Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 5443–5447, June 1991
Biochemistry

GTPase domains of ras p21 oncogene protein and elongation factor Tu: Analysis of three-dimensional structures, sequence families, and functional sites

(guanine nucleotide binding/structure comparison/multiple sequence alignment/contact analysis/exchange factor)

ALFONSO VALENCIA*, MORTEN KJELDGAARD†, EMIL F. PAI‡, AND CHRIS SANDER*

*European Molecular Biology Laboratory, Meyerhofstrasse 1, D-6900 Heidelberg, Federal Republic of Germany; †Department of Chemistry, University of Aarhus, Langelandsgade 140, DK-8000 Aarhus C, Denmark; and ‡Max-Planck-Institute für Medizinische Forschung, Abteilung Biophysik, Jahnstrasse 29, D-6900 Heidelberg, Federal Republic of Germany

Communicated by H. A. Scheraga, December 13, 1990

ABSTRACT **GTPase domains are functional and structural units employed as molecular switches in a variety of important cellular functions, such as growth control, protein biosynthesis, and membrane traffic. Amino acid sequences of more than 100 members of different subfamilies are known, but crystal structures of only mammalian ras p21 and bacterial elongation factor Tu have been determined. After optimal superposition of these remarkably similar structures, careful multiple sequence alignment, and calculation of residue–residue interactions, we analyzed the two subfamilies in terms of structural conservation, sequence conservation, and residue contact strength. There are three main results. (*i*) A structure-based alignment of p21 and elongation factor Tu. (*ii*) The definition of a common conserved structural core that may be useful as the basis of model building by homology of the three-dimensional structure of any GTPase domain. (*iii*) Identification of sequence regions, other than the effector loop and the nucleotide binding site, that may be involved in the functional cycle: they are loop L4, known to change conformation after GTP hydrolysis; helix α2, especially Arg-73 and Met-67 in ras p21; loops L8 and L10, including ras p21 Arg-123, Lys-147, and Leu-120; and residues located spatially near the N and C termini. These regions are candidate sites for interaction either with the GTP/GDP exchange factor, with a GTPase-affected function, or with a molecule delivered to a destination site with the aid of the GTPase domain.**

GTPase Domains. GTPase domains, or G domains, are used as molecular switches in the regulation of various cellular functions. The two states of the switch are the GTP and GDP forms. Details of the functional cycle are known in some cases: the acceleration of the GTPase rate by binding to another protein, dissociation of a molecular complex as a result of the GTP/GDP transition, or catalysis of GDP/GTP exchange by another protein (for reviews, see refs. 1 and 2).

Function and Structure of ras p21. Mammalian ras p21, an oncoprotein, is involved in growth control (3). The G domain comprises the first 166 of 189 residues. The high-resolution (1.35 Å) x-ray structure of the G domain of human ras p21 with a bound GTP analogue (4, 5), as well as the structures of oncogenic mutants (6) and of the GTP- and GDP-bound forms (7, 8), have led to detailed molecular models of normal and altered GTPase activity. The intrinsic GTPase activity is lower by a factor of about 10 in the oncogenic mutants, such as G12V, as a result of blocking access to the active site. An effector protein, called GTPase activating protein (GAP) (9) or yeast cell-division-cycle protein CDC25 (10), binds to (at least) loop L1 (see Fig. 1) and accelerates the GTPase rate. However, the molecular partner directly activated by ras p21 (if it is not identical with GAP) has not yet been identified. A GDP-releasing factor of mammalian ras p21 has been identified (11–13).

Function and Structure of Elongation Factor Tu (EF-Tu). EF-Tu is involved in delivering the charged tRNA to the ribosome during protein synthesis (14). The GTPase domain comprises 200 of 393 residues (*Escherichia coli*). GTPase activity is accelerated by binding to the ribosome (15). GDP/GTP exchange is catalyzed by an exchange factor, elongation factor Ts (EF-Ts) (16) for EF-Tu or guanine nucleotide exchange factor (17) for eukaryotic initiation factor 2 (see ref. 14). The first G-domain crystal structure was that of *E. coli* EF-Tu (18, 19), which identified the nucleotide binding site. The full EF-Tu structure includes domains 2 and 3 (M.K. and J. Nyborg, unpublished results). Domain 2 is located at the end of the C-terminal helix of the G domain. Domain 3 closes the overall ring-like structure. Helices IB and IC (α2 and α3 in ras p21 notation) form the interface between the G domain and domain 3. Neither the sites of interaction with tRNA, with EF-Ts, nor with the ribosome are completely known (20).

Structural and Functional Similarities. Sequence similarity between the G domains of ras p21 and EF-Tu had been detected by 1984 (21). Its structural interpretation (22) was confirmed once the correct crystal structure of ras p21 became known (4, 5): the striking similarity of three-dimensional structure and the near-identity of the GTPase mechanism. Structural and functional analogies, extending in part to adenylate kinase and the phosphate-binding chemotaxis protein CheY, were explored by comparison of crystal structures and calculation of low-energy peptide conformations (23–25).

Further Functional Analogies? (*i*) Conceivably, GAP (for p21) and the ribosome (for EF-Tu) have analogous functions—both turn on the GTPase. (*ii*) The recently discovered releasing factor for ras p21 (11–13) appears to be analogous to EF-Ts. For other components of the elongation system, the ras p21 analogues, if they exist, remain to be found: (*iii*) a "load" to be delivered (tRNA for EF-Tu) and (*iv*) complex-forming proteins (domains 2 and 3 for EF-Tu). Where do these possible partners interact with p21?

Sequence Subfamilies. There are currently about 40 known homologues of EF-Tu and more than 50 for ras p21. The EF-Tu family includes human gst1-hs, yeast sup2 suppressor protein, bacterial elongation factors 1α, selb selenocysteine translation factor, lepa protein, nodulation protein Q, initiation factor IF-2, elongation factor EF-2, and tetracycline-resistance proteins. The ras p21 family includes proteins from

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: GAP, GTPase activating protein; EF-Ts, elongation factor Ts; EF-Tu, elongation factor Tu.

the ras (rap, ral, and rsr), rab (ypt, sec4, and smgp25), and rho (rac) subfamilies, from various species ranging from yeast to mammals (26). Multiple sequence alignment reveals strictly conserved residues of structural or functional importance.

Our goals are to (*i*) derive a structure-based alignment of the two families, (*ii*) define a common conserved core for model-building purposes, and (*iii*) help identify key residues involved in the functional cycle as a guide to experiments.

METHODS

We introduce a methodology for the comparative analysis of two protein families that combines superposition of three-dimensional structures with multiple sequence alignment and contact analysis.

Structural Conservation: Methods for Structural Alignment. To define residues equivalent in three dimensions (see stereoview, Fig. 2*A*; sequence alignment, Fig. 3), we used an automatic (27) or an interactive (28) procedure to minimize the rms difference of C^{α} positions after optimal superposition. Ambiguities (rms deviations >2.5 Å after the first step) were resolved by requiring that equivalent residues have a similar pattern of contacts with their nearest neighbors (for definition of contact strength, *ContNum*, see legend to Fig. 3). Structural variation is defined as the difference in C^{α} position of equivalent residues after optimal superposition. Structural conservation (*StrCons*) is defined as a constant (3.5 Å) minus structural variation.

Sequence Conservation: Multiple Sequence Alignment. Multiple sequence alignments and sequence variability *within each family* were taken from ref. 29. Each family alignment includes only sequences deemed structurally homologous to ras p21 (54 proteins) or EF-Tu (36 proteins). Sequence variability at a residue position is calculated from the amino acid similarity averaged over all sequence pairs at that position. Sequence conservation (*SeqCons*) is defined as a constant minus sequence variability.

Combining Sequence Conservation, Contact Strength, and Structural Conservation. Sequence-conserved residues are divided into several categories. Those with high contact strength are likely to be structurally important. Of these, residues with high structural conservation are likely to have an important structural role and are defined to be part of the "structural core," while those with low structural conservation are likely to have a different structural role for reasons related to the particular function of each protein and are defined to be part of a "structural fine-tuning" region. Sequence-conserved residues with low contact strength are accessible to solvent on the protein surface and are more likely to be conserved for nonstructural reasons (i.e., they may very well interact with a substrate or another protein); these residues are considered part of "functional interface" regions.

By using empirical cutoffs (see Fig. 3) to divide the value range of the three variables (*SeqCons*/*ContNum*/*StrCons*) into high (hi) and low (lo) values, the three residue classes of interest thus are the structural core (hi/hi/hi), the structural fine-tuning region (hi/hi/lo), and the functional interface region (hi/lo/any value). Note that calculation of sequence conservation and contact strength is done separately in each family, while structural conservation is calculated from the structural alignment of the two (for normalization, see legend to Fig. 3).

RESULTS

Optimal Structural Superposition of ras p21 and EF-Tu. The common G domain consists of a central six-stranded β-sheet surrounded by five α-helices (Fig. 1). The C-terminal half of the domain has a regular β/α/β/α/β/α topology. The N-terminal half has an unusual antiparallel strand pair at the edge of the otherwise parallel sheet. The first strand and helix form a characteristic phosphate-binding motif. In the structural

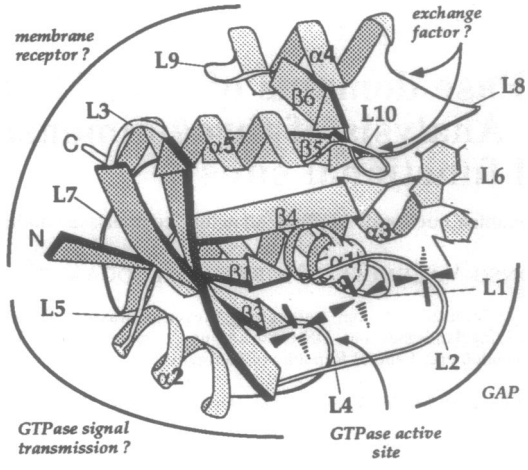

FIG. 1. Structural cartoon of the G domain of p21 ras (4). β-Strands are drawn as arrows (β1–β6), α-helices as helical ribbons (α1–α5), and loops as double lines (L1–L10). The GTP location is approximate. Known and hypothetical interaction sites are labeled. This figure was adapted from a drawing by Doug Lowy (National Institutes of Health).

superposition (Fig. 2*A*), there is a striking coincidence of the C^{α} trace in three dimensions for all strands of the β-sheet, for helices α1, α3, and α5, and for the GNP binding loops. These loops are at the C terminus of the β-strands [right side of structure in Fig. 1; loop L1: p21 residues 10–17 (GAGGVGKS)and EF-Tu residues 18–25 (GHVDHGKT); loop L8: p21 residues 116–119 (NKCD) and EF-Tu residues 135–138 (NKCD); loop L10: p21 residues 145 and 146 (SA) and EF-Tu residues 173 and 174 (SA)]. The simplest explanation for the structural similarity, considering the different cellular role of p21 and EF-Tu, is a strong selective pressure related to the proper functioning of the GTPase conformational switch.

Overall Structural Differences. In spite of the striking similarity of the C^{α} trace of the central β-sheet, there are several differences in the interaction with the surrounding helices. In both proteins, the packing is primarily hydrophobic. However, there are compensating mutations that conserve approximately the total volume of several side chains. An example is in a hydrophobic cluster in the area of the tight turn between β4 and α3. In ras p21, the central residue of this cluster is Phe-78 on β4, whereas in EF-Tu, the central residue of the cluster is Met-98 on loop L5. In ras p21, the following residues participate in the cluster: Val-7, Gly-75, Phe-78, and Ile-100, or, in terms of size, big, small, big, big. In EF-Tu, the residues occupying the same spatial positions are, respectively, Gly-15, Met-98, Ala-101, and Gly-122, or small, big, small, small. Another example of volume compensation is in the region between helix α5 and β-strands β1, β2, and β3.

Most C^{α} positional deviations are in loops at the beginning of the strands (left side of structure in Fig. 1), away from the GNP, as well as in helices α2 and α4. The largest difference in an element of secondary structure is the angular orientation and extent of helix α2 (Fig. 2*B*).

Structural Versus Sequence Alignment. The set of equivalent residue pairs resulting from the structural superposition can be interpreted as a one-dimensional sequence alignment (Fig. 3). In the trunk (107 residues), sequence identity is only 17%—so low that any simple pairwise sequence alignment method would produce regions of local alignment in disagreement with the structural superposition. For example, in helix α2, sequence alignment would match TG (p21 residues 74 and 75) with TG (EF-Tu residues 93 and 94), while structurally TG (p21 residues 74 and 75) corresponds to QM (EF-Tu residues 97 and 98), four residues downstream; in helix α3, QYREQ (p21 residues 95–99) matches QTREH (EF-Tu res-

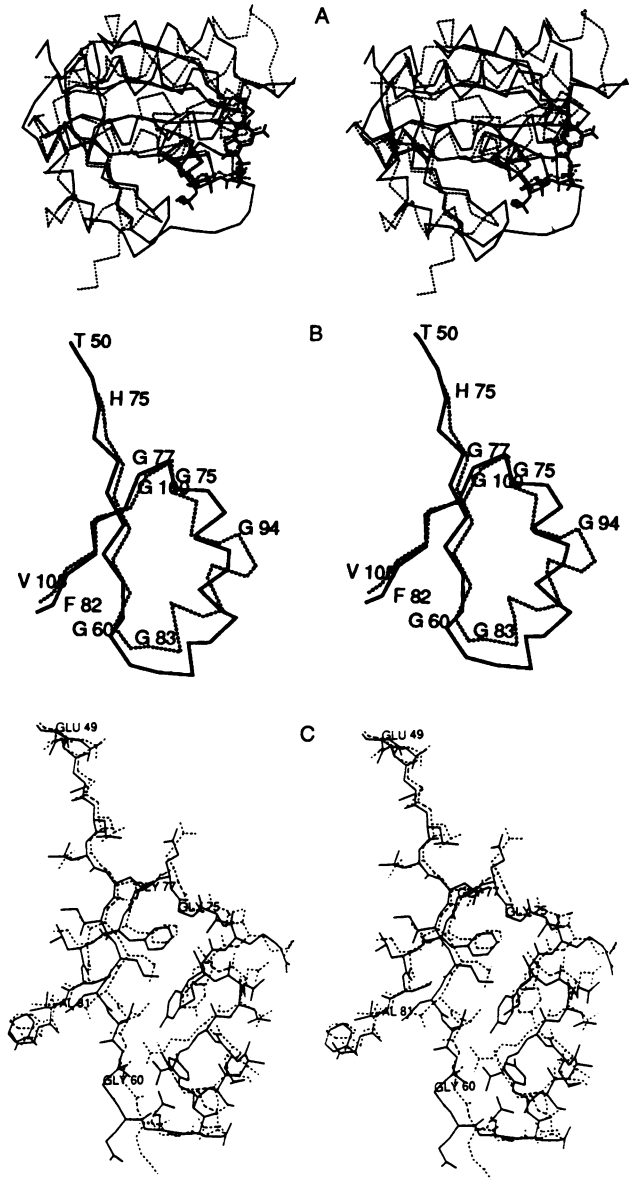

FIG. 2. Comparison of the crystal structures of the G domains of p21 and EF-Tu. Equivalent residues in the two proteins and the best superposition were determined by an automatic procedure (27). The coordinates for p21 are from the high-resolution guanosine 5'-[β,γ-imido]triphosphate-bound form (4, 5); those for EF-Tu are from the GDP-bound form at 2.6 Å resolution (18). (*A*) Optimal superposition of ras and EF-Tu. A stereoview of the C^{α} trace is shown: the solid line is p21 and the dotted line is EF-Tu. GTP/GDP is at the right. The rms position difference is 1.36 Å for 107 C^{α} atom pairs. (*B*) Ras and EF-Tu: details in the region of helix α2. A stereoview of p21 (solid line) and EF-Tu (dotted line) is shown. Note the different orientation of the helical axes of helix α2 relative to the β-strands. In EF-Tu, helix α2 is involved in an interdomain contact; an analogous interaction may occur between p21 and an associated protein. (*C*) GTP and GDP forms of p21: details in the region of helix α2. For comparison with *B*, a stereoview of the GTP form (solid line) and GDP form (dotted line) of p21-G12V (Val at position 12), as captured by time-resolved x-ray crystallography (8) is shown. Small but distinct conformational changes are visible (e.g., in the side chain orientation of Thr-58, Arg-68, and Tyr-71) but not a reorientation of the helical axis as in *B*. Apparently, the differences between p21 and EF-Tu in this region (see *B*) do not primarily reflect changes between the triphosphate (p21) and diphosphate (EF-Tu) forms.

idues 114–118) in sequence, while in structure it matches EHILL (EF-Tu residues 117–121), shifted by three residues. Both correspondences, however, may be "right." The sequence alignment may reflect an actual evolutionary pathway, yet the structure may have shifted by a few residues in response to altered intraprotein interactions.

Definition of a Conserved Structural Core. Earlier definitions of structural core are in terms of conserved secondary structure elements (32) or conserved hydrophobic residues (33). Here, we have defined structural core in terms of important interactions and require not only conservation of sequence and structure but also sizable intraprotein contacts (see *Methods* and Fig. 3). The core residues (Fig. 3) include primarily residues in the interior strands of the β-sheet, in helix α1, and in part of the loop regions in contact with the bound nucleotide. For ras p21, 32 of 166 residues are in the core; for EF-Tu, 35 of 177 are in the core. Of the 116 equivalent residues, 21 are in both cores and thus define a common core (thick bars just above and below the sequences in Fig. 3). The common core overlaps with the regions shown in deletion experiments to be essential for ras p21 function (34).

Conserved Nucleotide-Binding Pocket. The structurally most conserved regions include the GNP-binding residues in the four sequence-conserved "boxes" (solid bars between the two sequences in Fig. 3): GXXXXGKS, DXXGQ/DXXGH, NKXD, and SAK/SAL (p21/EF-Tu). The hydrogen bonds between Asp-119 (p21) or Asp-138 (EF-Tu) in the NKXD box and the guanine base are probably a main factor in determining the specificity of the nucleotide-binding pocket of G domains. Strong evolutionary pressure has maintained remarkably similar ways of binding nucleotides.

The most significant difference in the nucleotide-binding site is in loop L4 (see below). Other differences are as follows: (*i*) In the binding of the phosphate tail, the side chain of Thr-26 forms a hydrogen bond to the pro-S oxygen of the α-phosphate in EF-Tu but not in p21 (Thr changed to Ala). (*ii*) In the binding of the guanine base, the stabilizing hydrogen bond in ras p21 between Thr-144 and Asn-116 does not exist in EF-Tu (Thr changed to Gly). (*iii*) At the current level of refinement of EF-Tu, the orientation of the base-binding Asp-138 is slightly different from that in p21. (*iv*) In both proteins, the guanine base is located in a hydrophobic pocket. The bottom of the pocket is formed by the aliphatic part of Lys in the NKXD loop. In ras p21 the roof of the pocket is formed by Phe-28 of the effector loop L2, but in EF-Tu it is formed by Leu-175 in helix α5 (L in the SALK box). The equivalent physical space is occupied by a side chain from a nonequivalent part of each protein. (*v*) Unfortunately, the functionally important effector loop (L2), including the YDP "flip peptide" (ras p21 positions 34–36), which changes conformation as a result of hydrolysis (8, 9), cannot be compared structurally, as the corresponding residues were removed from EF-Tu by a tryptic cut of 14 residues (17).

The Region of Helix α2: Structural Changes as a Result of GTPase Activity. In ras p21, the largest conformational changes between the GTP and GDP forms are in the effector loop (L2) and in the region following the DXXGQ box (residues 57–61) (7, 8). In the high-resolution structure (5), the electron density in loop L4 can be explained by two alternate conformations of Gln-61 (large temperature factors for a single model). Reorientation of the side chain of Gln-61 appears to be the rate-limiting step in hydrolysis. In our analysis this region is correctly picked out as functionally important solely on the basis of conservation and contact criteria (Figs. 3 and 4*A*). It is plausible that His-84 in EF-Tu participates in catalysis in a manner analogous to Gln-61 of ras p21.

Interestingly, helix α2 is oriented differently in p21 and EF-Tu (Fig. 2*B*), in spite of the similar sequence length L4–α2–L5 (Fig. 3). However, its patterns of contacts are preserved, in particular the fairly strong set of contacts between helices α2 and α3, involving two to four turns on each helix. The different orientation of helix α2 in ras and

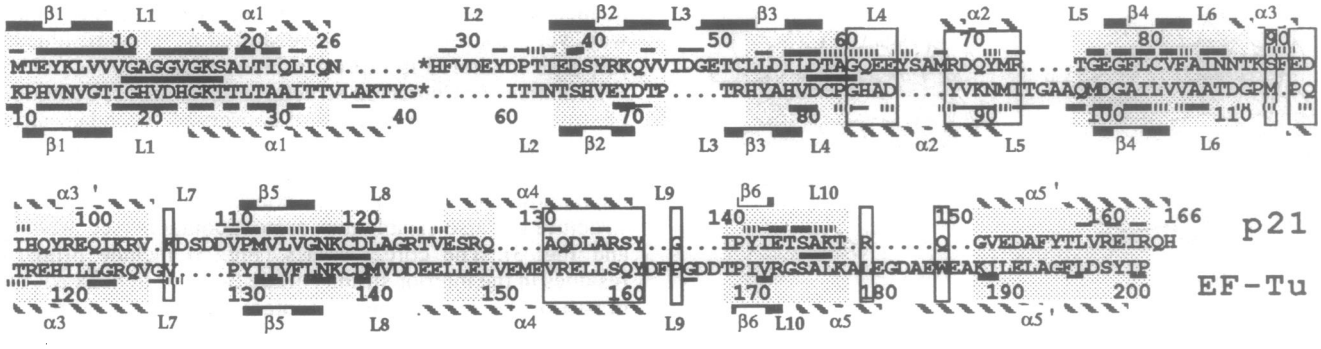

FIG. 3. Structural alignment of p21 (upper sequence) and EF-Tu. Residues in boxes are equivalent. There are three levels of confidence: (*i*) shaded boxes, residue pairs aligned by using the automatic procedure (the "trunk"); (*ii*) boxes, residues aligned by contact environment; (*iii*) outside the boxes, residues without a clear structural equivalent (arbitrary alignment). Contact similarity was judged by visual inspection of contact maps (program CONAN) (30) or by a two-stage dynamic programming algorithm (ref. 31; C.S. and M. Scharf, unpublished results). Residue–residue contact strengths are calculated as a sum over potential energies (6–9 potential), not counting contacts inside a secondary structure element. Bars at the very top and bottom of the structures are secondary structure elements labeled β1, β2, etc. for strands and α1, α2, etc. for helices. Intervening loops are labeled L1, L2, etc. Amino acid residues are one-letter code. EF-Tu residue numbers are from the Brookhaven Protein Data Bank (set 1ETU). The stars mark the 20-residue segment missing in EF-Tu. Conserved residues are marked by bars near the sequences. Black bars between the two sequences indicate regions of sequence conservation involved in nucleotide binding. The thick bars just above and below the sequences indicate core residues; bars made of vertical lines just above and below the sequences indicate fine-tuning residues; thin bars just above and below the sequences indicate potential interface residues. The variables *SeqCons*, *StrCons*, and *ContNum* are mapped onto the interval [0, 1] (0 = minimum, 1 = maximum value in the particular protein); *ContNum* is scaled independently for each residue type, so *ContNum* values express the contact strength relative to other residues of the same type. Cutoffs are fixed empirically (i.e., somewhat arbitrarily) at 0.55 for *SeqCons* and at 0.3 for all other variables. For example, core residues have *SeqCons* > 0.55, *StrCons* > 0.3, and *ContNum* > 0.3; in addition, in each residue category, residues must also pass a threshold for a combined quantity: *SeqCons* + *StrCons* + *ContNum* > 1.5 for the core region and *SeqCons* + *StrVar* + *ContNum* > 1.0 for the fine-tuning region.

EF-Tu (Fig. 2*B*) does not reflect the difference between the GDP and GTP forms, as the conformational changes in this region between the GTP and GDP forms of ras p21 (7, 8) are much less pronounced (Fig. 2*C*). A more likely explanation is the different role of helix α2 in interdomain/interprotein interaction (see below).

Helices α2 and α3: Role in External Interaction. In EF-Tu, residues on helices α2 and α3 are involved in an interdomain interface with residues in three β-strands in domain 3 (ref. 18; M.K. and J. Nyborg, unpublished results): Val-88, Lys-89, Ile-92, Thr-93, and Gln-124. In EF-Tu this region is a possible binding site for the aminoacyl-tRNA (20). It is quite plausible that as a result of GTP hydrolysis the conformational change in loop L4 and helix α2, expected in analogy to p21, affects the interface between helices α2/α3 and domain 3, with a resulting release of tRNA and/or dissociation from the ribosome. Our analysis without prior knowledge of the details of the three-domain structure, correctly predicted that His-84, Val-88, Lys-89, Ile-92, and Thr-93 are interface residues in EF-Tu (thin bars just above and below the sequences in Fig. 3). This remarkable agreement (four of five) encourages similar predictions for p21.

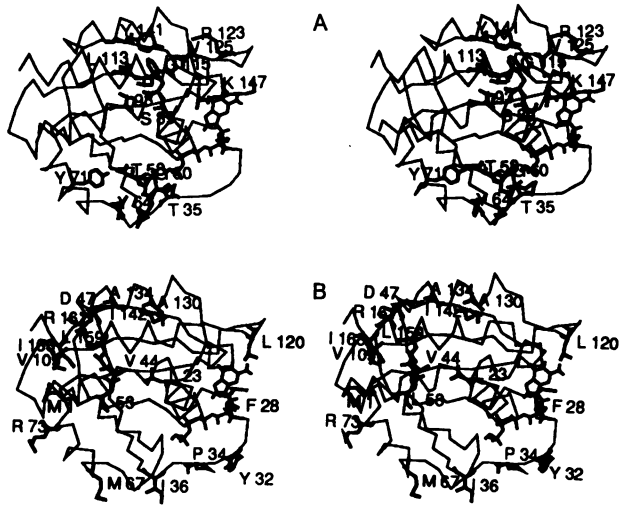

FIG. 4. Fine-tuning and interface residues in ras p21 (side chains are shown). (*A*) Fine-tuning residues are candidates for possible functional importance specific to each family. (*B*) Interface residues may interact with other proteins. These residues are good candidates for mutational analysis.

The solvent-exposed residues (underlined) in the equivalent region of p21 are on helix α2, residues 63–74 (EYSAMRDQYMRT) and on the C-terminal part of helix α3, residues 93–103 (IHQYREQIKRV). Of these, Met-67 and Arg-73 are classified in our conservation analysis as interface residues (Figs. 3 and 4*B*)—i.e., as likely candidates for maintaining a possible functional external interaction. In fact, mutations in residues Glu-63, Ser-65, Ala-66, Met-67, Gln-70, and Arg-73 affect binding of an antibody known to neutralize the biological transforming activity of ras p21 (35). Binding of the antibody is also reported to interfere with GDP/GTP exchange (36). It may be useful to mutate the residues underlined above, especially Met-67 and Arg-73, in the search for an activity affected by the GTPase switch.

Loops L8 and L10: Regulation of Nucleotide Binding? Protein–base contacts, made mainly by the NKXD box in loop L8 and the SA box (indirectly) in loop L10, determine the high affinity and specificity of binding. Mutations in these loops can strongly affect nucleotide binding (37, 38). Adenylate kinase and phosphofructokinase, which bind nucleotide less tightly than p21 and EF-Tu, lack the "extra" helix (α4) and strand (β6) between loops L8 and L10 and do not have analogues of the NKXD and SA boxes (39–41).

Our analysis points to fine-tuning residues (vertical lines just above and below the sequences in Fig. 3) near these boxes that provide structural support near the guanine base binding pocket in a way specific to each subfamily and that may be used to modulate nucleotide affinity. In ras p21 these residues are (Fig. 4*A*) Arg-123 and Glu-143 (salt bridge) as well as Phe-82, Phe-90, Tyr-141, Leu-113, and Val-125 (nonpolar cluster). The EF-Tu family may have a larger spread in nucleotide affinity, as its sequences are less conserved in this region; for example, the elongation factors 1α do not have the SAK/SAL box.

In G domains, loops L8/L10 may interact with the nucleotide exchange factors, such as EF-Ts for EF-Tu. For ras, such factors have been purified in animals (11–13) and in yeast (sdc25) (42). Information on the site of interaction comes from mutation experiments; e.g., the activity of bovine brain ras guanine nucleotide exchange factor (13) is not affected by a Gly-12 → Val or Ala-59 → Thr mutation in p21. In contrast, a variant of yeast ras2, mutated (Thr → Ile) at what would be Thr-144 of ETSAK in p21, has an enhanced rate of GDP/GTP exchange (43). We speculate that dissociation of bound nucleotide requires "loosening" of protein–base contacts in loops L8/L10, brought about by interaction of these loops with the exchange factor.

At least in one set of experiments (13), an exchange factor appears to act equally on the ras-like p21 proteins Rras, rap1A, rab1B, and rho. This suggests that the interacting residues should be strongly conserved in the p21 family. For mutational tests of interaction with the exchange factors, the best candidates in p21 are therefore the most conserved solvent-exposed residues on loop L8, Leu-120 (in close proximity to the guanine base; Fig. 4*B*) and Arg-123, as well as Lys-147 in loop L10.

Odd-Numbered Loops: Possible Function. In G domains, the odd-numbered loops and the N and C termini are on the side away from the nucleotide (Fig. 1). Here, ras p21 and EF-Tu differ in detail. Based on our conservation analysis, several sets of residues in this region may have functional importance.

(*i*) Conserved interface residues (Fig. 3). In ras p21 they are Met-1, Asp-47, Gly-48, Arg-73, Val-109, Ile-142, Leu-159, Arg-161, and Leu-163 (Fig. 4*B*); in EF-Tu they are Asp-17, Thr-71, Ile-92, Thr-93, Gly-94, Gly-126, Gly-164, Leu-195, and Pro-200.

(*ii*) A cluster of charged residues in ras p21: Asp-47, Glu-49, Glu-76, and residues 104–108 (KDXDD). The cluster is not present in EF-Tu.

(*iii*) Residues that stand out in mutation experiments. Simultaneous mutations of Gly-75 and Gly-77 are able to inactivate ras p21 (44).

In ras p21, the odd-numbered loops are not far from the sequence-diverged C-terminal "specificity" tail involved in membrane attachment. A reasonable hypothesis is that some of these residues interact with a membrane-associated protein, possibly a receptor.

CONCLUSION

The integrated use of multiple sequence alignment, superposition of three-dimensional structures, and contact analysis has led to insight into the relation between function and evolutionary conservation of G domains. One should be aware, however, that this analysis can be used to generate reasonable hypotheses about the role of particular residues but cannot prove such a role. We encourage mutational experiments to prove or disprove the hypotheses put forward here. The method may be useful for other protein families, particularly as a guide to experiments.

We thank Reinhard Schnedier for HSSP; Gert Vriend for WHAT IF; Michael Scharf for CONAN; Brian Clark, Ottavio Fasano, Wolfgang Kabsch, Jens Nyborg, Ilme Schlichting, and Alfred Wittinghofer for stimulating discussions; and Ken Holmes for initiating and supporting the ras project. Because of space limitations, some of the original references have been replaced by reviews or more recent references in which the original work is cited.

1. Bourne, H. R., Sanders, D. A. & McCormick, F. (1990) *Nature (London)* **348,** 125–132.
2. Bourne, H. R., Sanders, D. A. & McCormick, F. (1990) *Nature (London)* **349,** 117–127.
3. Barbacid, M. (1987) *Annu. Rev. Biochem.* **56,** 779–827.
4. Pai, E. F., Kabsch, W., Krengel, U., Holmes, K. C., John, J. & Wittinghofer, A. (1989) *Nature (London)* **341,** 209–214.
5. Pai, E. F., Krengel, U., Petsko, G. A., Goody, R. S., Kabsch, W. & Wittinghofer, A. (1990) *EMBO J.* **9,** 2351–2359.
6. Krengel, U., Schlichting, I., Scherer, A., Schumann, R., Frech, M., John, J., Kabsch, W., Pai, E. F. & Wittinghofer, A. (1990) *Cell* **62,** 539–548.
7. Milburn, M. V., Tong, L., deVos, A. M., Brunger, A., Yamaizumi, Z., Nishimura, S. & Kim, S.-H. (1990) *Science* **247,** 939–945.
8. Schlichting, I., Almo, S. C., Rapp, G., Wilson, K., Petratos, K., Lentfer, A., Wittinghofer, A., Kabsch, W., Pai, E. F., Petsko, G. A. & Goody, R. S. (1990) *Nature (London)* **345,** 309–315.
9. McCormick, F., Adari, H., Trahey, M., Halenbeck, R., Koths, K., Martin, G. A., Crosier, W. J., Watt, K., Rubinfeld, B. & Wong, G. (1988) *Cold Spring Harbor Symp. Quant. Biol.* **53,** 849–869.
10. Broek, D., Takashi, T., Michaeli, T., Levin, L., Birchmeier, C., Zoller, M., Power, S. & Wigler, M. (1987) *Cell* **48,** 789–799.
11. Wolfman, A. & Macara, I. G. (1990) *Science* **248,** 67–69.
12. Downward, J., Riehl, R., Wu, L. & Weinberg, R. A. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 5998–6002.
13. Huang, Y. K., Kung, H. F. & Kamata, T. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 8008–8012.
14. Riis, B., Rattan, S. I. S., Clark, B. F. C. & Merrick, W. C. (1990) *Trends Biochem. Sci.* **15,** 420–484.
15. Kaziro, Y. (1978) *Biochim. Biophys. Acta* **535,** 95–107.
16. Miller, L. & Weissbach, H. (1977) in *Molecular Mechanism of Protein Biosynthesis*, eds. Weissbach, H. & Pestka, S. (Academic, New York), pp. 223–273.
17. Panniers, R., Rowlands, A. G. & Hendhaw, E. (1988) *J. Biol. Chem.* **263,** 5519–5525.
18. Clark, B. F. C., Kjeldgaard, M., LaCour, T. F. M., Thirup, S. & Nyborg, J. (1990) *Biochim. Biophys. Acta* **1050,** 203–208.
19. Jurnak, F. (1985) *Science* **130,** 32–36.
20. Parmeggiani, A., Swart, G. W. M., Mortensen, K. K., Jensen, M., Clark, B. F. C., Dente, L. & Cortese, R. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 3141–3145.
21. Halliday, K. R. (1984) *J. Cyclic Nucleotide Protein Phosphorylation Res.* **99,** 435–448.
22. McCormick, F., Clark, B. F. C., LaCour, T. F. M., Kjeldgaard, M., Nørskov-Lauritsen, L. & Nyborg, J. (1985) *Science* **230,** 78–82.
23. Chen, J. M., Lee, G., Murphy, R. B., Carty, R. P., Brandt-Rauf, P. W., Friedman, E. & Pincus, M. R. (1989) *J. Biomol. Struct. Dyn.* **6,** 859–875.
24. Chen, J. M., Lee, G., Murphy, R. B., Brandt-Rauf, P. W. & Pincus, M. R. (1990) *Int. J. Pept. Protein Res.* **36,** 1–6.
25. Chen, J. M., Lee, G., Brandt-Rauf, P. W., Murphy, R. B., Gibson, K. D., Scheraga, H. A., Rackovsky, S. & Pincus, M. R. (1990) *Int. J. Pept. Protein Res.* **36,** 247–254.
26. Valencia, A., Chardin, P., Witinghofer, A. & Sander, C. (1991) *Biochemistry*, in press.
27. Vriend, G. & Sander, C. (1991) *Proteins* **9,** in press.
28. Jones, T. A., Zou, J.-Y., Cowan, S. & Kjeldgaard, M. (1990) *Acta Crystallogr.* **A47,** 110–119.
29. Sander, C. & Schneider, R. (1991) *Proteins* **9,** 56–68.
30. Scharf, M. (1989) Thesis (University of Heidelberg, Heidelberg).
31. Taylor, W. & Orengo, C. A. (1989) *J. Mol. Biol.* **208,** 1–22.
32. Chothia, C. & Lesk, A. M. (1986) *EMBO J.* **5,** 823–826.
33. Hubbard, T. J. P. & Blundell, T. L. (1987) *Protein Eng.* **3,** 159–171.
34. Willumsen, B. M., Papageorge, A. G., Kungk, H.-F., Bekesi, E., Robin, T., Johnsen, M., Vass, W. C. & Lowy, D. R. (1986) *Mol. Cell. Biol.* **6,** 2654–2654.
35. Sigal, I. S., Gibbs, J. B., D'Alonzo, J. S. & Scolnick, E. M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4725–4729.
36. Hattori, S., Clanton, D. J., Sath, T., Nakamura, S., Kaziro, Y., Kawakita, M. & Shih, T. Y. (1987) *Mol. Cell. Biol.* **7,** 1999–2002.
37. Hwang, Y.-W., Sanchez, A. & Miller, D. L. (1989) *J. Biol. Chem.* **264,** 8304–8309.
38. Clanton, D. J., Hattori, S. & Shih, T. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 5076–5080.
39. Reinstein, J., Vetter, I. R., Schlichting, I., Roesch, P., Wittinghofer, A. & Goody, R. (1990) *Biochemistry* **29,** 7440–7450.
40. Miller, L. & Weissbach, H. (1977) in *Nucleic Acid Protein Recognition*, ed. Vogel, H. J. (Academic, New York), pp. 409–440.
41. Feuerstein, J., Goody, R. S. & Wittinghofer, A. (1987) *J. Biol. Chem.* **262,** 8455–8458.
42. Crechet, J.-B., Poullet, P., Mistou, M.-Y., Parmeggiani, A., Camonis, J., Boy-Marcotte, E., Damak, F. & Jacquet, M. (1990) *Science* **248,** 866–868.
43. Crechet, J.-B., Poullet, P., Camonis, J., Jacquet, M. & Parmeggiani, A. (1990) *J. Biol. Chem.* **265,** 1563–1568.
44. Fasano, O., Crechet, J.-B., De Venditis, E., Zahn, R., Feger, G., Vitelli, A. & Parmeggiani, A. (1988) *EMBO J.* **7,** 3375–3383.