Three-dimensional structure of the Ras-interacting domain of RalGDS

The Ras-interacting domains of the the protein-kinase Raf and the Ral guanine nucleotide dissociation stimulator, RalGDS, lack extensive sequence similarity, but their overall three-dimensional structures are very similar to each other. Mutational analysis indicated that three residues in the RalGDS domain are critical for its interaction with Ras.

Sir — The Ha-*ras* proto-oncogene plays a critical role in regulating cell growth, motility and differentiation and is mutationally activated in many types of cancer[1]. The Ras protein is a membrane-localized guanine nucleotide-binding protein (with a relative molecular mass of ~21,000) and is the prototype of a family of small GTPases that function as molecular switches[2]. This family includes a number of proteins whose biological functions are not clearly understood, including Rap1A, which may modulate Ras signaling[3], and Ral (Ras-like) proteins, which appear to regulate phospholipase D[4–6]. The biological activities of Ras are mediated by interactions of the Ras protein with downstream signalling proteins referred to as Ras effectors[7]; similarly other members of the Ras family are believed to interact with effectors that are specific for each GTPase[3]. Ras and its relatives cycle between an active GTP-bound state, in which they bind effector molecules, and an inactive GDP-bound state[8]. The loading of GTP is accelerated by guanine nucleotide exchange factors (GEFs) or guanine nucleotide dissociation stimulators (GDSs), and GTP hydrolysis is accelerated by GTPase-activating proteins (GAPs)[8]. Binding of GTP results in major conformational changes in two highly conserved stretches of Ras, switch I and II[9]. In the case of Ha-Ras, switch I spans residues 30–37, which constitute most of the effector loop that is necessary for interaction with Ras effectors[10]. Switch II of Ha-Ras, consisting of residues 60–76, has been implicated in interactions with Ras guanine nucleotide exchange factors and in modulating the strength of effector binding[11].

Because of the many biological activities of Ras, the identity of its effectors and the structural basis for their interaction with Ras are questions of considerable interest. The best characterized of these effectors is the protein kinase Raf. Binding of Ras to a segment within the regulatory region of Raf appears to be the initial event in activation of Raf, and is a crucial step in cellular transformation initiated by oncogenic Ras[12,13]. The region of Raf responsible for binding to Ras is designated the Raf Ras-binding domain (Raf-RBD)[14]. As a result of binding to Ras, Raf is translocated to the plasma membrane, where the kinase is activated through phosphorylation by an as yet unknown mechanism[15,16]. Activation of Raf initiates a kinase cascade leading to the phosphorylation and activation of certain transcription factors, resulting in the induction of early response genes that are believed to mediate mitogenic signals[17–20].

Recently a second potential effector of Ras has been identified. This protein is a guanine nucleotide dissociation stimulator specific for the Ral GTPase, and is referred to as RalGDS[21]. We and others originally recognized RalGDS as a potential effector of Ras when it was recovered in two-hybrid screens using activated Ras as a bait[22–24]. The interaction between Ras and RalGDS was shown to be mediated by the C-terminal, non-catalytic segment of RalGDS, designated here the RalGDS Ras-interacting domain (RalGDS-RID). The interaction could be detected both *in vivo*, using the two-hybrid system, and *in vitro*, with purified recombinant proteins. As in the case of Raf, the interaction is dependent on activation of Ras by GTP, and is blocked by mutations that affect Ras effector function. While the catalytic (exchange-stimulating) domain of the RalGDS is specific for Ral, and does not regulate nucleotide exchange on Ras, the RalGDS-RID interacts with the Ras effector domain but not with the corresponding domain of Ral[22–24]. Like Raf, RalGDS also binds to other members of the Ras family that share an identical effector region, including R-Ras, K-Ras, and Rap1A. Recent reports suggest that Ras stimulates the exchange activity of RalGDS *in vivo*, leading to activation of Ral, and that this activation can contribute to transformation by Ras[4,25].

There is little sequence similarity between the Ras-binding domains of Raf and RalGDS. However, similarities in the disposition of hydrophobic residues in the RalGDS-RID and the Raf-RBD have been noted and it has been suggested that these domains are part of a large family of homologous Ras-associating (RA) domains[26].

Although the Ras effector domain can interact with more than one effector molecule, and although these effector molecules can interact with more than one GTPase, these interactions also display some specificity. Mutations within the Ras effector domain can differentially affect signalling through the Raf and Ral pathways[25,27]. In addition, Raf and RalGDS display different affinities for Ras and Rap1A respectively. The affinity of the Raf-RBD for Ras-GTP is nearly 70-fold higher than its affinity for Rap-GTP, whereas for RalGDS-RID the affinities are reversed, with Rap1A binding 100-fold tighter than Ras[14]. Recently the crystal structure of Rap1A-GTP complexed to Raf-RBD has been solved[28]. This structure has revealed specific interactions between amino acids within the effector region of Rap1A and Raf-RBD. Here we report the crystal structure of the RID of RalGDS. The structure, together with a mutational analysis of RalGDS-RID, has allowed us to compare the binding domains from two different Ras effector molecules.

The Ras-interacting domain (RID) of rat RalGDS constitutes RalGDS residues 767–864, which are designated here as residues 3–100 of RID; residues Gly 1 and Ser 2 are derived from the linker in the GST fusion protein. Expression, purification and crystallization of the RID have been described previously[29]. The structure was determined by the multiwavelength anomalous diffraction (MAD) method[30] on crystals of the seleno-methionyl mutant of the protein. Details of data collection and analysis, model building and structure refinement are presented in Table 1. A portion of the
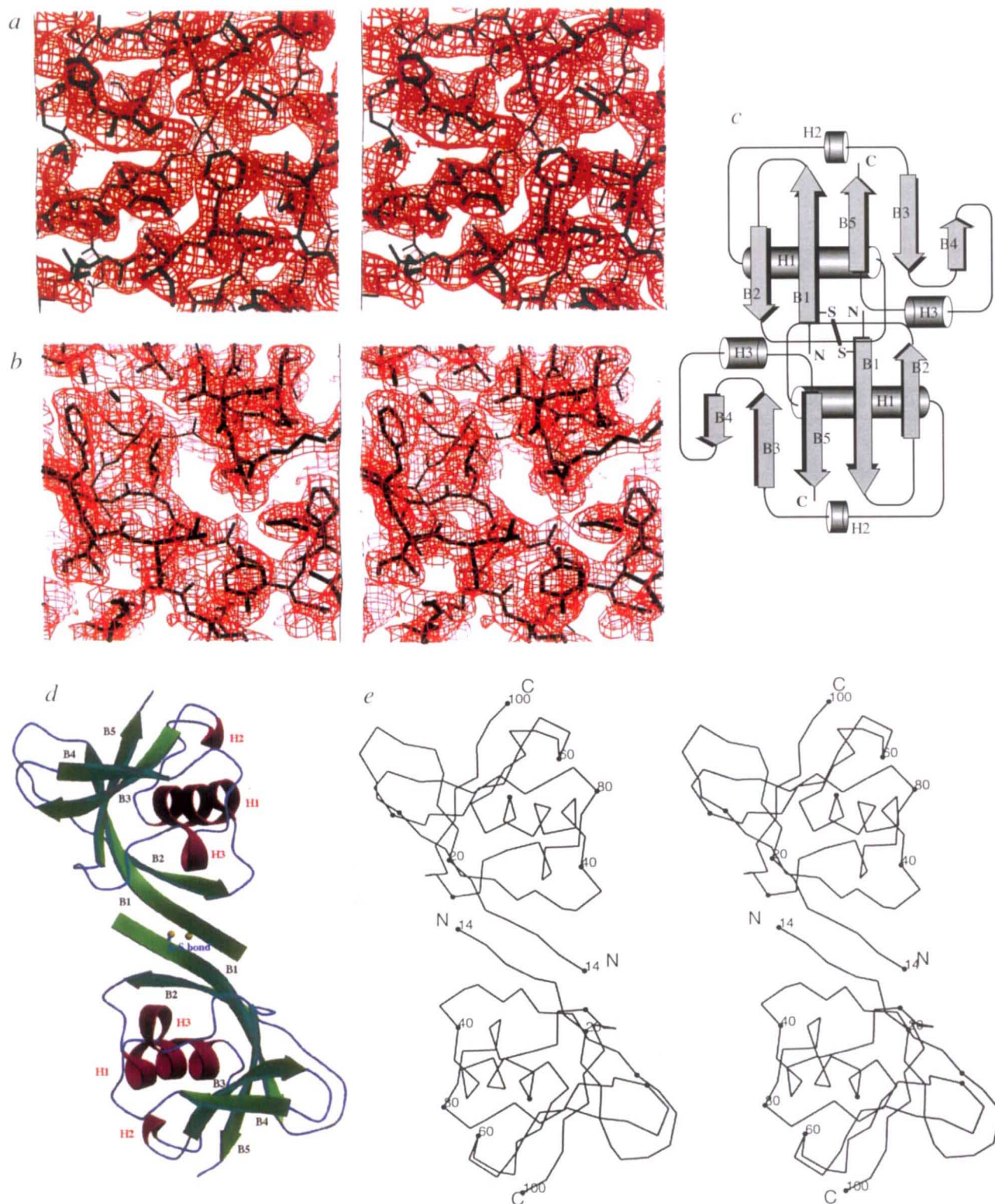
correspondence

Fig. 1 Stereo pair of the electron density map of a region in RalGDS-RID. *a*, The MAD-phased electron density map at 3.0 Å resolution is contoured at 1σ, with the current model displayed for comparison. *b*, The 2F_o - F_c electron density map at 2.4 Å resolution is contoured at 1σ, with the current model displayed for comparison. The pictures were generated by the computer program O[43]. *c*, Topological structure, where β-strands and α-helices are represented by arrows and cylinders respectively. *d*, Dimer structure of the Ras-interacting domain of RalGDS shown in ribbon diagram, where β-strands and α-helices are in green and red respectively, and a disulphide bond between two monomers is indicated. The secondary structure elements as assigned by DSSP[44] are labelled. The ribbon representation was created using the program MOLSCRIPT[45]. *e*, Stereo view of a Cα trace of the structure in the same orientation as in (*d*). Every tenth residue is marked by a small filled circle and every twentieth residue is labelled.
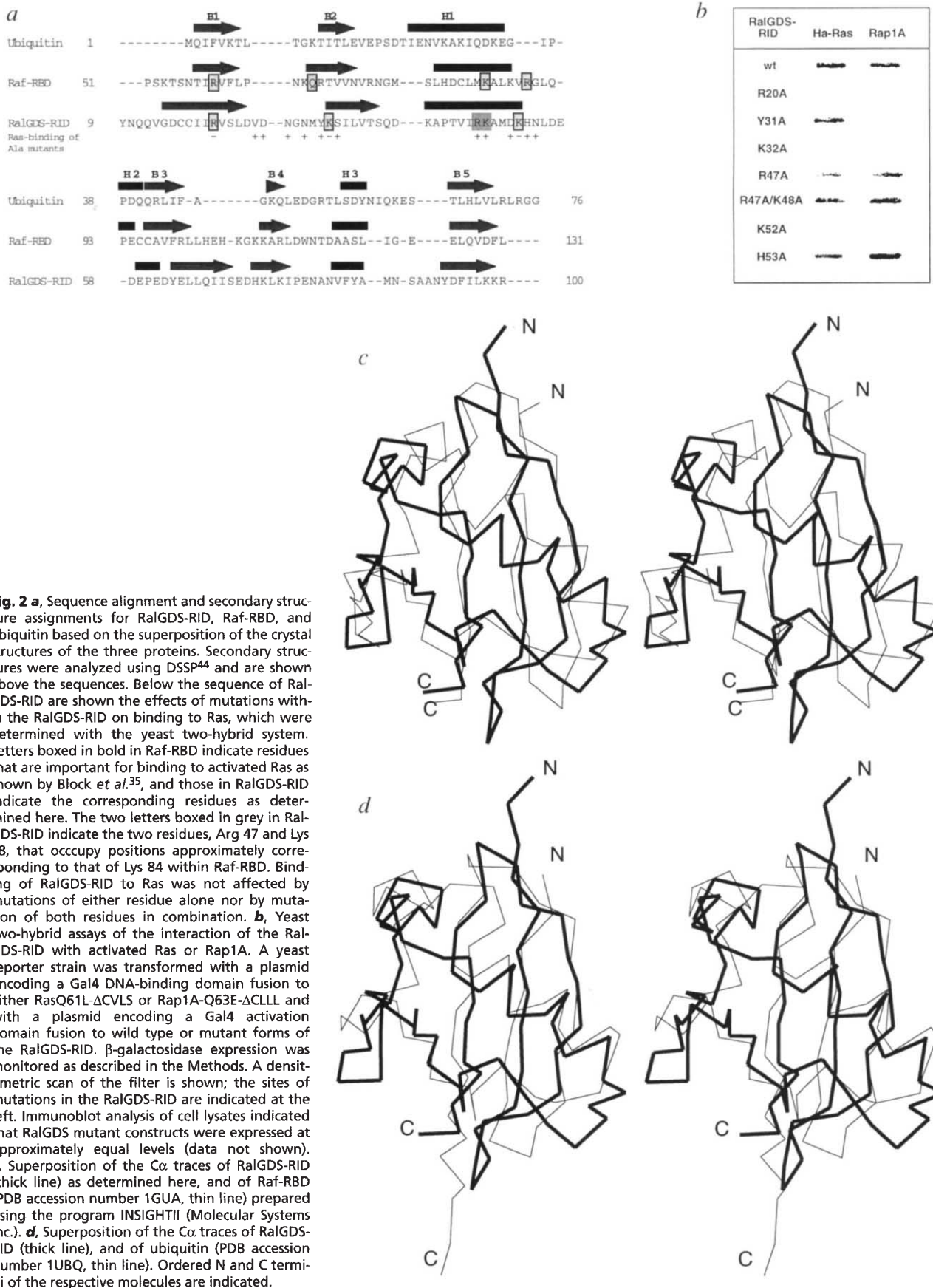
correspondence

Fig. 2 *a*, Sequence alignment and secondary structure assignments for RalGDS-RID, Raf-RBD, and ubiquitin based on the superposition of the crystal structures of the three proteins. Secondary structures were analyzed using DSSP[44] and are shown above the sequences. Below the sequence of RalGDS-RID are shown the effects of mutations within the RalGDS-RID on binding to Ras, which were determined with the yeast two-hybrid system. Letters boxed in bold in Raf-RBD indicate residues that are important for binding to activated Ras as shown by Block *et al.*[35], and those in RalGDS-RID indicate the corresponding residues as determined here. The two letters boxed in grey in RalGDS-RID indicate the two residues, Arg 47 and Lys 48, that occcupy positions approximately corresponding to that of Lys 84 within Raf-RBD. Binding of RalGDS-RID to Ras was not affected by mutations of either residue alone nor by mutation of both residues in combination. *b*, Yeast two-hybrid assays of the interaction of the Ral-GDS-RID with activated Ras or Rap1A. A yeast reporter strain was transformed with a plasmid encoding a Gal4 DNA-binding domain fusion to either RasQ61L-ΔCVLS or Rap1A-Q63E-ΔCLLL and with a plasmid encoding a Gal4 activation domain fusion to wild type or mutant forms of the RalGDS-RID. β-galactosidase expression was monitored as described in the Methods. A densitometric scan of the filter is shown; the sites of mutations in the RalGDS-RID are indicated at the left. Immunoblot analysis of cell lysates indicated that RalGDS mutant constructs were expressed at approximately equal levels (data not shown). *c*, Superposition of the Cα traces of RalGDS-RID (thick line) as determined here, and of Raf-RBD (PDB accession number 1GUA, thin line) prepared using the program INSIGHTII (Molecular Systems Inc.). *d*, Superposition of the Cα traces of RalGDS-RID (thick line), and of ubiquitin (PDB accession number 1UBQ, thin line). Ordered N and C termini of the respective molecules are indicated.
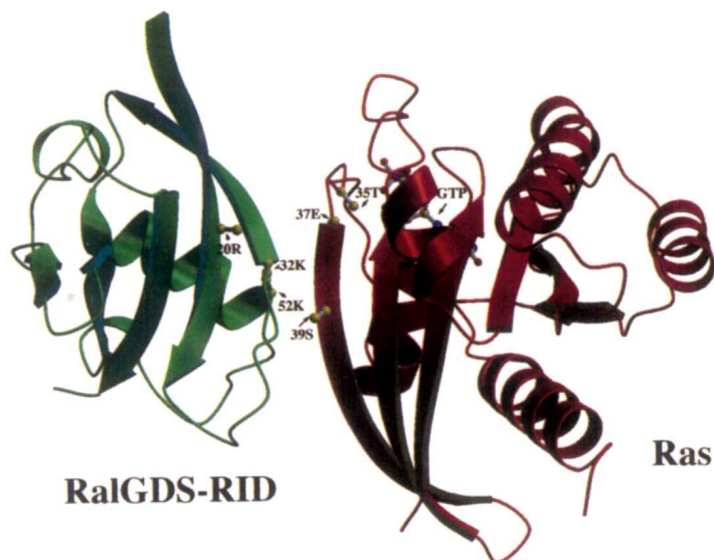
correspondence

Fig. 3 A model of RalGDS-RID binding to active Ras, based on the structural similarities of RalGDS-RID with Raf-RBD and Ras with Rap1A using the crystal structure of the complex between Raf-RBD and active Rap1A[28] as the template. The coordinates of RalGDS-RID were superimposed onto those of Raf-RBD in the structure of the complex, and active Rap1A was replaced by active Ras[46]. Residues within RalGDS-RID important for binding to Ras are shown and labelled. The figure was generated by the program MOLSCRIPT[45]. The coil region before the N terminus of the β-strand B2 may converts to a β-strand upon binding to Ras, thus extending B2, to form an intermolecular β-sheet between Ras and RalGDS-RID.

electron density map computed with the experimental MAD phases and the phases calculated from the atomic model are shown in Fig. 1a and 1b respectively.

RID forms a homodimer in solution (L.H. and S.-H.K., unpublished results) and in the crystal. The interdimer surface is composed of two cysteines (Cys 16 in each monomer) forming an intermolecular disulphide bond and two interacting intermolecular antiparallel β-sheets. The majority of monomer RID consists of a five-stranded mixed β-sheet (B1–B5) interrupted by a 12 residue α-helix (H1) and two additional small α-helices (H2 and H3) (Fig. 1c,d,e). The N-terminal 13 residues are disordered. These N-terminal residues are not required for binding to active Ras (F.H. and G.S.M., unpublished results). The backbone topology of the RalGDS-RID monomer begins with an anti-parallel pair of β-strands (B1, residues 14–24, and B2, residues 32–38) arranged in a hairpin configuration. The turn (residues 25–29) is not well defined by the electron density map. H1 is composed of residues 41–52. Following the C terminus of the helix, a linker-segment runs from 53–58. H2 (residues 59–61) is a three residue 3_{10}-helix. B3 (residues 63–70) is arranged in an antiparallel manner with the remaining two β-strands (B4, residues 73–77 and B5, residues 93–99) in a partial Greek-key motif. H3 (residues

83–86) is a one-turn α-helix. The C-terminal β-strand (residues 93–99) forms a parallel pair with the N-terminal β-strand (residues 14–24) to complete the β-sheet. The hydrophobic core is formed by residues Ile 19, Val 21, Ile 34, Val 36, Ala 42, Val 45, Ile 46, Ala 49, Leu 65, Ile 77, Val 83, Phe 84, Ala 86, Phe 95 and Leu 97. Almost all of these residues belong to the consensus hydrophobic residues in ubiquitin and all RA domain sequences[26].

The root mean square (r.m.s.) difference between the Cα positions of all corresponding residues on the two monomers in the homodimer of RalGDS-RID, except those on poorly defined loops (residues 24–30 and residues 54–58) and the N-terminal 13 residues, is 0.635 Å. The surface loop consisting of residues 56–58 is well defined in one monomer, but not in the other. Among the well defined Cα residues in the two monomers, the largest differences occur at the termini (residues 14 and 15 near the N terminus, residues 99 and 100 at the C terminus), and in surface loops (residues 51 and 52 immediately C-terminal to H1, and residues 71 and 72 in the loop between B3 and B4).

Comparison with Raf-RBD, ubiquitin

Despite the low level of sequence homology (<15% sequence identity) among the primary sequences of RalGDS-RID, Raf-RBD[28,31], ubiquitin[32], the [2Fe-2S] hete-

rocyst ferrodoxin[33], and the IgG-binding domain of protein G from Streptococcus[34], they all exhibit similar topology. Among these RalGDS-RID is most similar to Raf-RBD (r.m.s. difference of 0.908 Å for 30 Cα atoms), and ubiquitin (r.m.s. difference of 1.004 Å for 28 Cα atoms). The spatial positioning of B1, B2 and B5 is nearly identical in the three proteins, and in all of them H1 is packed diagonally against part of the β-sheets. Fig. 2a shows the sequence alignment of RalGDS-RID, Raf-RBD and Ubiquitin based on their crystal structures.

When compared with the Raf-RBD structure[28] (Fig. 2c), the largest structural differences are found in the regions corresponding to the surface loops of RalGDS-RID: residues 25–31 in the loop connecting strands B1 and B2 (4–8 Å), residues 53–58 in the loop connecting helices H1 and H2 (3–6 Å), and the loop connecting H3 and B5 (5–8 Å). Residues 71 and 72 in RalGDS-RID connecting B3 and B4 form the only loop shorter than the corresponding loop in Raf-RBD; these loops deviate from each other by about 2–4 Å.

Comparison with the ubiquitin structure[32] (Fig. 2d) reveals that all loops in RalGDS-RID are longer than those of ubiquitin. The largest differences occur also on surface loops: residue 25–31 in the loop connecting strands B1 and B2 (4–7 Å), residues 53–58 in the loop connecting helices H1 and H2 (3–6 Å), residues 71–72 between B3 and B4 (4–8 Å), and the loop between H3 and B5 (4–7 Å).

Interaction of Ras with RalGDS-RID

Recently the crystal structure of the Ras homologue Rap1A complexed with Raf-RBD has been solved[28]. The two proteins interact with each other through intermolecular β-sheet formation. The crystal structure has revealed specific interactions between amino acids within the effector region of Rap1A and residues within the Raf-RBD located in B1, B2, and the end of the H1 helix. The residues within the Raf-RBD that are important for binding to Ras have been identified by mutational analysis[35]. The residues important for the binding affinity of Raf-RBD include residues Arg 59, Gln 66, Lys 84 and Arg 89, located within B1, the N-terminal segment of B2 and H1. Since we found that B1, B2, and the C-terminal end of H1 of RalGDS-RID and Raf-RBD are structurally similar, we examined the effect of mutations in these regions of the RalGDS-RID on their ability to interact with Ras and Rap1A in the yeast two-hybrid system; the results are summarized in Fig. 2a,

correspondence

Table 1 Structural Data Statistics

a, Statistics for data collection of Se-Met mutant crystal at Brookhaven X4A, 100K (30–2.3 Å) space group C2, cell a = 103.789 Å, b = 30.551 Å, c = 51.461 Å, β = 94.893°

Wavelength (Å)	Reflections[1] (N)	Completeness (%)	Signal ($I/\sigma(I)$)	R_{sym}[2] (%)
0.9879 (remote)	12801	93.2	13	7.7
0.9793 (edge)	12863	93.4	12	8.5
0.9791 (peak)	12853	93.3	12	8.4
0.9686 (remote)	12990	93.5	12.5	8.1

b, Statistics of phase determination
MAD structure factor ratios[3]

			Observed Ratios							
		$30.0 < d < 5.0$ Å				$5.0 < d < 2.5$ Å				Scattering factors (e)
Wavelength (Å)	0.9879	0.9793	0.9791	0.9686	0.9879	0.9793	0.9791	0.9686	f'	f''
a. 0.9879	0.044 (0.041)	0.058	0.050	0.032	0.051 (0.040)	0.056	0.048	0.042	-4.3	0.5
b. 0.9793		0.077 (0.048)	0.044	0.061		0.071 (0.049)	0.042	0.054	-10.9	2.9
c. 0.9791			0.116 (0.060)	0.051			0.099 (0.058)	0.046	-8.4	4.9
d. 0.9686				0.083 (0.046)				0.074 (0.046)	-3.7	3.3

MAD phasing[4]

$R(|^\circ F_T|) = 0.069$ $<\Delta(\Delta\phi)> = 23.38°$ $<m> = 0.92$
$R(|^\circ F_A|) = 0.313$ $<\sigma(\Delta\phi)> = 7.27°$

c, Refinement statistics[5] for data taken at Brookhaven X12B 100K(10–2.4 Å) space group C2, a = 105.282 Å, b = 30.714 Å, c = 51.326 Å, β = 94.566°

Wavelength (Å)	Reflections[1] (N)	Completeness (%)	Signal ($I/\sigma(I)$)	R_{sym}[2] (%)
1.0086	6485	99.3	16.6	9.0

Model: homodimer with 87 residues in a monomer, 58 water molecules

d-spacings (Å)	Reflections (N)	R-value (%)	Free R-value (%)	R.m.s. deviations Bonds (Å)	Angles (°)
6-2.4	6209	21.2	29.8	0.014	2.5

[1]Unique reflections in space group C2.
[2]$R_{sym} = 100 \times (\Sigma_{hkl} \Sigma_i |I_i - <I>|) /(\Sigma_{hkl} \Sigma_i I_i)$, where I_i is the ith measurement and $<I>$ is the weighted mean of all measurement of I.
[3]Table values represent $<\Delta|F|^2>^{1/2} / <|F|^2>^{1/2}$, where $\Delta|F|$ is the absolute value of the Bijvoet difference at one wavelength (diagonal elements) or of the dispersive differences at two wavelengths (off-diagonal element). The values in the parentheses are the ratios for centric Bijvoet reflections, which would be equal to zero for perfect data and serve as an estimate of the noise in the anomalous signals. The listed scattering factors are the refined values for selenium.
[4]$R = 100 \times \Sigma_{hkl} \Sigma_i ||F_i| - <|F|>| /\Sigma_{hkl}|F|$. $^\circ F_T$ is the structure factor due to normal scattering from all atoms. $^\circ F_A$ is the structure factor due to normal scattering from the anomalous scatters only, and $\Delta\phi$ (is the phase difference between $^\circ F_T$ and $^\circ F_A$. $\Delta(\Delta\phi)$ is the difference between two independent determinations of $\Delta\phi$. Values given are based on calculations that did not include reflections with m = 0; $<m>$ is the mean figure of merit including reflections with m = 0.
[5]A subset of the data (10%) was excluded from the refinement and used for the free R-value calculation until the final round of refinement, in which all the data ($F > 2\sigma$) were used. R-value = $100 \times \Sigma_{hkl} ||F_o| - |F_c|| / \Sigma_{hkl} |F_o|$.

and representative assays are shown in Fig. 2*b*. Interaction of the RalGDS-RID with Ras was abrogated by mutation of Arg 20, Lys 32, and Lys 52, which correspond to Arg 59, Gln 66, and Arg 89 of Raf-RBD respectively (Fig. 2*a,b*). Mutation of residues adjacent to these critical residues did not affect the interaction of RalGDS-RID with Ras (Fig. 2*a,b*), although a small decrease in affinity would not be detected by the two-hybrid filter assay. These findings indicate that most residues in RalGDS-RID that mediate binding to Ras are located at sites that correspond to those that mediate binding by Raf-RBD, despite the fact that the two domains lack extensive amino acid sequence similarity. Interestingly, however, we were not able to identify a residue within H1 of RalGDS-RID that corresponds to Lys 84 of Raf-RBD. It has been demonstrated that Lys 84 of Raf-RBD forms a salt bridge with Glu 31 of Ras, but is not involved in the interaction with Rap1A, which contains a Lys at position 31[36]. The affinity for the Raf-Ras interaction ($K_d = 0.018$ mM) is 70 times higher than that for the Raf-Rap1A interaction ($K_d = 1.2$ mM) or for the interaction between Ras and RalGDS ($K_d = 1.0$ mM). The absence of a residue corresponding to Lys 84 may account for the lower affinity of the RalGDS-RID for Ras. The effects of the mutations on binding to Rap1A were similar to the effects on binding to Ras, with the exception of the Y31A mutation which did not detectably affect binding to Ras, but completely abrogated binding to Rap1A. This suggests that Tyr 31 participates in a critical interaction with Rap1A that is not required for the interaction with Ras, and it is likely that this residue contributes to

correspondence

the higher affinity of the RalGDS-RID for Rap1A.

In summary, despite the lack of extensive sequence similarity between RalGDS-RID and Raf-RBD, they have the same overall structure, and the residues on RalGDS-RID that are required for binding to Ras are located at sites that correspond to Ras-binding residues within the Raf-RBD. Therefore, we suspect that the structure of the complex of active Ras and RalGDS-RID may be similar to that of active Rap1A and Raf-RBD (Fig. 3), and this similarity may be the structural basis for the existence of multiple effectors of Ras. The features that determine the binding specificity of the Ras-binding domains may be apparent when the structure of the complex of Ras and RalGDS-RID is solved.

Methods

The Ras interaction domain of rat RalGDS (amino acids 767–864) was expressed as a glutathione-S-transferase fusion protein. Expression, purification and crystallization of the wild type protein and the seleno-methionine derivative have been described[29]. In brief, the RalGDS-RID was cloned into the pGEX-2T vector to generate the expression plasmid pGEX98. The expressed protein was purified by binding to glutathione agarose (Pharmacia), followed by thrombin cleavage and chromatography on Mono Q and Superose 12 HR (Pharmacia). The crystals were grown by the vapour-diffusion method from a solution containing 5 mg ml^{-1} protein, 0.5 mM DTT, 0.5 mM PMSF, 0.5 mM EDTA, 0.1 M calcium acetate, 0.05 mM Tris-HCl pH 8.5, and 10% PEG 8,000 in a drop equilibrated against 0.2 M calcium acetate, 0.1 M Tris-HCl pH 8.5, and 20% PEG 8,000. The seleno-methionine mutant crystals were grown under similar conditions. The crystals are in space group C2, with unit-cell dimensions a = 108.8 Å, b = 30.7 Å, c = 51.3 Å, β = 91.7° for data collected at 277 K and a = 103.8 Å, b = 30.55 Å, c = 51.4 Å, β = 94.9° at 100 K.

The plasmids encoding the Gal4 DNA-binding domain fusion with Ha-ras, pGB-H-rasQ61L-ΔCVLS, and the Gal4 activation domain fusion with the RalGDS-RID, pGAD-ralGDS98, have been described[22]. The plasmid encoding the Gal4 DNA-binding domain fusion with Rap1A was constructed by cloning a PCR-amplified fragment encoding the first 182 amino acids of Rap1A-Q63E into the Gal4 DNA-binding domain vector pGBT9[37], resulting in plasmid pGB-rap1A-Q63E-ΔCLLL. The alanine substitution mutants of RalGDS-RID were created by site-directed mutagenesis using the PCR with pGAD-ralGDS98 as a template. Mutagenesis was performed by a two-step PCR amplification[38] using the appropriate mutagenesis primers and the same two outside primers that had been used to subclone pGAD-ral-GDS98. In some cases mutations were introduced by a single PCR using the appropriate mutagenesis primer spanning the unique NcoI site at codon 813 of rat RalGDS. All constructs were verified by sequencing.

Two-hybrid assays were performed essentially as described[22]. Briefly, the yeast reporter strain Y153 was transformed with Gal4 DNA-binding domain and activation domain constructs using the lithium acetate procedure[39] and were plated on selective medium lacking leucine and tryptophane to select for both plasmids. The colonies were screened after incubation for three days at 30 °C. Expression of β-galactosidase was measured with a filter assay using the chromogenic substrate X-Gal (5-bromo-4-chloro-2-indolyl β-D-galactoside). Filters were incubated in staining solution at 30 °C and colonies remaining white after 24 h were scored negative.

All collected X-ray diffraction data sets were processed with DENZO[40] and reduced with SCALEPACK[40]. We used the MADSYS program package (algebraic formalism) developed by Hendrickson and coworkers to obtain electron density maps[30]. MAD data sets were used for phase calculation. The IFAI values (structure factors of normal scattering from the anomalous scatters) of the MAD data were used in a Patterson synthesis to determine the position of the anomalous scatters (selenium positions). Four selenium sites were found in one asymmetric unit. Since RalGDS-RID only contains three methionines in a monomer, this finding suggests that there is a dimer in one asymmetric unit, and that two out of three methionines are fully substituted by selenomethionines and not disordered. The third methionine is located at the surface and its side chain B-factor is higher than that of the other two. 30% of the unit cell volume is estimated to be solvent. From the position of the four selenium atoms and their distance to each other, we determined the non-crystallographic symmetry between the two monomers.

The initial MAD-derived phases to 3.0 Å Bragg spacings were subsequently improved through solvent flattening, non-crystallographic symmetry averaging (two-fold), and multiple cycles of model building and refinement. However, when MAD data sets were collected, there was ice trapped in the cold stream nozzle, and X-ray diffraction from the ice degraded part of the data, resulting in a noisy MAD data set. Thus the final model was refined against 2.4 Å native data taken at 100 K at Brookhaven beam line X12B. Positional refinement, simulated annealing refinement, and temperature refinement were performed with the program X-PLOR 3.1[41]. Because of the weak electron densities in the N-terminal region (N-terminal 13 residues corresponding to RalGDS 767–777, or RID 3–13, plus two linker residues Gly and Ser), we did not construct a model for this region. The electron densities are also weak for two loops (residues 25–29 and 56–58) in one monomer; and only one loop (residues 25–29) in the other. No non-crystallographic symmetry (NCS) constraints or restraints were used in the final refinement. During refinement, the free R-factor[42] was monitored using 10% of total reflection as a test data set. After the free R-factor was reduced below 32%, 58 water molecules were identified from well defined electron densities in the F_o - F_c map. The structure has been refined to 2.4 Å resolution with a crystallographic R-value of 21.2% and a free R-value of 29.8%. All backbone dihedral angles fall within allowed regions in a Ramachandran plot. Coordinates (ID code: 1LXD) and structure factors have been deposited into the Brookhaven PDB data bank.

Lan Huang[1], Xiangwei Weng[1], Franz Hofer[2], G. Steven Martin[2] and Sung-Hou Kim[1]

[1]Department of Chemistry and Structural Biology Division of the E. O. Lawrence Berkeley National Laboratory, University of California, Berkeley, California 94720, USA. [2]Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA.

Correspondence should be addressed to S.-H. K. email: shkim@lbl.gov

correspondence

1. Barbacid, M. *Ann. Rev. Biochem.* **56**, 779–827 (1987).
2. Hall, A. *Curr. Opin. Cell Biol.* **5**, 265–268 (1993).
3. McCormick, F. *Mol. Reproduction Devel.* **42**, 500–506 (1995).
4. Feig, L.A., Urano, T. & Cantor, S. *Trends Biochem. Sci.* **21**, 438–441 (1996).
5. Feig, L.A. & Emkey, R. In *The ras Superfamily of GTPases* (eds Lacal, J.C. & McCormick, F.) 247–257 (CRC Press, Boca Raton, Florida; 1993).
6. Jiang, H., *et al. Nature* **378**, 409–412 (1995).
7. Katz, M.E. & McCormick, F. *Curr. Opin. Genet. Dev.* **7**, 75–79 (1997).
8. Boguski, M.S. & McCormick, F. *Nature* **366**, 643–654 (1993).
9. Kim, S.-H., Privé, G.G. & Milburn, M.V. In *Handbook of Experimental Pharmacology, Vol 108II: GTPase in Biology* (eds Dicky, B.F. & Birnbaumer, L.) 177–194 (Springer-Verlag, Berlin, Germany; 1993).
10. Lowy, D.R. & Willumsen, B.M. *Ann. Rev. Biochem.* **62**, 851–891 (1993).
11. Moodie, S.A., *et al. Oncogene* **11**, 447–454 (1995).
12. Van Aelst, L., Barr, M., Marcus, S., Polverino, A. & Wigler, M. *Proc. Natl. Acad. Sci. USA* **90**, 6213–6217 (1993).
13. Vojtek, A.B., Hollenberg, S.M. & Cooper, J.A. *Cell* **74**, 205–214 (1993).
14. Herrmann, C., Horn, G., Spaargaren, M. & Wittinghofer, A. *J Biol Chem* **271**, 6794–6800 (1996).
15. Leevers, S.J., Paterson, H.F. & Marshall, C.J. *Nature* **369**, 411–414 (1994).
16. Stokoe, D., Macdonald, S.G., Cadwallader, K., Symons, M. & Hancock, J.F. *Science* **264**, 1463–1467 (1994).
17. Macdonald, S.G., *et al. Mol. Cell. Biol.* **13**, 6615-6620 (1993).
18. Dent, P., *et al. Science* **257**, 1404–1407 (1992).
19. Hill, C.S., *et al. Cell* **73**, 395–406 (1993).
20. Marais, R., Wynne, J. & Treisman, R. *Cell* **73**, 381–393 (1993).
21. Albright, C.F., Giddings, B.W., Liu, J., Vito, M. & Weinberg, R.A. *EMBO Journal* **12**, 339–347 (1993).
22. Hofer, F., Fields, S., Schneider, C. & Martin, G.S. *Proc. Natl. Acad. Sci. USA* **91**, 11089–11093 (1994).
23. Spaargaren, M. & Bischoff, J.R. *Proc. Natl. Acad. Sci. USA* **91**, 12609–12613 (1994).
24. Kikuchi, A., Demo, S.D., Ye, Z.H., Chen, Y.W. & Williams, L.T. *Mol. Cell. Biol.* **14**, 7483–7491 (1994).
25. White, M.A., Vale, T., Camonis, J.H., Schaefer, E. & Wigler, M.H. *J. Biol. Chem.* **271**, 16439–16442 (1996).
26. Ponting, C.P. & Benjamin, D.R. *Trends Biochem. Sci.* **21**, 422–425 (1996).
27. White, M.A., *et al. Cell* **80**, 533–541 (1995).
28. Nassar, N., *et al.* The 2.2 *Nature* **375**, 554–560 (1995).
29. Huang, L., Jancarik, J., Kim, S.H., Hofer, F. & Martin, G.S. *Acta Crystallographica* **D52**, 1033–1035 (1996).
30. Hendrickson, W.A. *Science* **254**, 51–58 (1991).
31. Emerson, S.D., *et al. Biochemistry* **34**, 6911–6918 (1995).
32. Vijay-Kumar, S., Bugg, C.E. & Cook, W.J. *J Mol. Biol.* **194**, 531–544 (1987).
33. Jacobson, B.L., Chae, Y.K., Markley, J.L., Rayment, I. & Holden, H.M. *Biochemistry* **32**, 6788–6793 (1993).
34. Derrick, J.P. & Wigley, D.B. *Nature* **359**, 752–754 (1992).
35. Block, C., Janknecht, R., Herrmann, C., Nassar, N. & Wittinghofer, A. *Nature Struct. Biol.* **3**, 244–251 (1996).
36. Nassar, N., *et al. Nature Struct. Biol.* **3**, 723–729 (1996).
37. Bartel, P., Chien, C.-T., Sternglanz, R. & Fields, S. In *Cellular Interactions in Development: a Practical Approach* (ed. Hartley, D.A.) 153–179 (Oxford University Press, Oxford; 1993).
38. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. & Pease, L.R. *Gene* **77**, 51–59 (1989).
39. Schiestl, R.H. & Gietz, R.D. *Curr. Genet.* **16**, 339–346 (1989).
40. Otwinowski, Z. In *Data Collection and Processing* (eds Swayer, L., Issacs, N. and Bailey, S.) 56–62 (SERC Daresbury Laboratory, Warrington, England; 1993).
41. Brunger, A.T. *X-PLOR, a System for X-Ray Crystallography and NMR* (Yale University Press, New Haven, Connecticut; 1993).
42. Brunger, A.T. *Nature* **355**, 472–475 (1992).
43. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. *Acta Crystallogr.* **A42**, 110–119 (1991).
44. Kabsch, W. & Sander, C. *Biopolymers* **22**, 2577–2637 (1983).
45. Kraulis, P.J. MOLSCRIPT: *J. Appl. Crystallogr.* **24**, 946–950 (1991).
46. Milburn, M.V., *et al. Science* **247**, 939–945 (1990).

Crystal structure of a trapped phosphoenzyme during a catalytic reaction

The crystal structure of the fructose-2,6-bisphosphatase domain trapped during the reaction reveal a phosphorylated His 258, and a water molecule immobilized by the product, fructose-6-phosphate. The geometry suggests that the dephosphorylation step requires prior removal of the product for an 'associative in-line' phosphoryl transfer to the catalytic water.

The reciprocal regulation of the two opposite carbohydrate metabolic pathways, glycolysis and gluconeogenesis by hepatic fructose-2,6-bisphosphate (Fru-2,6-P_2) is the major mechanism for the control of energy storage and blood glucose levels in mammalian systems[1–3]. The cellular concentration of Fru-2,6-P_2 is regulated by glucagon and insulin through phosphorylation and dephosphorylation of the bifunctional enzyme, 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (6-PF-2-K/Fru-2,6-P_2ase). The bifunctional enzyme consists of the kinase domain at residues 1–250 and the bisphosphatase domain at residues 251–470[1–3]. The kinase domain catalyzes the formation of the β-anomer of Fru-2,6-P_2 using ATP and Fru-6-P as substrates[2,4]. The C-terminal bisphosphatase domain catalyzes the hydrolysis of the sugar bisphosphate into Fru-6-P and inorganic phosphate[1,2,10]. Under hypoglycemic conditions, glucagon increases the bisphosphatase activity through cAMP-dependent phosphorylation at Ser 32, resulting in elevated gluco-

neogenic flux. To increase glycolytic fluxes, the bisphosphatase activity is decreased by insulin-dependent dephosphorylation of Ser 32, resulting in the bifunctional enzyme with predominantly kinase activity[1–3,5,6]. Thus, manipulation of activity of either or both domains would have major metabolic consequences and possible therapeutic application in diabetes for restoration of euglycemia.

Recently, the X-ray crystal structures of the rat liver bisphosphatase domain[7] and the rat testis bifunctional enzyme[8] have been determined. The overall conformation and geometry of the separate bisphosphatase domain were nearly identical to that of the bifunctional enzyme[8], consistent with the kinetic studies which suggested that the separate domain could be used as a powerful tool for structure/function studies[9–11].

Like other members of the histidine phosphatase family, Fru-2,6-P_2ase carries out a phosphoryl transfer reaction using histidine as the key catalytic residue[12,13]. The bisphosphatase forms a phosphoen-

zyme intermediate (E–P) in which His 258 is transiently phosphorylated by the 2-P of the substrate[1,2]. Transient kinetic studies of the catalytic reaction of the bisphosphatase activity have shown that the catalytic E–P intermediate is formed in the millisecond time range[14], and ^{31}P-NMR studies have shown it is stable for up to eight hours at pH 8.0 at concentrations of 3.3 mM protein and 26.4 mM substrate[11].

The E–P intermediate

The E–P intermediate was trapped by incubating the bisphosphatase crystal in substrate and then flash-freezing. The presence of the intermediate phospho-amide bond (N–P) between the 3-N of His 258 and the phosphorous is clearly visible in the electron density omit maps (Fig. 1). The map of the intermediate was calculated using only the protein coordinates (R_{free}/R_{crys} = 0.312/0.236) which were refined from the native structure (see Methods). A refined bond length of 1.90 Å (Figs 2, 3a) was comparable to both a calculated value of 1.76 Å (ref. 15) and a