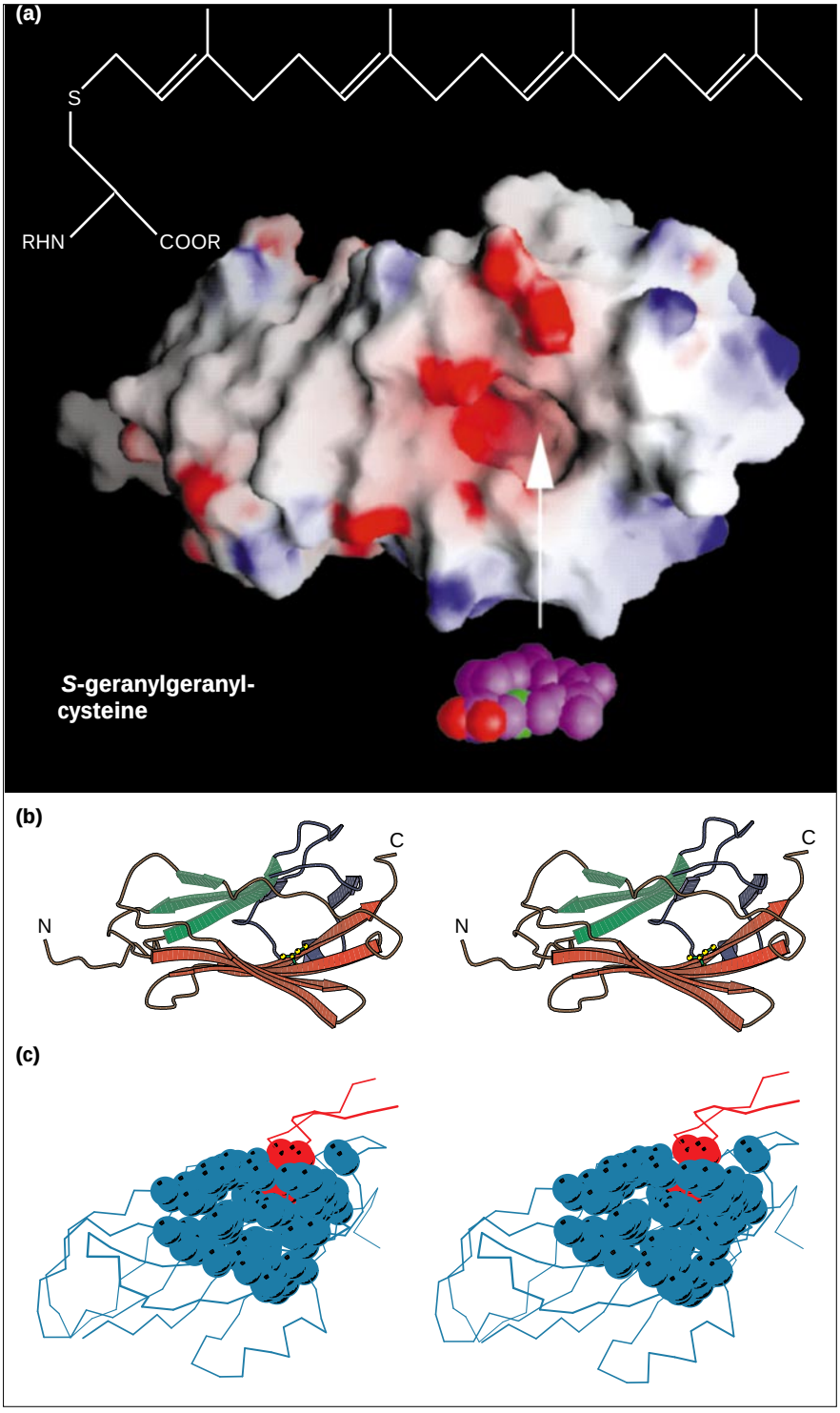
The function of the flexible N-terminal domain

Addition of recombinant non-isoprenylated rac1 to ^{15}N-labelled rhoGDI led to specific changes in the chemical shifts of a limited number of cross-peaks in the ^{15}N–^{1}H HSQC spectrum. These changes involve both some of the sharp resonances from the N-terminal region and some of the resonances from the structured domain (Fig. 6). When the experiment was repeated with truncated rhoGDI, the spectrum of rhoGDI was unaffected (data not shown), demonstrating that the N-terminal domain is required for binding rac. Recognition of rac by rhoGDI thus involves both domains of the rhoGDI molecule.

In discussing the structure of rabGDI, Schalk *et al.* suggested that the complex with rab might be formed by an extended β sheet [26], by analogy to the only G-protein–effector structure reported to date, the rap–raf complex [31]. In the latter structure, the main interaction is between the edge of the effector β strand of rap1A and the edge of a β strand from the raf domain. Although this is an edge-to-edge β-sheet interaction, many of the hydrogen bonds are between sidechains and most of the mainchain interactions expected in a β sheet are missing; there is also an interaction between the helix in the raf protein and the same effector β strand of rap. There are several exposed β-sheet edges in the rhoGDI structure (Fig. 5b), but it remains to be seen whether this type of sheet–sheet interaction is general. The sequence in rhoGDI (residues 32–35; Gln-Lys-Ser-Ile) most similar to the ras-binding region of raf (66–69; Gln-Arg-Thr-Val) lies in the flexible N-terminal region (Fig. 4). If the effector region of the G protein is involved in binding to rhoGDI, the N-terminus of rhoGDI could wrap round to bind the effector region of rac or rho, while the immunoglobulin-like domain binds the C-terminus. From the structure of rac-1 [28] and a model of cdc42 [32], it is unlikely that the ordered domain could interact both with the C-terminus and with the effector loop, because they are likely to be separated by too large a distance.

The first 25 residues of full length rhoGDI, which show little sequence conservation between rhoGDI-1 and rhoGDI-2 (Figure 4), are not required for binding to isoprenylated cdc42 [25]. This suggests that the interaction with rac probably involves residues 26–58 (including the residues 32–35, noted above), and the first 25 residues may remain mobile in the complex. Indeed, the HSQC spectrum shows that many of the sharp cross-peaks remain

Figure 5

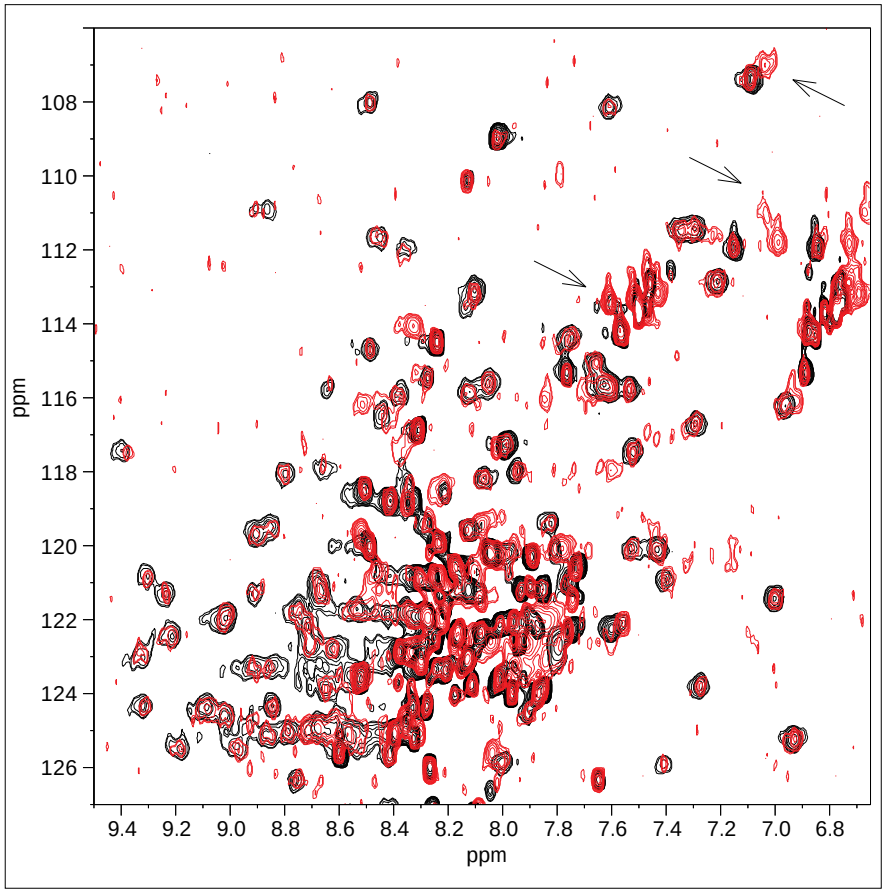


lsoprenyl-binding pocket of rhoGDI. (a) GRASP [44] representation of the accessible surface of the ordered domain of rhoGDI. Regions of positive and negative electrostatic potential are shown in blue and red, respectively. The cavity is near the centre of the surface (marked with an arrow), close to a region of negative potential. Geranylgeranylcysteine is shown below the structure to allow its size to be compared with that of the cavity. (b) Stereo pair of the structure of rhoGDI in the same orientation as (a). Colour scheme and secondary structure elements are depicted as for Figure 3b. (c) Stereo pair of the mainchain trace of molecule A (blue) and residues 164–177 of a symmetry generated copy of molecule C (red). The sidechain atoms of Leu170 from the copy of molecule C are shown as red spheres, and the sidechain atoms of residues 77, 102, 104, 110, 112, 128, 130, 132, 139, 140, 142, 144, 163, 165, 171, 175, 179, 194, 196 and 198 of molecule A are drawn as blue spheres to show the cavity. The red of the inserted leucine residue 170 (molecule C) can be seen in the gaps which result from the fact that the mainchain atoms are not drawn as spheres. Drawn using MOLSCRIPT [42].

sharp on addition of rac, and hence that the N-terminal region retains a large degree of flexibility in the complex. The rhoGDI-2 protein has been shown to be cleaved specifically in the N-terminal region by two members of the ICE/CED-3 family of cysteine proteinases (Fig. 4) [33,34]. In particular, cleavage by IL-1β-converting enzyme (ICE) leads to inactivation of the GDI [33], and it has been proposed that cleavage and inactivation of rhoGDI-2 may have a regulatory role in inflammation and apoptosis [33,34]. It is striking that the ICE cleavage site

Figure 6

The ^{15}N–^{1}H HSQC spectra of intact rhoGDI alone (black) and in the presence of an equimolar amount of rac1 (red). Examples of resonances affected by rac binding are indicated by arrows.



is in a position exactly equivalent, in terms of the sequence alignment in Figure 4, to the site of trypsin cleavage in rhoGDI-1 reported here (Fig. 4). Given the strong sequence similarity between rhoGDI-1 and rhoGDI-2, it is very likely that the N-terminal region of rhoGDI-2 is also flexible, and is also essential for binding to rho family G proteins. It remains to be seen whether cleavage at this site has any physiological significance as a method of rhoGDI-1 (as well as rhoGDI-2) inactivation *in vivo*.

Biological implications

Members of the rho family of ras-related GTPases, including rho, rac and cdc42, are involved in the reorganization of the actin cytoskeleton in response to external stimuli, in cell transformation by ras, in cytokinesis and in focal adhesion formation. Additionally, rac has a specific role in the activation of the phagocytic NADPH oxidase. There are two known groups of guanine nucleotide dissociation inhibitors (GDIs) that inhibit release of guanine nucleotides from members of the rab family (rabGDI and its homologues) and the rho family (rhoGDI and its homologues). These two groups of GDIs show no sequence homology with each other. GDIs can also control the partitioning of their cognate G proteins between the membrane and cytosol.

Our studies of rhoGDI reveal that it consists of two regions, a flexible N-terminal region and an ordered immunoglobulin-like domain, and we report the 2.5 Å crystal structure of the ordered domain. The structure of rhoGDI has no similarity to that of rabGDI. The ordered region of rhoGDI is responsible for binding to the C-terminal isoprenyl group of the post-translationally modified GTPases. We identify a possible site for this interaction, in a cavity between the **β** sheets of the ordered domain. This cavity provides a structural explanation of how rhoGDI controls the partitioning of the G proteins between the membrane and the cytosol, suggesting that it competes with the membrane for binding the G protein isoprenyl group. Both the folded domain and the flexible N-terminal arm are essential for the rac-binding function of rhoGDI. The flexible domain may interact with the nucleotide-binding loop of the G protein, preventing dissociation of the nucleotide. A comparison of rhoGDI with the closely related D4/Ly-GDI (rhoGDI-2) reveals that proteolytic cleavage between

the flexible and structured regions of rhoGDI occurs at an equivalent position to cleavage of rhoGDI-2. Because cleavage of rhoGDI-2 leads to its inactivation, cleavage of rhoGDI may have a role in regulating the activity of members of the rhoGDI family — and hence in the regulation of G proteins of the rho family.

Materials and methods

Purification and crystallization

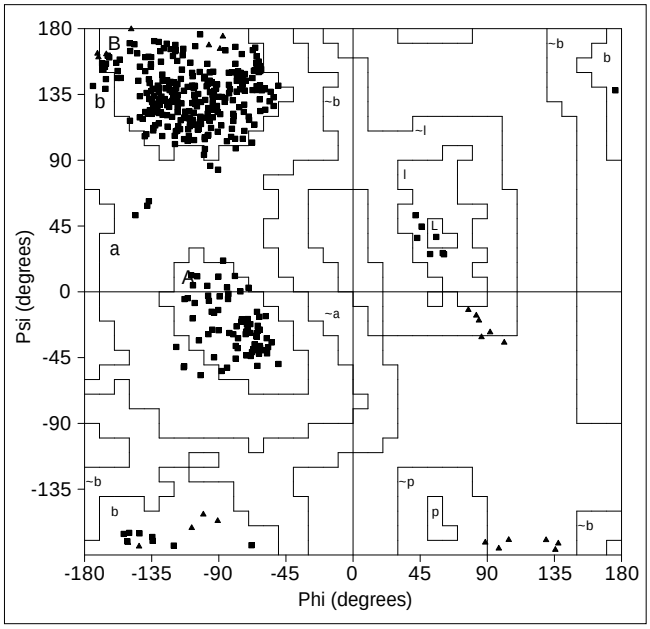
A GST–rhoGDI fusion protein was expressed in *E. coli* using the vector pGEX2T and purified using glutathione Sepharose; rhoGDI was released by thrombin treatment of the fusion protein bound to the column. Full-length material was further purified by anion-exchange (MonoQ) and gel-filtration chromatography. The truncated protein was prepared by treatment with 0.001% (w/v) trypsin for 2 hours at 4°C, the reaction was stopped by addition of 0.5 mM PMSF and the protein purified by gel filtration. N-terminal sequencing was carried out by the Leicester Protein and Nucleic Acid Chemistry Laboratory. Crystals were grown by the hanging-drop method. Equal volumes of the precipitant well (2 M ammonium sulphate in 0.1 M Tris pH 8.5, 1 mM DTT, 1 mM sodium azide) and the protein stock (10 mg ml^{-1} protein in 5 mM Tris pH 8.5, 1 mM DTT, 1 mM sodium azide) were mixed and placed on a coverslip sealed over 1 ml of precipitant mixture. Crystals grew at all temperatures tried between 4 and 22°C; the final data sets were from crystals grown at 4°C. The best crystals used for final data collection grew to $0.8 \times 0.3 \times 0.3$ mm.

For heavy-atom soaks, crystals were transferred to a mother liquor of 2 M ammonium sulphate in 0.1 M Tris, pH 8.5, 1 mM sodium azide, at room temperature. They were transferred to 2 M ammonium sulphate in 0.1 M Tris, pH 8.5, 1 mM sodium azide, 25% glycerol for freezing. The crystals could be stored unfrozen for several weeks in the glycerol mixture without significant deterioration. Selenomethionine protein (rhoGDI) was produced by growth of a methionine auxotroph in medium supplemented by selenomethionine. ^{15}N-labelled protein was prepared by growth in minimal medium containing $^{15}NH_4Cl$ as the sole nitrogen source. Recombinant rac1 was also expressed in *E. coli* as a GST fusion and purified in a similar manner.

Crystal structure determination

Potential heavy-atom derivatives were screened using the protein crystallography facilities at Leicester (Rigaku RU200HB rotating anode generator and Raxis IIc image plate system). Final diffraction data was collected at 100K on beamline BW7A at EMBL DESY Hamburg at wavelengths chosen to optimize the anomalous scattering (Table 1). The data were integrated and averaged with DENZO and SCALEPACK [35]. The three mercury sites were found from the anomalous difference data using SHELX [36]. The CCP4 [37] package was used to obtain the selenium sites by difference Fourier calculations. All nine selenium sites arose as the top nine sites in a difference Fourier of the selenium–native differences. The initial map from MLPHARE had a figure of merit of 0.39. Solvent flattening and histogram matching using the program dm [38] gave an interpretable map and improved the figure of merit to 0.84. Averaging between copies in the asymmetric unit was not used. A model was manually built into one copy using the program O [39]. The noncrystallographic symmetry (ncs) was used to generate the other two copies of this model. This model gave an initial R factor of 40% (R_{free} of 42%) Refinement was carried out against the selenomethionine data set, using selenium scattering factors, as this was to the highest resolution. The ncs restraints were removed early in the refinement, as it was clear there were significant differences in the N terminal region. Alternate cycles of refinement using the restrained geometry regularization (prepstage), simulated annealing (slow cool) or B-factor refinement protocols in X-PLOR [40] and manual rebuilding and addition of water molecules and sulphate ions using F_o–F_c maps in O was continued until no further interpretation of the largest remaining differences was possible. This gave a final R factor of 22.5% (R_{free} of 29.8%). The structure was also checked by recalculating the R factors using the native data and sulphur scattering factors. This only gave a small increase in R factors (Table 1) without refining any parameters against this data. Analysis of the structure using PROCHECK [41] confirmed that the final geometry was at least as good as expected for a structure of this resolution and that all the residues lay in the allowed regions of the Ramachandran plot (Fig. 7). The overall G factor was 0.1 in PROCHECK [41]

Figure 7



Ramachandran plot of all residues produced by PROCHECK [41]. Glycines are represented by triangles, other residues by squares.

NMR spectroscopy

Samples contained 0.1–0.5 mM uniformly ^{15}N-labelled rhoGDI in 20 mM phosphate, pH 6.3. All experiments were carried out at 293K on Bruker DMX500 and AMX600 spectrometers. Two-dimensional ^{15}N–^{1}H HSQC (or HMQC) spectra were acquired with gradient water suppression. Unlabelled rac1, *N*-acetyl-*S*-farnesylcysteine or *N*-acetyl-*S*-geranylgeranylcysteine were added in amounts equimolar with rhoGDI. Geranylgeranylcysteine caused nonspecific line-broadening and precipitation.

Accession numbers

The coordinates have been deposited in the Protein Data Bank with an accession code of 1RHO.

Acknowledgements

This work was supported by grants from the Wellcome Trust, Medical Research Council, Biotechnology and Biological Sciences Research Council and European Union.

References

1. Boguski, M.S. & McCormick, F. (1993). Proteins regulating ras and its relatives. *Nature* **366**, 643–654.
2. Hall, A. (1994). Small GTP binding proteins and the regulation of the actin cytoskeleton. *Ann. Rev. Cell Biol.* **10**, 31–54.
3. Nobes, C.D. & Hall, A. (1995). Rho, rac and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellopodia and filopodia. *Cell* **81**, 53–62.

4. Machesky, L.M. & Hall, A. (1996). Rho: a connection between membrane receptor signalling and the cytoskeleton. *Trends Cell Biol.* **6**, 304–310 .
5. Symons, M. (1996). Rho family GTPases: the cytoskeleton and beyond. *Trends Biochem. Sci.* **21**, 178–181
6. Narumiya, S. (1996). The small GTPase Rho: cellular functions and signal transduction. *J. Biochem* **120**, 215–228.
7. Lim, L., Masner, E., Leung, T. & Hall, C. (1996). Regulation of phosphorylation pathways by p21 GTPases. *Eur. J. Biochem.* **242**, 171–185.
8. Abo, A., Pick, E., Totty, N., Tehan, C.G. & Segal, A.W. (1991). Activation of the NADPH oxidase involves the small GTP-binding protein $p21^{rac1}$. *Nature* **353**, 668–670.
9. Abo, A., Webb, M.R., Grogan, A. & Segal, A.W. (1994). Activation of NADPH oxidase involves the dissociation of p21 from its inhibitory GDP/GTP exchange protein (rhoGDI) followed by its translocation to the plasma membrane. *Biochem. J.* **298**, 585–591.
10. Fukomoto, Y., *et al.*, & Takai, Y. (1990). Molecular cloning and characterization of a novel type of regulatory protein (GDI) for the rho proteins, ras p21-like small GTP-binding proteins. *Oncogene* **5**, 1321–1328.
11. Leilas, J.M., *et al.*, & Lim, B. (1993). cDNA cloning of a human mRNA preferentially expressed in hematopoietic cells and with homology to a GDP-dissociation inhibitor for the rho GTP-binding proteins. *Proc. Natl. Acad. Sci. USA* **90**, 1479–1483.
12. Scherle, P., Behrens, T. & Staudt, L.M. (1993). Ly-GDI, a GDP-dissociation inhibitor of the rhoA GTP binding protein is expressed preferentially in lymphocytes. *Proc. Natl. Acad. Sci. USA* **90**, 7568–7572.
13. Zalcman, G., *et al.*, & Olofsson, B. (1996). RhoGDI-3 is a new GDP dissociation inhibitor (GDI). *J. Biol. Chem.* **271**, 30366–30374.
14. Clarke, S. (1992). Protein isoprenylation and methylation at carboxy-terminal cysteine residues. *Ann. Rev. Biochem.* **61**, 355–386.
15. Casey, P.J. & Seabra, M.C. (1996). Protein prenyltransferases. *J. Biol. Chem*. **271**, 5289–5292.
16. Hori, Y., *et al.*, & Takai, Y. (1991). Post translational modification of the C-terminal region of the rho protein are important for its interaction with the membranes and inhibitory GDP/GTP exchange proteins. *Oncogene* **6**, 515–522.
17. Musha, T., Kawata, M. & Takai,Y. (1992). The geranylgeranyl moiety but not the methyl moiety of the smg25a/rab3a protein is essential for the interaction with membrane and its inhibitory GDP/GTP exchange protein. *J. Biol. Chem.* **267**, 9821–9825.
18. Araki, S., Kikuchi, A., Hata, Y., Isomura, M. & Takai, Y. (1990). Regulation of reversible binding of smg25A, a ras p21-like GTP-binding protein, to synaptic plasma membranes and vesicles by its specific regulator protein, GDP dissociation inhibitor. *J. Biol. Chem.* **265**, 13007–13015.
19. Isomura, M., Kikuchi, A., Ohga, N. & Takai, Y. (1991). Regulation of binding of rhoB p20 to membranes by its specific regulatory protein GDP dissociation inhibitor. *Oncogene* **6**, 119–124.
20. Lang, P., Gesbert, F., Delspine-Carmagnat, M., Stancou, R., Pouchelet, M. & Bertoglio, J. (1996). Protein kinase A phosphorylation of rhoA mediates the morphological and functional effect of cyclic AMP in cytotoxic lymphocytes. *EMBO J.* **15**, 510–519.
21. Kleywegt, G.J. & Jones, T.A. (1994). A superposition. *ESF/CCP4 Newsletter* **31**, 9–14.
22. Holm L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **223**, 123–138.
23. Murzin, A.G., Brenner, S.E., Hubbard, T. & Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536–540.
24. Bork, P., Holm, L. & Sander, C. (1994). The immunoglobulin fold. Structural classification, sequence patterns and common core. *J. Mol. Biol.* **242**, 309–320.
25. Platko, J.V., Leonard, D.A., Adra, C.N., Shaw, R.J., Cerione, R.A. & Lim, B. (1995). A single residue can modify target-binding of the rho-subfamily GDP dissociation inhibitors. *Proc. Natl. Acad. Sci. USA* **92**, 2974–2978.
26. Schalk, I., *et al.*, & Balch, W.E. (1996). Structure and mutational analysis of Rab GDP-dissociation inhibitor. *Nature* **381**, 42–48.
27. Nomanbhoy, T.K. & Cerione, R.A. (1996). Characterization of the interaction between rhoGDI and Cdc42Hs using fluorescence spectroscopy. *J. Biol. Chem.* **271**, 10004–10009.
28. Hirshberg, M., Stockley, R.W., Dodson, G. & Webb, M.R. (1997). The crystal structure of human rac1, a member of the rho-family, complexed with a GTP analogue. *Nature Struct. Biol.* **4**, 147–152.
29. Flower, D.R. (1996). The lipocalin protein family: structure and function. *Biochem. J.* **318**, 1–14.
30. Flower, D.R. (1993). Structural relationship of streptavidin to the calycin protein superfamily. *FEBS Letts* **333**, 99–102.
31. Nassar, N., Horn, G., Hermann, C., Scherer, A., McCormick, F. & Wittinghofer, A. (1995). The 2.2 Å crystal structure of the ras-binding domain of the serine/threonine kinase c-raf1 in complex with rap1A and a GTP analogue. *Nature* **375**, 554–560.
32. Sutcliffe, M.J., Feltham, J., Cerione, R.A. & Oswald, R.E., (1994). Modelling predicts an additional switch when GTP binds to the cdc42Hs protein. *Prot. Pept. Letts.* **1**, 84–91.
33. Danley, D.E., Chuang, T.-H. & Bokoch, G.M. (1996). Defective rho GTPase regulation by IL-1β-converting enzyme-mediated cleavage of D4 GDP dissociation inhibitor. *J. Immunol.* **157**, 500–503.
34. Na, S.Q., *et al.*, & Danley, D.E. (1996). D4-GDI, a substrate of CPP32, is proteolyzed during Fas-induced apoptosis. *J. Biol. Chem.* **271**, 11209–11213.
35. Otwinowski, Z. (1993). Oscillation data reduction program. In *Data Collection and Processing* (Sawyer, L., Issacs, N. & Bailey, S., eds) pp 56–62. SERC Daresbury Laboratory, Warrington, UK.
36. Sheldrick, G.M., Dauter, Z., Wilson, K.S., Hope, H. & Sieker, L.C. (1993). The application of direct methods and patterson interpretation to high-resolution native protein data. *Acta Cryst. D* **49**, 18–23.
37. CCP4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760–763.
38. Cowtan, K.D. & Main, P. (1996). Phase combination and cross validation in iterated density modification calculations. *Acta Cryst. D* **52**, 43–48.
39. Jones, T.A. & Kjeldgaard, M. (1995). O: the manual (version 5.10). Uppsala University, Sweden.
40. Brünger, A.T. (1992). *X-PLOR Version 3.1: a System for X-Ray Crystallography and NMR*. Yale University Press, New Haven, CT, USA.
41. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the sterochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.
42. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structure. *J. Appl. Cryst.* **24**, 946–950.
43. Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acid Res.* **22**, 4673–4680.
44. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296.