The structural basis of the activation of Ras by Sos

P. Ann Boriack-Sjodin*, S. Mariana Margarit†, Dafna Bar-Sagi† & John Kuriyan*‡

** Laboratories of Molecular Biophysics, and ‡ Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10021, USA*
† Department of Molecular Genetics and Microbiology, State University of New York at Stony Brook, Stony Brook, New York 11794, USA

The crystal structure of human H-Ras complexed with the Ras guanine-nucleotide-exchange-factor region of the Son of sevenless (Sos) protein has been determined at 2.8 Å resolution. The normally tight interaction of nucleotides with Ras is disrupted by Sos in two ways. First, the insertion into Ras of an α-helix from Sos results in the displacement of the Switch 1 region of Ras, opening up the nucleotide-binding site. Second, side chains presented by this helix and by a distorted conformation of the Switch 2 region of Ras alter the chemical environment of the binding site for the phosphate groups of the nucleotide and the associated magnesium ion, so that their binding is no longer favoured. Sos does not impede the binding sites for the base and the ribose of GTP or GDP, so the Ras–Sos complex adopts a structure that allows nucleotide release and rebinding.

The Ras proteins are highly conserved guanine-nucleotide-binding enzymes that couple cell-surface receptors to intracellular signalling pathways controlling cell proliferation and differentiation[1,2]. Ras acts as a molecular switch by cycling between active GTP-bound and inactive GDP-bound states. Whether GDP or GTP is bound to Ras is determined by the action of two classes of regulatory proteins: guanine-nucleotide-exchange factors, and GTPase-activating proteins. Exchange factors promote the activation of Ras by catalysing the exchange of GDP for GTP, whereas GTPase-activating proteins control the conversion of Ras to the inactive state by stimulating the hydrolysis of GTP to GDP[2].

Cell-surface receptors that signal through tyrosine kinases activate Ras by stimulating the guanine-nucleotide exchange reaction[3–5]. Genetic and biochemical studies have indicated that this reaction is controlled by the Ras guanine-nucleotide exchange factor Son of sevenless (Sos)[6]. Following ligand binding, Sos is brought from the cytoplasm to the activated receptor in a phosphotyrosine-dependent manner through adapter proteins such as Grb2. Grb2 contains SH3 domains that are bound constitutively to a carboxy-terminal proline-rich region of Sos, and the Grb2–Sos complex is recruited to activated receptors by interactions between the SH2 domain of Grb2 and phosphotyrosine residues on the receptor[7]. Because Ras is localized to the membrane, receptor activation results in an increase in the effective concentration of Sos in the vicinity of Ras, facilitating the exchange of bound guanine nucleotide for free cellular guanine nucleotides.

The region of Sos (relative molecular mass of about 150,000; M_r ~150K) that is functional for nucleotide exchange on Ras spans

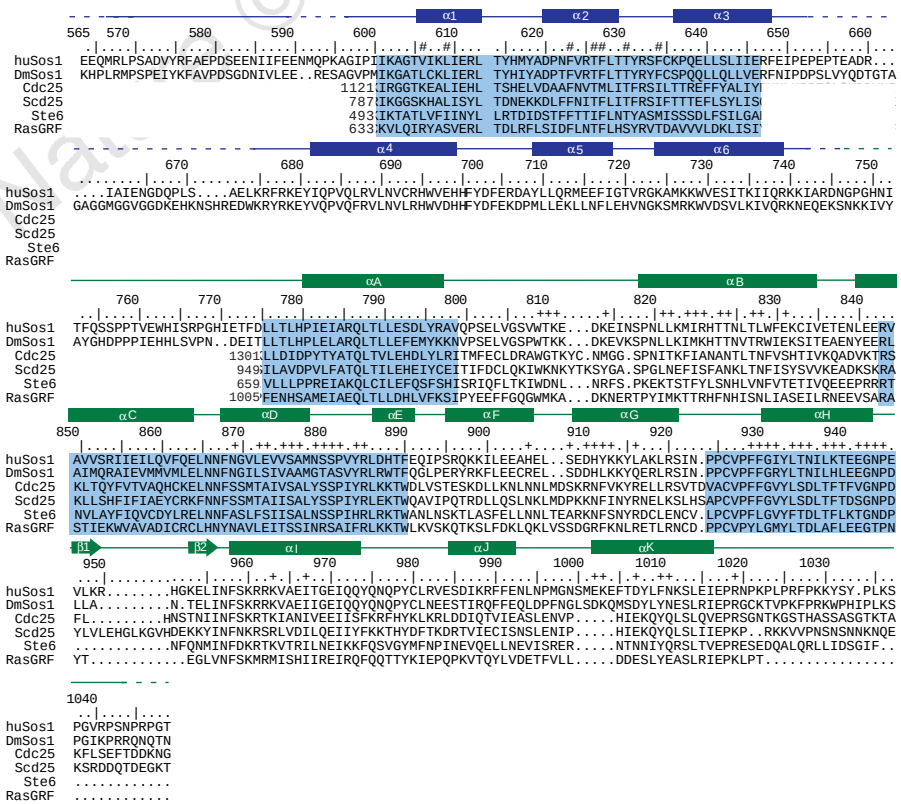

Figure 1 Sequence alignment of Ras-binding exchange factors with residue numbers of human (hu) Sos1 indicated. The secondary structure elements (α-helices are shown as rectangles, β-sheets as arrows, coil regions as a solid line, disordered residues as broken lines) of the N-domain are dark blue, and those of the catalytic domain are green. Conserved regions are shaded with pale blue. Residues of Sos at the Ras interface are indicated with +; residues in the N-domain that form the hydrophobic core with the catalytic domain are indicated with #. Sequences have been omitted where the sequence similarity between exchange factors is low. DmSos1, *Drosophila melanogaster* Sos1.

Table 1 Data collection, structure determination and refinement statistics

Data set	Resolution (Å)	Observations (total/unique)	R_{sym}* (%)	Completeness (%)	R_{iso}† (%)	No. of sites	Phasing power‡
Native	2.8	283,114/26,562	5.3 (28.7)	99.3 (99.9)			
PCMB§	3.1	616,299/19,766	7.7 (28.9)	93.4 (87.9)	22.8	5	1.68
Baker's dimercurial§	3.25	640,979/16,988	7.6 (19.3)	96.9 (98.6)	21.5	6	0.80
EMTS§	3.25	634,455/17,174	7.3 (35.1)	95.4 (97.2)	20.0	4	1.94
Trimethyl lead acetate§	2.9	523,356/23,842	5.9 (27.0)	98.2 (99.3)	22.5	6	0.77
Selenomethionine	3.0	543,633/21,735	6.5 (29.7)	96.1 (93.2)	15.4	10	1.44
Gold cyanide§	3.05	368,880/20,927	5.4 (28.6)	98.3 (96.5)	33.5	3	1.46
Plantinum terpyridine	3.25	240,941/16,915	10.6 (36.7)	98.9 (99.3)	18.5	6	1.35
Osmium chloride	3.1	202,555/20,025	6.8 (29.4)	99.6 (99.9)	23.5	4	1.55
PCMB/trimethyl lead acetate§	3.3	571,428/16,485	7.8 (30.0)	99.8 (99.5)	26.6	9	1.91
Overall figure of merit‖ = 0.70							

Refinement statistics

Data set	Resolution (Å)	Reflections	Total atoms	R-factor/R_{free}¶	r.m.s.d. from ideal values	
					bonds (Å)	angles (°)
Native	30–2.8	26,502	5,010	0.222/0.281	0.007	1.26

* $R_{sym} = 100 \times \Sigma|I - \langle I \rangle|/\Sigma I$, where I is the integrated intensity of a given reflection. For R_{sym} and completeness, numbers in parentheses refer to data in the highest resoution shell.
† $R_{iso} = 100 \times \Sigma|F_{PH} - F_P|/\Sigma F_P$, where F_{PH} and F_P are the derivative and native structure factor amplitudes, respectively.
‡ Phasing power $= \Sigma|F_{PH(calc)}|^2/\Sigma\{F_{PH(obs)} - F_{P(calc)}|^2\}^{1/2}$.
§ Anomalous data were used in the phasing of these derivatives.
‖ Figure of merit $= \langle|\Sigma P(\alpha)e^{i\alpha}/\Sigma P(\alpha)|\rangle$, where α is the phase and $P(\alpha)$ is the phase probability distribution.
¶ R-factor $= \Sigma||F_P| - |F_{P(calc)}|/\Sigma|F_P|$; R_{free} was calculated with 5% of the data.

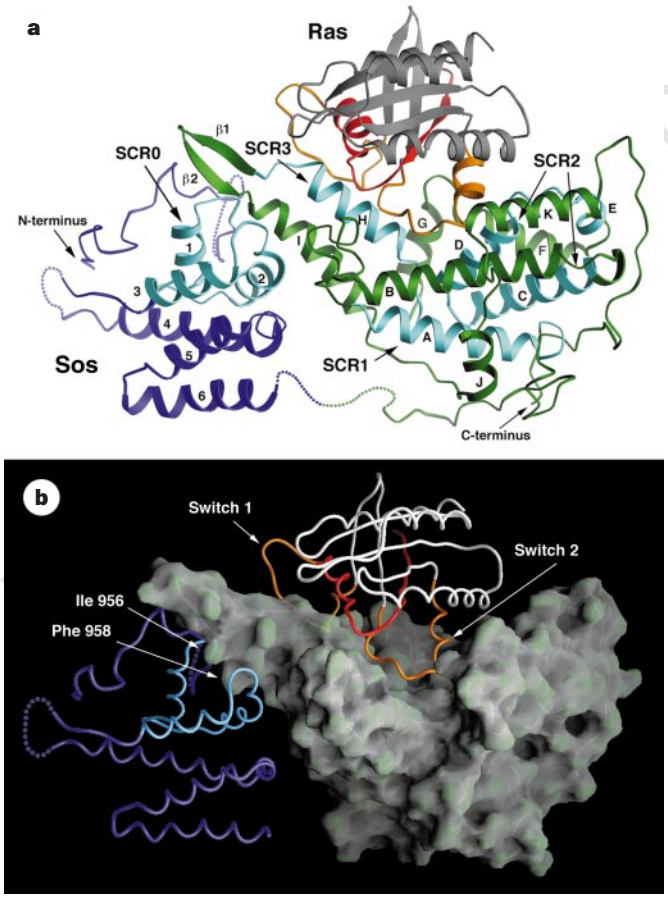

Figure 2 The complex of human H-Ras with the exchange-factor region of human Sos1. **a**, The N-domain of Sos (residues 568–741) is shown blue; the catalytic domain (residues 752–1044) is green; the Switch 1 and Switch 2 segments and the P-loop region of Ras (as defined here) are orange and red, respectively; conserved regions (SCRs) among Ras-family exchange factors are cyan[2,17]. Disordered residues of Sos are shown as dotted lines. This and all other ribbon diagrams were generated using RIBBONS[44]. **b**, The Ras–Sos complex is shown with the catalytic domain of Sos depicted as a molecular surface. Conserved residues Ile 956 and Phe 958 in the catalytic domain that form a hydrophobic interface with the N-domain are labelled. This and all other figures with molecular surfaces were generated using GRASP[45].

about 500 residues, and contains blocks of sequence that are conserved in other Ras-specific nucleotide exchange factors such as Cdc25, Sdc25 and Ras guanine-nucleotide-release factor (GRF)[2] (Fig. 1). However, in contrast to the high degree of structural conservation seen in the GTPases of the Ras family and others[1], there are distinct families of nucleotide exchange factors for the various guanine-nucleotide-binding protein (G protein) families that are structurally unrelated to each other[8–13]. The structure of only one nucleotide exchange factor bound to its cognate G protein has yet been determined: that of ribosomal EF-Tu bound to its exchange factor EF-Ts[12,13].

To clarify the molecular mechanism of the activation of Ras by Sos, we have determined the crystal structure of the complex of a C-terminally truncated form of human Harvey-Ras (residues 1–166, hereafter referred to as Ras) with the guanine-nucleotide-exchange-factor region of human Sos1 (residues 564–1049, hereafter referred to as Sos). The structure reveals that Sos interacts extensively with Ras and stabilizes it in a nucleotide-free state by displacing the residues that coordinate the magnesium ion and the phosphate groups of the nucleotide and by partly occluding the magnesium-binding site. The structure also suggests a pathway for the rebinding of nucleotides to Ras, with consequent release of Sos, a process that is crucial for Sos to function as a nucleotide exchanger rather than a binding inhibitor.

General features of the structure

Ras and Sos form a tight complex in the absence of nucleotides. Crystals of this complex were obtained with one Ras–Sos complex (M_r 75K) in the asymmetric unit. The structure was determined by multiple isomorphous replacement and the molecular model was refined against data to 2.8 Å, resulting in a crystallographic R-value of 22.2% (free R-value of 28.1%; Table 1). The final model includes 439 residues of Sos (spanning residues 568–1044), 166 residues of Ras and 26 water molecules. There is no nucleotide or magnesium present in the crystals. The structure of Ras resembles that of the nucleotide-bound forms described previously[19,46], except for localized changes in the vicinity of the nucleotide-binding site. The structure of Sos, described below, seems to be unrelated to that of any other protein of known structure, based on a DALI[14] search of the protein data bank.

The structure of the Ras exchange-factor region of Sos consists of two distinct α-helical structural domains (Fig. 2). The amino-terminal domain (N-domain; residues 568–741, α-helices α1–α6) does not interact with Ras, and seems to have a purely structural

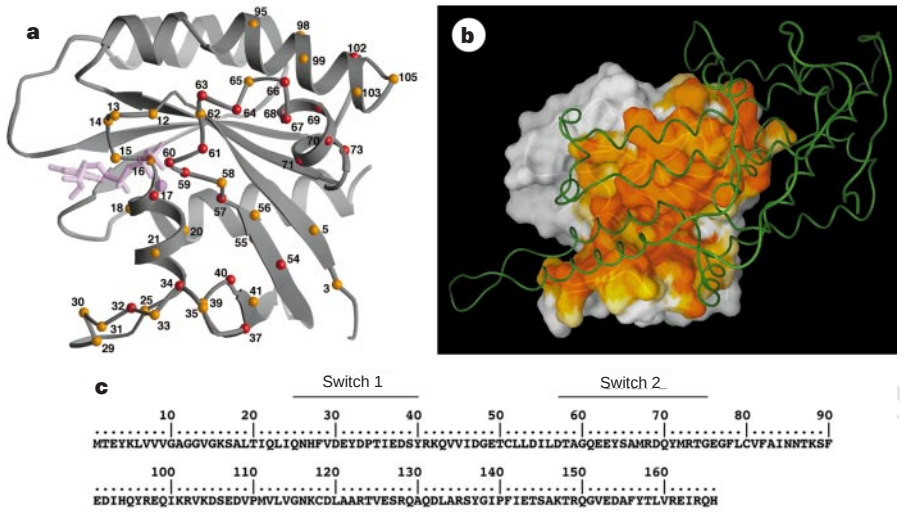

Figure 3 Interface surfaces of the Ras–Sos complex. **a**, Residues of Ras that form direct interactions with Sos are shown as red spheres; additional residues at the interface are orange spheres. The nucleotide is shown for reference only and is not present in the structure. **b**, The interface surface of Ras; the orientation is the same as in **a**. The surface is coloured using a gradient: bright orange indicates atoms <4 Å from Sos, white indicates atoms >7 Å from Sos, lighter shades of orange indicate intermediate distances. Sos (N-domain deleted) is shown as a green ribbon. **c**, The primary sequence of Ras.

role. The C-terminal domain (residues 752–1044, α-helices αA–αK) contains within it all the residues that interact with Ras, and this region will be referred to as the catalytic domain. Analyses of the exchange-factor activity of Cdc25 and Sos have shown that the relatively well-conserved C-terminal catalytic domain suffices for catalytic activity[15,16]. However, recombinant fragments of Sos that span the catalytic domain, but which lack some or all of the N-domain, are expressed poorly and have low solubility (data not shown), whereas the fragment used for the structure determination includes both domains and results in relatively high yields of soluble protein.

The structure of the catalytic domain consists of a series of helical hairpins that pack against each other (Fig. 2). A notable feature of the catalytic domain is the protrusion of a helical hairpin, formed by helices αH and αI, out of the core of the domain. Helix αH plays an important role in the nucleotide-exchange mechanism, and the main structural role for the N-domain appears to be the stabilization of the hairpin that presents this helix to Ras. Helices α1 and α2 of the N-domain together form a small hydrophobic groove into which two conserved hydrophobic side chains from helix αI of the catalytic domain are inserted (Ile 956 and Phe 958; Fig. 2b). The packing of these side chains into the α1–α2 groove, together with several adjacent interdomain interactions, is likely to be important for the stability and correct placement of the hairpin structure (Fig. 2).

The structure of the N- and catalytic domains of Sos is likely to be a good model for the general architecture of related guanine-nucleotide-exchange factors, such as Cdc25, Sdc25 and RasGRF (Fig. 1). Three regions of sequence conservation (structurally conserved regions, or SCRs) within the catalytic domain had been identified previously[2], and these are important either for the structural integrity of the domain (SCR1, helix αA and SCR2, helix αC) or for the interaction with Ras (SCR2, helix αD and SCR3; Fig. 2a). The region of the N-domain spanning helices α1, α2 and α3 is highly conserved among Ras-specific nucleotide exchange factors (SCR0 in Fig. 2a)[17]. The hydrophobic nature of the groove between helices α1 and α2 is conserved, as are the residues on the catalytic domain that interact with the groove and the adjoining surface of the N-domain, suggesting that the interaction between the N-domain and the catalytic domain is conserved.

Structure of the Ras–Sos interface

The overall shape of the catalytic domain of Sos is that of an oblong bowl (Fig. 2), with Ras bound at the centre of the bowl. The regions of Ras that interact most closely with Sos include the phosphate-binding P-loop (residues 10–17) and surrounding segments (including strand β1 and helix α1), the Switch 1 region (defined here as residues 25–40) and the Switch 2 region (defined here as residues 57–75). Additional interactions are seen with helix α3 (residues 95–105; Fig. 3a, b). The interface between Ras and Sos is primarily hydrophilic and very extensive, with 3,600 Å^2 of surface area buried in the complex. At the heart of the interface between Ras and Sos is a cluster of three hydrophobic side chains from the Switch 2 region of Ras (Tyr 64, Met 67 and Tyr 71) which are buried into the hydrophobic core of Sos at the base of the binding site. Surrounding this hydrophobic anchor is an array of polar and charged interactions between Sos and Ras that results in almost every external side chain of Switch 2 being coordinated by Sos (Fig. 4).

The most obvious effect of Sos binding to Ras is the opening of the nucleotide binding site as a result of the displacement of Switch 1 of Ras by the insertion of the helical hairpin formed by αH and αI of Sos (Fig. 5). An important consequence of the insertion of helix αH into the Ras active site is that it introduces a hydrophobic side chain (Leu 938), which blocks magnesium binding, and an acidic side chain (Glu 942), which overlaps the site where the α-phosphate of the nucleotide would otherwise be bound.

The Switch 2 region is held very tightly by Sos. Although this region tends to be poorly ordered in the nucleotide-bound forms of Ras[19,46], its conformation is very well defined in the Sos complex. The temperature factors for all atoms of Switch 2 are low (34 Å^2, compared with the average value of 49 Å^2 for all of Ras). The C-terminal end of Switch 2 is farthest from the nucleotide binding site, and side chains in this region interact with Sos but do not change their positions significantly. However, closer to the nucleotide-binding site the interactions between Sos and the side chains of Switch 2 result in a restructuring of the polypeptide backbone connecting β3 and α2 (Fig. 4a, b). This restructuring of the backbone is a crucial determinant of nucleotide exclusion, as it prevents the coordination of magnesium and phosphate (see below).

Switch 1 and Switch 2 are the only regions of Ras in which structural changes are directly induced by Sos. Comparison of the

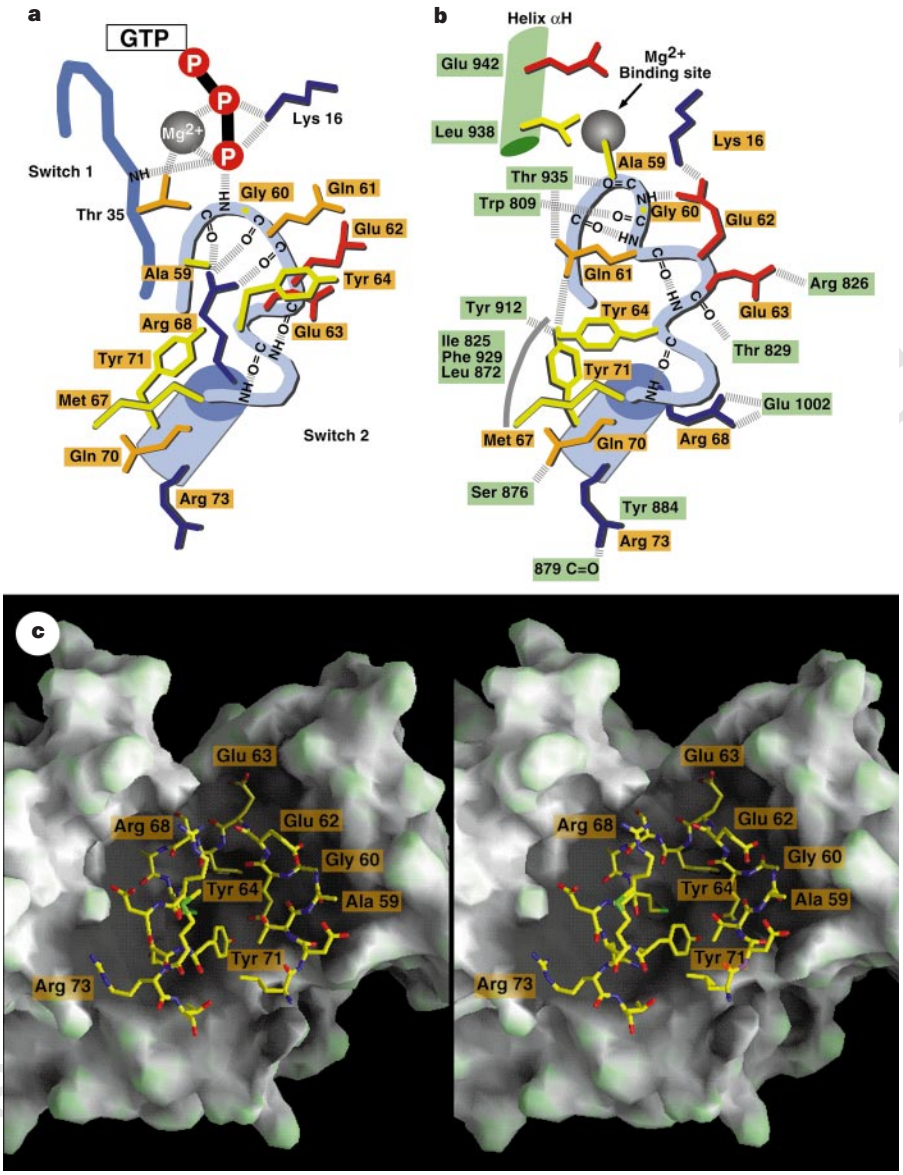

Figure 4 A schematic representation of the differences in the Switch 2 regions of a Ras-GTP analogue and Ras-Sos. **a**, Ras-GTP analogue (5P21)[39]. **b**, Ras-Sos. Selected polar interactions are shown as broken lines, hydrophobic interactions as solid arcs. **c**, A stereodiagram of the Switch 2 region of Ras superimposed on the surface of Sos. Selected residues of Ras are indicated.

structure of Ras in the Ras–Sos complex with that of the nucleotide-bound forms shows that there are also structural changes in one of the loops that binds the nucleotide base (loop 8, between β5 and α5, residues 118–123) and the phosphate-binding P-loop. However, the changes in these loops seem to be a simple consequence of the absence of nucleotide.

Structural changes in the Switch 1 region

The change in the Switch 1 region of Ras when bound to Sos is drastic (Fig. 6). Switch 1 normally rises up from the end of α1 towards the P-loop region, to sandwich the nucleotide between it and the rest of Ras. Strand β2 in Switch 1 forms an antiparallel interaction with strand β3 in the central β-sheet of Ras. In the Sos complex, the C-terminal region of helix α1 is shortened by about one helical turn, the antiparallel β-sheet interaction with β3 is completely disrupted, strand β2 is partly melted, and Switch 1 is completely removed from the nucleotide-binding site.

One important aspect of the insertion of the helical hairpin of Sos into the Switch 1 region is that it does not result in a significant occlusion of the guanine and ribose binding sites (Fig. 5d). Instead, this structural distortion breaks the network of direct and water-mediated interactions between Switch 1 and the nucleotide. For example, in the nucleotide-bound forms of Ras, Phe 28 interacts with the guanine base through a perpendicular aromatic–aromatic interaction (Fig. 5a). Mutation of Phe 28 to leucine results in a significant increase in the intrinsic rate of dissociation of nucleotide from Ras[18]. In the Sos complex, the Cα of Phe 28 moves 9.6 Å and the side chain no longer interacts with the nucleotide-binding site (Fig. 5b).

Two side chains presented by helix αH of Sos, Leu 938 and Glu 942, directly impede the binding of magnesium and phosphate, respectively. The carboxylate group of Glu 942 of Sos is positioned near the location of the α-phosphate of GTP or GDP in the nucleotide-bound forms of Ras. In the Sos complex, Glu 942 forms a hydrogen bond with Ser 17 of Ras, a ligand of the magnesium ion in the nucleotide complexes of Ras, preventing the binding of the phosphates as well as the magnesium ion. Leu 938 further increases the hydrophobicity of the magnesium binding site

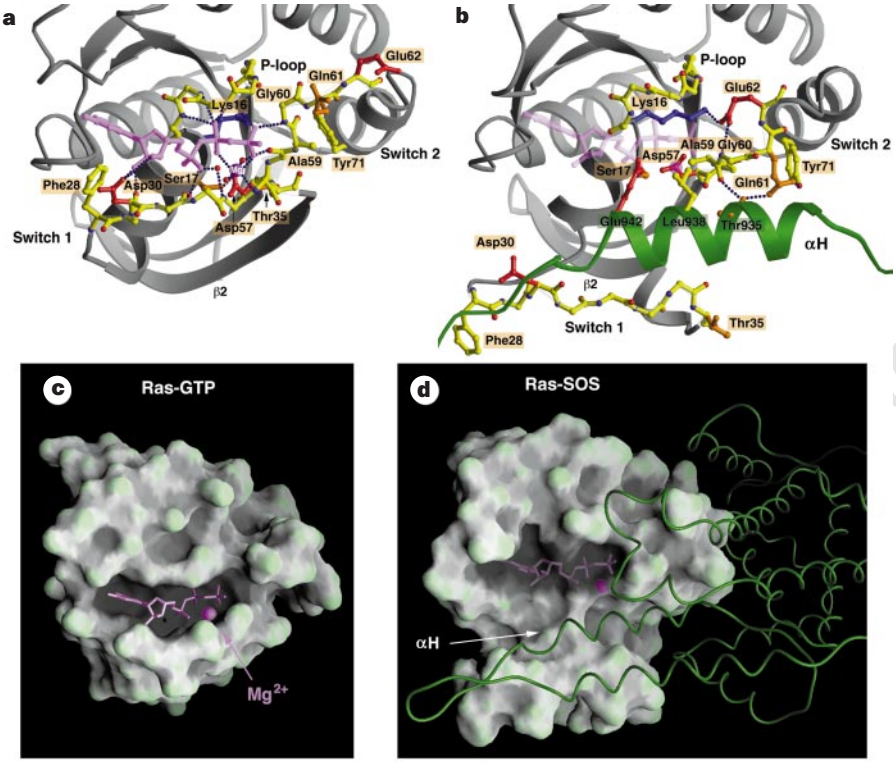

Figure 5 Interactions at the nucleotide-binding site. **a**, Selected interactions between Ras and a GTP analogue (5P21)[39]. Water molecules are depicted as red spheres. **b**, Selected interactions between Ras and Sos are shown in the same orientation as in **a**. Only helix αH and selected side chains of Sos are shown. **c**, The nucleotide-binding site on the surface of Ras in a Ras-GTP analogue (5P21)[39] is shown. The side chain of Tyr 32 was deleted from the surface calculations for clarity. **d**, The surface of Ras in the Ras–Sos complex is shown with the backbone of Sos (N-domain deleted) as a green ribbon. Ras is in a slightly different orientation from that in **c**. In **b** and **d**, the nucleotide is shown for reference only.

in the Sos complex, which is also occupied by Ala 59 from Switch 2 (see below).

Changes in the Switch 2 region

The sensitivity of Switch 2 to the presence of either GTP or GDP is a consequence of the coordination of the terminal phosphate of GTP by the backbone amide nitrogen of Gly 60 (Figs 4a and 5a; refs 19, 46). In the following discussion, we compare the conformation of the Switch 2 region in the Ras–Sos complex to that seen in the GTP-bound form of Ras, because the Switch 2 region is not well ordered in the GDP-bound form[19]. The Switch 2 region of Ras makes important interactions with GTP and not with GDP[19,46]. Nevertheless, structural changes that are induced in Switch 2 by Sos result in the exclusion of both GDP and GTP, because they affect magnesium binding as well as the conformation of Lys 16 in the P-loop, a crucial phosphate ligand.

Three specific features of the Switch 2 conformation are correlated with the disruption of nucleotide binding. First, the backbone conformation in the central region of Switch 2 is compressed in the Ras–Sos complex because of the formation of three consecutive β-turns (Fig. 4a, b). As a consequence of the first β-turn, the methyl side chain of Ala 59 is turned in towards the phosphate-binding site and occludes the position that would be occupied by the Mg^{2+} ion in nucleotide complexes. Second, the formation of the second β-turn results in the side chain of Glu 62 coordinating both the amide nitrogen of Gly 60 and the side chain of Lys 16, two residues involved in phosphate coordination (Figs 4 and 5). Finally, in the GTP-bound form, the polypeptide backbone of the β3–α2 loop adopts a conformation in which three carbonyl groups (that of 59, 60 and 61) are pointed inwards. The resulting anion hole coordinates the side chain of Arg 68, and positions the methyl group of Ala 59 away from the magnesium-binding site. In the Sos complex, Arg 68 is removed from this internal location by interactions with Glu 1002 of Sos, and the anion hole is disrupted by the formation of the first two β-turns in the structure (Fig. 4a, b).

Comparison with EF-Tu and EF-Ts

The structure of only one other GTPase complexed to its exchange factor is yet known, that of EF-Tu bound to EF-Ts. EF-Tu is a GTPase that contains a nucleotide-binding domain that is topologically similar to Ras[20], as well as two additional domains. However, the mechanism of nucleotide release by its exchange factor EF-Ts is quite different from that seen here for Ras and Sos. The complex between EF-Tu and EF-Ts, like the Ras–Sos complex,

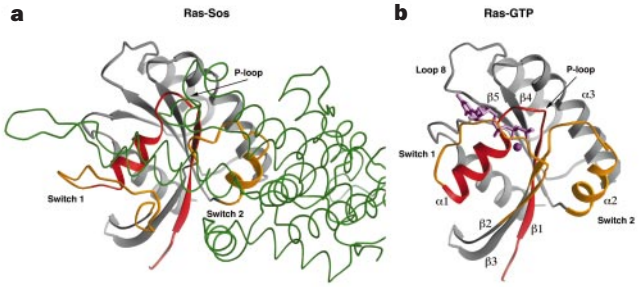

Figure 6 Comparison of Ras complexed with Sos and a GTP analogue. **a**, Sos; **b**, GTP analogue (PDB code 5P21)[39]. The colouring scheme for Ras is the same as in Fig. 2. Sos (N-domain deleted) is shown as a green ribbon. Secondary structure elements of Ras important in nucleotide binding and Sos binding are labelled.

stimulates nucleotide release by disrupting the interactions of the phosphate groups of the nucleotide, leaving the binding site for the base and ribose unimpeded[12,13]. Reorientation of a peptide bond in the phosphate-binding P-loop, induced by EF-Ts binding, results in the placement of a carbonyl oxygen in a position where it would collide with the β-phosphate of the nucleotide. In addition, global changes in the positions of the three domains result in conformational changes in the Switch 2 region that remove side chains that interact with the magnesium ion by means of water molecules. In contrast to Ras, the Switch 1 region of nucleotide-bound EF-Tu does not interact extensively with the nucleotide-binding site. Consequently, disruption of the Switch 1 structure does not seem to be a major component of the mechanism of EF-Ts action.

Implications for the reaction mechanism

Biochemical studies of Ras exchange factors have shown that the complex of Ras with these proteins is stable in the absence of nucleotides and is dissociated by the rebinding of either GDP or GTP[16–18,21,22]. The principal role for the exchange factor is to facilitate nucleotide release, and it does not seem to control significantly the preferential rebinding of GTP over GDP[16,22,23]. Cellular concentrations of GTP are ~10-fold higher than GDP, which results in the loading of GTP onto Ras.

The mechanism of nucleotide release by the catalytic domain of murine Cdc25 ($Cdc25^{Mm}$) has been investigated recently using fluorescently labelled nucleotides[16]. The affinity of $Cdc25^{Mm}$ for nucleotide-free Ras ($K_d = 4.6$ nM) is found to be several orders of magnitude higher than that for nucleotide-bound Ras, and the maximal acceleration by $Cdc25^{Mm}$ of the rate of dissociation of nucleotide is more than 10^5. Kinetic analysis of nucleotide association shows that the reaction proceeds by the formation of a ternary complex of a loosely bound nucleotide and Ras–$Cdc25^{Mm}$ followed by conversion to a form in which the nucleotide is tightly bound to Ras[16]. In light of the structure of the Ras–Sos complex, the first step can be interpreted as the interaction of the base and the ribose of the nucleotide with the part of the Ras binding site that is not occluded by Sos. The second step would involve a conformational change in the Switch 2 segment and release of Switch 1, resulting in the restructuring of a competent binding site for phosphate and magnesium, and the subsequent dissociation of Sos. The mechanism by which $Cdc25^{Mm}$ displaces the nucleotide does not depend solely on expulsion of the magnesium[16]. This is consistent with the Ras–Sos structure, as the mechanism involves interference with phosphate binding as well as displacement of the magnesium ion.

Analysis of Ras mutations

The structure of the Ras–Sos complex is consistent with several mutations in Ras that have highlighted the importance of the Switch 1 and Switch 2 regions in the interaction with nucleotide exchange factors[18,24–30]. The importance of helix α3 (residues 102–105) has also been noted[25,27,28]. The importance of Switch 2 for the recognition of the exchange factor is demonstrated by the analysis of mutations in residues that are not directly involved in nucleotide binding, but which affect GDP–GTP exchange. For example, mutation of Glu 62 and Glu 63 to histidine had no significant effect on the stability of the Ras–GDP complex[18]. However, both Ras mutants were severely compromised in their ability to be activated by Sdc25 (ref. 18). In the structure of the Ras–Sos complex, Glu 62 and 63 of Ras are both seen to be crucial to the interaction with Sos (Figs 4b and 5b).

Of particular interest are dominant-negative mutants of Ras that appear to act by binding to and sequestering nucleotide-exchange factors[31,32]. The most straightforward explanation of the action of these mutations is that they destabilize nucleotide binding[22,32,33], increasing the apparent affinity of Ras for Sos or other exchange factors. That the dominant-negative mutant Ras proteins act by binding to the exchange factor is also suggested by the fact that mutations at residues important for the formation of the Ras–Sos interface result in a reversion of the dominant-negative behaviour[26,29]. For example, in the yeast *Saccharomyces cerevisiae* protein Ras2, the mutation Ser 24 → Asn (corresponding to Ser 17 → Asn in human Ras) is dominant negative[32]. Substitution of Arg 80, Asn 81 (Arg 73, Thr 74 in human Ras) with Asp 80, Asp 81 in mutant (Ser 24 → Asn) Ras2 results in a loss of sensitivity to Sdc25 and reversion of the dominant-negative phenotype[29]. In the Ras–Sos complex, Arg 73 (Arg 80 in Ras2) is involved in interactions with two residues of Sos (Fig. 4b), and mutation to Asp would clearly be disruptive.

Discussion

As a nucleotide-exchange factor, Sos functions under two apparently conflicting imperatives. The interaction between Sos and Ras must be strong enough to dislodge the tightly bound nucleotide, but the Ras–Sos complex must also be poised for subsequent displacement by incoming nucleotides. The structure of the Ras–Sos complex shows that Ras and Sos meet these demands by forming a tight complex that is anchored at one end of the nucleotide-binding site, where phosphate and magnesium are normally bound. The interface between Sos and Ras is mainly hydrophilic, suggesting a ready unzippering through water-mediated displacements of the coordinating side chains. The main interacting elements of Sos avoid direct occlusion of the nucleotide-binding site, except the region where the terminal phosphate groups and the magnesium ion are bound. This feature allows incoming nucleotides to reverse the process by competing for the groups that ligate the phosphate and metal ion.

The proliferative status of cells is critically dependent on the activation state of Ras. The structure of the Ras–Sos complex suggests at least two avenues for the design of inhibitors to block the activation of Ras by Sos. Nucleotide analogues that are designed to recognize the altered nucleotide-binding site in the Ras–Sos complex may help to stabilize the complex, mimicking the action of dominant-negative alleles of Ras. Alternatively, hydrophobic compounds that bind to the core hydrophobic region at the heart of the Ras binding site on Sos may effectively inhibit Sos action. □

Methods

Expression and purification. *Escherichia coli* cells (BL21-DE3) were transformed with a pProEX HTb vector (Life Technologies) containing Ras (residues 1–166) linked to an N-terminal polyhistidine tag using the *Bam*HI and *Xho*I restriction sites. Protein production was induced with 250 μM IPTG at a cell density of absorbance $A_{600} = 0.5$. Protein was expressed at 30 °C for 6 h. Cells were collected by centrifugation, resuspended in 20 mM Tris, pH 8.0, 300 mM NaCl at 4 °C, flash frozen and stored at −80 °C until needed. Once thawed, cells were lysed using a French press (EmulsiFlex-C5, Avestin), cell debris was removed by centrifugation, and the resulting cell lysate loaded onto a charged nickel binding column (HisBind; Novagen) pre-equilibrated wit 20 mM Tris, pH 8.0, 500 mM NaCl and 20 mM imidazole. Protein was eluted using an imidazole gradient. Fractions containing Ras were pooled, dialysed into buffer A (20 mM Tris, pH 8.0, 100 mM NaCl), and concentrated. The polyhistidine tag was cleaved by tobacco etch virus (TEV) protease in buffer A and 2 mM β-mercaptoethanol at 4 °C for 48 h. After cleavage, protein was passed over a charged nickel-binding column pre-equilibrated with buffer A to remove uncleaved protein. Fractions containing pure Ras were pooled and concentrated. Expression and purification of Sos (residues 564–1049) was performed as above, with an additional purification step using a HiQ (Biorad) column pre-equilibrated with buffer A. Fractions containing Sos were concentrated in buffer A. The Ras–Sos complex was formed by incubating concentrated Sos with a 3- to 5-fold excess of Ras in buffer A for 1 h at 4 °C. Protein was loaded onto a Sephadex 75 gel filtration column (Pharmacia Biotech) pre-equilibrated with buffer A. Fractions containing complex were pooled and concentrated to 10 mg ml^{-1}. Approximately 30 mg purified complex could be obtained from 16 litres of *E. coli* cell culture.

Crystallographic analysis. Hanging drops of Ras–Sos complex (2.5 μl,

10 mg ml^{-1}) in buffer A were mixed with an equal amount of reservoir buffer containing 2.7–3.2 M sodium formate and 100 mM Tris buffer, pH 8.0, and kept at 4 °C. Crystal showers appeared after 1–2 days with large single crystals growing to full size ($0.3 \times 0.3 \times 0.15$ mm^3) within 2–3 weeks. The crystals contain one heterodimeric complex per asymmetric unit and belong to space group $I422$ ($a = b = 142.7$ Å, $c = 207.9$ Å). Crystals were collected in 3.5 M sodium formate and 100 mM Tris buffer, pH 8.0, and cryoprotected in harvesting buffer with 10% (w/v) sucrose and 10% (v/v) ethylene glycol before flash freezing in liquid propane. Heavy-atom derivatives were prepared by soaking crystals in harvesting buffer containing the following heavy-atom solutions (soak durations were 1 day unless indicated otherwise): ethyl mercurithiosalicylate (EMTS), Baker's dimercurial, plantinum terpyridine (12 h), 0.1 mM; osmium chloride (2 days), *p*-chloromercuribenzoic acid (PCMB), 1 mM; gold cyanide (2 days), 5 mM; trimethyl lead acetate (2 days), 10 mM; Table 1. All but one of the data sets used in this analysis were measured at Brookhaven National Laboratory on beamline X25 using the Brandeis 2×2 (four module) CCD-based detector[34]. A mercury derivative data set (PCMB) was measured at Cornell High Energy Synchrotron Source on beamline F2 using the Q1 CCD-based detector (ADSC). Data processing was performed using Denzo, and data reduction was performed using Scalepack[35]. MIR phases were calculated using MLPHARE as implemented in the CCP4 suite of programs[36]. Solvent flattening was performed using DM[37].

Model building and refinement. Model building was performed using O[38]. Coordinates for Ras bound to a GTP analogue (5P21)[39] were obtained from the Protein Data Bank[40], and the molecule, with Switch 1 and Switch 2 regions deleted, was fit into density. After an initial round of model building and positional refinement using CNS[41] with bulk solvent corrections and anisotropic *B*-factor scaling protocols, phase combination methods using Sigma[42] resulted in a much improved map into which the Switch 2 region of Ras, the entire catalytic domain and the N-terminal helices of Sos were built. Electron density maps based on multiple simulated annealing models[43] allowed the remaining regions of Ras and Sos to be placed into density. Residues 564–567 (N-terminal), 591–597, 654–675, 742–751 and 1045–1049 (C-terminal) are disordered and not modelled in Sos; no residues of Ras are disordered. The Ramachandran plot shows that 89% of all residues are in the most favoured regions and no residues are in disallowed regions.

Received 6 May; accepted 10 June 1998.

1. Bourne, H. R., Sanders, D. A. & McCormick, F. The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349,** 117–127 (1991).
2. Boguski, M. S. & McCormick, F. Proteins regulating Ras and its relatives. *Nature* **366,** 643–654 (1993).
3. Medema, R. H., de Vries-Smits, A. M., van der Zon, G. C. M., Maassen, J. A. & Bos, J. L. Ras activation by insulin and epidermal growth factor through enhanced exchange of guanine nucleotides on p21ras. *Mol. Cell. Biol.* **13,** 155–162 (1993).
4. Buday, L. & Downward, J. Epidermal growth factor regulates p21ras through the formation of a complex of receptor, Grb2 adapter protein, and Sos nucleotide exchange factor. *Cell* **73,** 611–620 (1993).
5. Gale, N. W., Kaplan, S., Lowenstein, E. J., Schlessinger, J. & Bar-Sagi, D. Grb2 mediates the EGF-dependent activation of guanine nucleotide exchange on Ras. *Nature* **363,** 88–92 (1993).
6. Bar-Sagi, D. The Sos (Son of sevenless) protein. *Trends Endocrin. Metab.* **5,** 165–169 (1994).
7. Schlessinger, J. How receptor tyrosine kinases activate Ras. *Trends Biochem. Sci.* **18,** 273–275 (1994).
8. Yu, H. & Schreiber, S. L. Structure of guanine-nucleotide-exchange factor human Mss4 and identification of its Rab-interacting surface. *Nature* **376,** 788–791 (1995).
9. Mossessova, E., Gulbis, J. M. & Goldberg, J. Structure of the guanine nucleotide exchange factor Sec7 domain of human Arno and analysis of the interaction with ARF GTPase. *Cell* **92,** 415–423 (1998).
10. Cherfils, J. *et al*. Structure of the Sec7 domain of the Arf exchange factor ARNO. *Nature* **392,** 101–105 (1998).
11. Renault, L. *et al*. The 1.7 Å crystal structure of the regulator of chromosome condensation (RCC1) reveals a seven-bladed propeller. *Nature* **392,** 97–101 (1998).
12. Wang, Y., Jiang, Y., Meyering-Voss, M., Sprinzl, M. & Sigler, P. B. Crystal structure of the EF-Tu·EF-Ts complex from *Thermus thermophilus*. *Nat. Struct. Biol.* **4,** 650–656 (1997).
13. Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S. & Leberman, R. The structure of the *Escherichia coli* EF-Tu-EF-Ts complex at 2.5 Å resolution. *Nature* **379,** 511–518 (1996).
14. Holm, L. & Sander, C. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233,** 123–138 (1993).
15. Chardin, P. *et al*. Human Sos 1: A gunaine nucleotide exchange factor for Ras that binds to GRB2. *Science* **260,** 1338–1343 (1993).
16. Lenzen, C., Cool, R. H., Prinz, H., Kuhlmann, J. & Wittinghofer, A. Kinetic analysis by fluorescence of the interaction between Ras and the catalytic domain of the guanine nucleotide exchange factor Cdc25^{Mm} *Biochemistry* **37,** 7420–7430 (1998).
17. Lai, C.-C., Boguski, M., Broek, D. & Powers, S. Influence of guanine nucleotides on complex formation between Ras and Cdc25 proteins. *Mol. Cell. Biol.* **13,** 1345–1352 (1993).
18. Mistou, M. Y. *et al*. Mutations of H-Ras p21 that define important regions for the molecular mechanism of the SDC25 C-domain, a guanine nucleotide dissociation stimulator. *EMBO J.* **11,** 2391–2397 (1992).
19. Milburn, M. V. *et al*. Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. *Science* **247,** 939–945 (1990).
20. Jurnak, F. Structure of the GDP domainof EF-Tu and location of the amino acids homologous to ras oncogene proteins. *Science* **230,** 32–36 (1985).
21. Powers, S., O'Neill, K. & Wigler, M. Dominant yeast and mammalian Ras mutants that interfere with CDC25-dependent activation of wild-type Ras in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **9,** 390–395 (1989).
22. Haney, S. A. & Broach, J. R. Cdc25p, the guanine nucleotide exchange factor for the Ras proteins of *Saccharomyces cerevisiae*, promotes exchange by stabilizing ras in a nucleotide free state. *J. Biol. Chem.* **269,** 16541–16548 (1994).
23. Klebe, C., Prinz, H., Wittinghofer, A. & Goody, R. S. The kinetic mechanism of Ran-nucleotide exchange catalyzed by RCC1. *Biochemistry* **34,** 12543–12552 (1995).
24. Verrotti, A. C. *et al*. Ras residues that are distant from the GDP binding site play a critical role in dissociation factor-stimulated release of GDP. *EMBO J.* **11,** 2855–2862 (1992).
25. Segal, M., Willumsen, B. M. & Levitzki, A. Residues crucial for Ras interaction with GDP-GTP exchangers. *Proc. Natl Acad. Sci. USA* **90,** 5564–5568 (1993).
26. Mosteller, R. D., Han, J. & Broek, D. Identification of residues of the H-Ras protein critical for functional interaction with guanine nucleotide exchange factors. *Mol. Cell. Biol.* **14,** 1104–1112 (1994).
27. Segal, M., Marbach, I., Willumsen, B. M. & Levitzki, A. Two distinct regions of Ras participate in functional interaction with GDP-GTP exchangers. *Eur. J. Biochem.* **228,** 96–101 (1995).
28. Leonardsen, L., DeClue, J. E., Lybaek, H., Lowy, D. R. & Willumsen, B. M. Rasp21 sequences opposite the nucleotide binding pocket are required for GRF-mediated nucleotide release. *Oncogene* **13,** 2177–2187 (1996).
29. Crechet, J.-B., Bernardi, A. & Parmeggiani, A. Distal switch II region of Ras2p is required for interaction with guanine nucleotide exchange factor. *J. Biol. Chem.* **271,** 17234–17240 (1996).
30. Quilliam, L. A. *et al*. Involvement of the switch 2 domain of Ras in its interaction with guanine nucleotide exchange factors. *J. Biol. Chem.* **271,** 11076–11082 (1996).
31. Feig, L. A. & Cooper, G. M. Inhibition of NIH 3T3 cell proliferation by a mutant Ras protein with preferential affinity for GDP. *Mol. Cell. Biol.* **8,** 3235–3243 (1988).
32. Chen, S.-Y., Huff, S. Y., Lai, C.-C., Der, C. J. & Powers, S. Ras-15A protein shares highly similar dominant-negative biological properties with Ras-17N and forms a stable, guanine-nucleotide resistant complex with CDC25 exchange factor. *Oncogene* **9,** 2691–2698 (1994).
33. Powers, S., Gonzales, E., Christensen, T., Cubert, J. & Broek, D. Functional cloning of Bud5, a Cdc25-related gene from S. cerevisiae that can suppress a dominant-negative Ras2 mutant. *Cell* **65,** 1225–1231 (1991).
34. Westbrook, E. M. & Naday, I. Charge-coupled device-based area detectors. *Methods Enzymol.* **276,** 244–268 (1997).
35. Otwinowski, Z. & Minor, W. Processing of x-ray diffraction data collected in oscillation model *Mehtods Enzymol.* **276,** 307–326 (1997).
36. CCP4 Suite: Programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
37. Cowtan, K. *Joint CCP4 and ESF-EACBM Newslett. Protein Crystallogr.* **31,** 34–38 (1994).
38. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119 (1991).
39. Pai, E. F. *et al*. Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *EMBO J.* **9,** 2351–2359 (1990).
40. Bernstein, F. C. *et al*. The protein data bank: a computer-based archival file for macromolecular structures. *Arch. Biochem. Biophys.* **185,** 584–591 (1978).
41. Brünger, A. T. *et al*. Crystallography and NMR system (CNS): A new software suite for macromolecular structure determination. *Acta Crystallogr. D* (in the press).
42. Read, R. J. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42,** 140–149 (1986).
43. Brünger, A. T., Adams, P. D. & Rice, L. M. New applications of simulated annealing in X-ray crystallography and solution NMR. *Structure* **5,** 325–336 (1997).
44. Carson, M. Ribbons 2.0. *J. Appl. Crystallogr.* **24,** 958–961 (1991).
45. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11,** 281–296 (1991).
46. Pai, E. F. *et al*. Structure of the guanine-nucleotide-binding domain of Ha-ras oncogene product p21 in the triphosphate conformation. *Nature* **341,** 209–214 (1989).

Acknowledgements. We thank L. Leighton and M. Uy for technical assistance; J. Bonanno, S. K. Burley, R. H. Chen, X. Chen, E. Conti, S. Corbalan-Garcia, D. Galron, B. Hall, M. Huse, S. Jacques, D. Jeruzalmi, C.-H. Lee, S. Soisson, H. Yamaguchi, S.-S. Yang and Y. Zhang for advice and help; L. Berman, R. Sweet, the staff of NSLS and CHESS for help with synchrotron data collection; and A. Wittinghofer and R. H. Cool for sharing a manuscript prior to publication. P.A.B.-S. is supported by grants for the Cystic Fibrosis Foundation and the NIH. D.B.-S. acknowledges grant support from the NIH.

Correspondence and requests for materials should be addressed to J.K. The atomic coordinates have been deposited in the Brookhaven Protein Data Bank under code 1BKD.