Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 8849–8853, December 1986
Biochemistry

A GTP-binding protein of *Escherichia coli* has homology to yeast RAS proteins

(DNA sequence/oncogenes/photoaffinity cross-linking/ribonuclease III)

JOOHONG AHNN, PAUL E. MARCH, HOWARD E. TAKIFF*, AND MASAYORI INOUYE

Department of Biochemistry, State University of New York at Stony Brook, Stony Brook, NY 11794

Communicated by Alexander G. Bearn, August 4, 1986

ABSTRACT **The DNA sequence of a gene (*era*) located immediately downstream of the gene (*rnc*) encoding ribonuclease III of *Escherichia coli* was determined and found to encode a protein of 316 amino acid residues. The amino acid sequence of this protein, Era, has significant similarity to the yeast RAS proteins. Overexpression of the Era protein was achieved and GTP cross-linking experiments demonstrated that the protein was indeed capable of binding GTP, as are the yeast and mammalian *ras* gene products. These data indicate that *ras*-related sequences occur not only in eukaryotes but also in prokaryotes.**

The *ras* family of oncogenes was discovered in Harvey and Kirsten murine sarcoma viruses and was subsequently shown to exist in mammalian cells as protooncogenes (for review see refs. 1 and 2). Further investigations demonstrated that certain tumor cells contain structurally mutated *ras* genes capable of tumorigenic transformation of NIH3T3 cells upon gene transfer (3–7). These findings have led to extensive investigations into the structure and function of *ras* genes. At least three genes (Ha-, Ki-, and N-*ras*) have been discovered in mammals and found to encode highly related proteins (8–10). In addition, DNA hybridization experiments have shown that organisms as widely divergent as *Drosophila* (11, 12), *Dictyostelium* (13), and *Saccharomyces* (14, 15) possess similar sequences. The same methodology was employed to demonstrate that the fission yeast *Schizosaccharomyces* contains DNA at a single locus that is highly related to *ras* (16). Interestingly, *Saccharomyces* has two genes (*RAS1* and *RAS2*), which encode proteins that are homologous to mammalian *ras* proteins (14, 15), and also another gene (*YP2*) encoding a protein that is less similar (17). Thus far, however, such observations have not been reported in bacteria.

We report here that *Escherichia coli* has a gene (*era*) that encodes a protein containing significant similarity to the yeast RAS proteins. Since mammalian and yeast *ras* gene products are known to bind GTP (18, 19), GTP cross-linking experiments were performed, and they revealed that the Era protein is a GTP-binding protein as well.

MATERIALS AND METHODS

Reagents and Materials. [α-^{32}P]dCTP (3000 Ci/mmol; 1 Ci = 37 GBq) and [α-^{32}P]rGTP (420 Ci/mmol) were supplied by Amersham. Restriction enzymes, T4 DNA ligase, and the large fragment of DNA polymerase I were supplied by New England Biolabs. Deoxy- and dideoxynucleotides were obtained from P-L Biochemicals.

DNA Sequencing. The *era* gene was obtained from the plasmid pSB, which contains a 4-kilobase (kb) *Eco*RI fragment obtained from the *E. coli* chromosome. This plasmid has the ability to complement genetic defects in a ribonuclease III-negative background (H.E.T. and D. Court, unpublished data). Fragments for sequencing were obtained by appropriate restriction enzyme digestion (as indicated in Fig. 1) followed by preparative polyacrylamide gel electrophoresis. The DNA fragments were cloned in sequencing vectors M13mp18 or M13mp19 by using the host strain JM103 (available from Bethesda Research Laboratories). DNA sequencing was performed by using the chain-termination method (23) except where indicated in Fig. 1, in which case the method of Maxam and Gilbert (24) was employed to sequence the cloned fragment.

Expression of the *rnc* Locus. The *rnc* locus was cloned in a high-expression vector under control of a *lac* inducer as described in *Results*. Such a plasmid harboring the *rnc* and *era* genes together (pJR002) was used to transform the *E. coli* K-12 strain SB221 (21). These cells were grown in M9–glucose medium supplemented with 0.2% Casamino acids, tryptophan at 50 μg/ml, and ampicillin at 50 μg/ml. Growth was at 37°C to early logarithmic phase (40 Klett units with a no. 808 filter), at which point the culture was divided. One-half of the culture received 2 mM isopropyl β-D-thiogalactoside (final concentration). After 2 hr of incubation, the cells were harvested by centrifugation and lysed by boiling in sample buffer [1% $NaDodSO_4$/10 mM sodium phosphate, pH 7.1/10% (vol/vol) glycerol/1% 2-mercaptoethanol/0.05% bromphenol blue] for 8 min. The samples were subjected to $NaDodSO_4$/polyacrylamide gel electrophoresis in a 17% gel as described previously (21).

GTP Photoaffinity Cross-Linking. The method employed to cross-link [α-^{32}P]rGTP to overexpressed GTP-binding proteins in *E. coli* crude cell extracts has been described in detail elsewhere (25, 26). Briefly, lysates were prepared by lysozyme treatment followed by detergent lysis. In these experiments SB221 harboring pJR002 were grown and induced with isopropyl thiogalactoside exactly as described above, and lysates were prepared upon harvesting the cells. Lysates were also prepared from SB221 containing the plasmid pJR001, which is identical to pJR002 except that excision of a 1.6-kb *Bam*HI fragment results in a deletion in the *era* gene. Thus pJR001 overexpresses ribonuclease III and not Era in the presence of isopropyl thiogalactoside.

The lysates were incubated in the dark for 1 hr at 0°C in the presence of 25 μCi of [α-^{32}P]rGTP and then exposed to ultraviolet irradiation at 254 nm for 30 min as described previously (25, 26). The lysates were subjected to $NaDodSO_4$/polyacrylamide gel electrophoresis (as above) followed by autoradiography, and the autoradiogram is presented in Fig. 5. Since label was not incorporated into the overexpressed ribonuclease III band it may be concluded that labeling is not an artifact of overexpression of protein that

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: kb, kilobase(s); bp, base pair(s); EF, elongation factor.

*Laboratory of Molecular Oncology, National Cancer Institute, Frederick, MD 21701-1013.

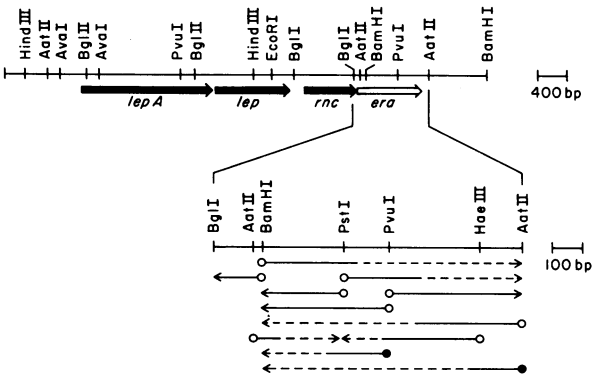

FIG. 1. Restriction map and sequencing strategy. The restriction map of the *era* gene and flanking regions is shown at the top. The thick black arrows indicate the position and orientation of the other known genes in this region (see refs. 20–22). The sequencing strategy employed is shown below the restriction map. Thin arrows with open circles indicate the restriction fragments that were employed to obtain sequence data by the chain-termination method (23). Thin arrows with the filled circles indicate restriction fragments that were sequenced by the method of Maxam and Gilbert (24). The broken lines represent regions of the cloned restriction fragments from which sequence data were not derived. bp, Base pairs.

otherwise would not bind GTP. Furthermore, under the conditions employed in this experiment we, and others, have not observed nonspecific labeling of proteins (25, 26).

RESULTS AND DISCUSSION

Identification of a Gene Downstream of the *rnc* Gene. Previously we have determined the DNA sequence of the gene (*rnc*) encoding ribonuclease III (20), which maps at 55.5 min on the *E. coli* chromosome (27). The DNA sequence demonstrated that there may be another cistron initiating at an ATG codon that overlaps with the termination codon (TGA) of the *rnc* gene (see Fig. 2 and ref. 20). To confirm this observation the restriction map of the downstream region was established (Fig. 1), and the DNA sequence of a 1-kb fragment adjacent to the *rnc* gene has been determined according to the strategy shown in Fig. 1.

As revealed in Fig. 2, the DNA sequence demonstrates that there is an open reading frame of 948 bp which initiates at the overlapping codons as previously predicted (20). The initiation codon is preceded by a Shine–Dalgarno consensus sequence (GGAG, see Fig. 2) for ribosome binding (28), indicating that the open reading frame would be translated not only by ribosomes reinitiating at the TGA–ATG overlap, but also by ribosomes initiating *de novo* upon recognition of the Shine–Dalgarno sequence. The open reading frame is terminated at a TAG codon at position 982 (Fig. 2) and therefore encodes a polypeptide of 316 amino acid residues (calculated molecular weight = 35,232).

Amino Acid Sequence Homology. The amino acid sequence deduced from the DNA sequence was compared to the sequences in the data base (3477 polypeptide files) maintained by the National Biomedical Research Foundation, using software available at the Molecular Biology Computer Research Resource.† Interestingly, this comparison revealed highly significant similarity between the predicted amino acid sequence shown in Fig. 2 and the yeast RAS proteins, which are apparent homologues of the mammalian *ras* oncogene family (15).

As shown in Fig. 3, the similarity between the predicted amino acid sequence and the yeast *RAS1* protein is found in

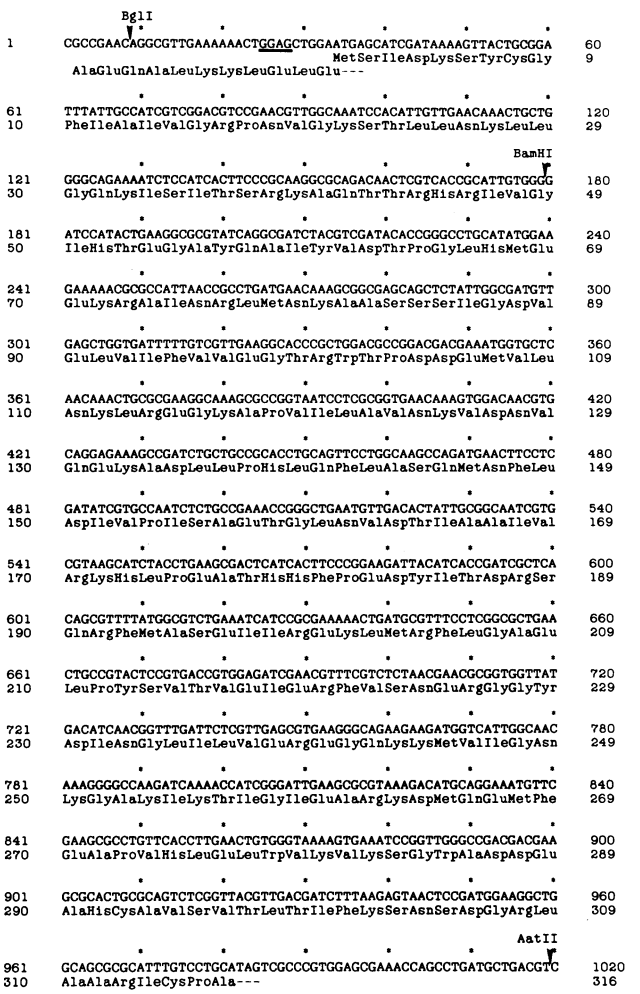

FIG. 2. DNA sequence of the *era* gene. The DNA sequence from the *Bgl* I site to the *Bam*HI site has been reported previously (20). The DNA sequence from the *Bam*HI site to the *Aat* II site was determined from both strands as shown in Fig. 1. The underline indicates the position of a probable ribosome binding site. The last 11 amino acid residues of ribonuclease III are shown (bp 1–34), followed by 316 amino acid residues resulting from the translation of the 948-bp open reading frame. It can be seen that the initiation codon of the *era* gene overlaps the termination codon of the *rnc* gene at bp 34–37.

an amino-terminal domain of 197 amino acid residues (196 amino acid residues for *RAS1*). Including conservative replacements, 91 positions can be aligned in this region for a similarity of 46% between these two proteins. The significance of the alignment presented in Fig. 3 is underscored by the observation that it was achieved with a minimum amount of insertion or gapping. Thus, the protein predicted from Fig. 2 is designated Era, for *E. coli* RAS-like protein, and the gene is designated *era*. Though only RAS1 is compared to Era in Fig. 3, an analogous alignment is obtained with *RAS2*, with which 42% of the residues match. The remaining carboxyl-terminal domain of 119 amino acid residues (113 in RAS1) exhibits only a very weak likeness, in which 19 positions can be aligned, resulting in a similarity of 16%, including conservative replacements. It is striking that the same two-domain structure has been observed in the yeast RAS1 and RAS2 proteins (15). There is extremely high similarity within the first 179 amino acid residues between RAS1 and RAS2, and almost no matching residues in the remaining carboxyl-terminal domain (consisting of 130 amino acid residues in RAS1 and 143 in RAS2), except for a short segment containing the last 8 amino acid residues. Alignment of the polypep-

†For information contact Dr. Temple Smith, Molecular Biology Computer Research Resource, Dana–Farber Cancer Institute, 44 Binney Street, Boston, MA 02115.

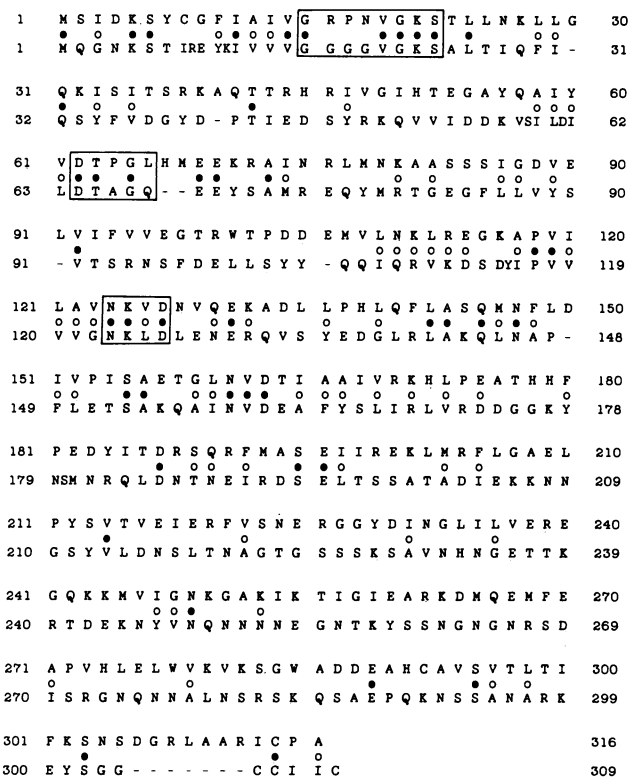

FIG. 3. Comparison of the amino acid sequences of the Era protein (upper sequence) and the yeast RAS1 protein (lower sequence). The filled circles denote identity between the amino acid sequences, and open circles indicate replacements by functionally related amino acids. The boxed residues indicate regions 1, 2, and 3 as shown in Table 1. These regions are common to all GTP-binding proteins. Region 1 is closest to the amino terminus, followed by 2 and 3, respectively. The yeast *RAS1* sequence data are from Powers *et al.* (15).

tides encoded by the yeast and mammalian *ras* genes demonstrates that this structural motif is a common feature of *ras* gene products (15). Thus they all possess a highly conserved domain of approximately 170–180 amino acids at their amino termini, followed by a sequence in which practically no similarity can be found. It should be noted, however, that mammalian *ras* genes encode only 188 or 189 amino acid residues, so their variable domain is much shorter than that encoded by the yeast *RAS* genes. In this sense Era (containing a total of 316 amino acid residues) is more like RAS1 (309 residues) and RAS2 (322 residues) than mammalian *ras* gene products.

The mammalian *ras* genes possess a cysteine residue at the fourth position from the carboxyl terminus through which the protein is bound to the cellular membrane by a lipid moiety (29). The importance of this cysteine residue was demonstrated by the observation that the ability to transform NIH3T3 cells was lost upon mutation of the cysteine codon in the Harvey murine sarcoma virus (29). Similarly, a cysteine residue was found in the equivalent position in the yeast RAS1 and RAS2 proteins (15). It is interesting that Era has a cysteine residue at the third position from the carboxyl terminus (Fig. 3). Chou–Fasman analysis (30) of potential secondary structure shows a very high probability for a β-turn structure at the amino-terminal side of the cysteine residue in RAS1, RAS2, and Era.

Overexpression and GTP-Binding of Era. GTP-binding proteins of *E. coli* share certain sequence determinants that were proposed to be important in binding guanine nucleotides (25, 31). Their importance has been confirmed in the case of elongation factor (EF)-Tu, for which x-ray crystallography has shown directly that these residues are intimately associated with the Mg-GDP binding pocket (32, 33). Similar sequences have been found in Era (boxed residues in Fig. 3), and both yeast and mammalian *ras* genes (see Table 1). These observations lead to the prediction that Era is a GTP-binding protein.

To examine the ability of Era to bind GTP we first attempted to identify the Era protein by using a high-expression vector containing the *era* gene. The two genes of the *rnc* operon were cloned together in a pIN-III($lpp^{P\text{-}5}$) vector [a high-expression vector developed in our laboratory (36)] by employing a *Hin*cII site located 24 bp upstream from the initiation codon of the *rnc* gene and the distal *Bam*HI site shown in Fig. 1. Thus expression of the *rnc* locus would be under the control of the strong $lpp^{P\text{-}5}$ promoter (36) as well as the *lac* promoter-operator instead of its own promoter. Furthermore, the production of the cloned gene products can be regulated by a *lac* inducer such as isopropyl β-D-thiogalactoside.

The proteins expressed from cells harboring this construction (pJR002) are shown in Fig. 4. After induction with isopropyl β-D-thiogalactoside for 2 hr the total cellular protein was separated by $NaDodSO_4$/polyacrylamide gel electrophoresis. Comparison to noninduced cells (Fig. 4, lane A) reveals that two bands are clearly produced specifically in the presence of isopropyl β-D-thiogalactoside (Fig. 4, lane B). Band R has a molecular weight of 26,000 and has been shown to be ribonuclease III, the *rnc* gene product (20). Band E has a molecular weight of 36,000, which is excellent agreement with predicted molecular weight for the Era protein (35,232). In pJR002 all of the Era protein is encoded on a 1.6-kb *Bam*HI fragment (see Fig. 1) except for the first 49 amino acid residues. When this fragment is excised from pJR002 the resultant construction (pJR001) expresses only band R (ribonuclease III) as predicted, demonstating that band E originates from the downstream sequence and confirming its identification as Era (data not shown). Fig. 4 indicates that Era and ribonuclease III are overproduced in an approximately 1:1 molar ratio, providing further evidence that they are cotranscribed, in support of the DNA sequence data.

The plasmids pJR001 and pJR002 were employed to investigate the ability of Era to bind GTP. We and others have utilized a photoaffinity labeling technique to demonstrate specific binding of [α-^{32}P]GTP to proteins in a crude cell lysate (25, 26). This method was employed to show that cells harboring pJR002 (Fig. 5, lane A), but not cells harboring pJR001 (lane B), produced a protein in the presence of isopropyl β-D-thiogalactoside that was very efficiently cross-linked by [α-^{32}P]GTP. As indicated in Fig. 5, this labeled species exactly comigrated with band E shown in Fig. 4, These data provide clear evidence that Era is a GTP-binding protein, since the only difference between pJR001 and pJR002 is the 1.6-kb *Bam*HI fragment, which is required for expression of Era as described above. In the absence of this fragment (pJR001) Era is not expressed and the ^{32}P-labeled species indicated in Fig. 5 is absent (compare lanes A and B).

Subcellular fractionation experiments revealed that most of the overproduced Era exists in the soluble fraction, while approximately 20–30% is associated with the membrane fraction (data not shown).

Possible Functions of Era. At present the function of Era is unknown. However, the Era protein is distinctly different from other *E. coli* GTP-binding proteins identified to date, including the LepA protein, which is a membrane-bound GTP-binding protein (25). These GTP-binding proteins all possess three well-defined regions (Table 1), which were shown by x-ray crystallographic analysis of EF-Tu to be closely associated with the bound guanine nucleotide (33). Region 1 is in close association with the phosphate groups, and in particular the well-conserved lysine residue may be

Table 1. Comparison of amino acid sequence in GTP-binding proteins

Protein	Sequence
	Region 1
E. coli EF-Tu	-Gly-His-Val-Asp-His-Gly-Lys-Thr- (18–25)
E. coli EF-G	-Ala-His-Ile-Asp-Ala-Gly-Lys-Thr- (17–24)
E. coli LepA	-Ala-His-Ile-Asp-His-Gly-Lys-Ser- (11–18)
E. coli IF2	-Gly-His-Val-Asp-His-Gly-Lys-Thr- (398–405)
S. cerevisiae RAS1	-Gly-Gly-Gly-Gly-Val-Gly-Lys-Ser- (17–24)
S. cerevisiae RAS2	-Gly-Gly-Gly-Gly-Val-Gly-Lys-Ser- (17–24)
Human Ha-ras, N-ras, Ki-ras	-Gly-Ala-Gly-Gly-Val-Gly-Lys-Ser- (10–17)
E. coli Era	-Gly-Arg-Pro-Asn-Val-Gly-Lys-Ser- (15–22)
	Region 2
E. coli EF-Tu	-Asp-Cys-Pro-Gly-His- (80–84)
E. coli EF-G	-Asp-Thr-Pro-Gly-His- (88–92)
E. coli LepA	-Asp-Thr-Pro-Gly-His- (77–81)
E. coli IF2	-Asp-Thr-Pro-Gly-His- (444–448)
S. cerevisiae RAS1	-Asp-Thr-Ala-Gly-Gln- (64–68)
S. cerevisiae RAS2	-Asp-Thr-Ala-Gly-Gln- (64–68)
Human Ha-ras, N-ras, Ki-ras	-Asp-Thr-Ala-Gly-Gln- (57–61)
E. coli Era	-Asp-Thr-Pro-Gly-Leu- (62–66)
	Region 3
E. coli EF-Tu	-Asn-Lys-Cys-Asp- (135–138)
E. coli EF-G	-Asn-Lys-Met-Asp- (142–145)
E. coli LepA	-Asn-Lys-Ile-Asp- (131–134)
E. coli IF2	-Asn-Lys-Ile-Asp- (498–501)
S. cerevisiae RAS1	-Asn-Lys-Leu-Asp- (123–126)
S. cerevisiae RAS2	-Asn-Lys-Ser-Asp- (123–126)
Human Ha-ras, N-ras, Ki-ras	-Asn-Lys-Cys-Asp- (116–119)
E. coli Era	-Asn-Lys-Val-Asp- (124–127)

The numbers in parentheses following the amino acid sequences represent the amino acid residue numbers. The references for the sequence data are EF-Tu (34), EF-G (35), *lepA* product (LepA) (21), initiation factor 2 (IF2) (31), RAS1 and RAS2 (15), Ha-ras (9), N-ras (8), and Ki-ras (10).

interacting with the electron-rich oxygen atoms. Region 2 is close to region 1, and apparently the conserved aspartic acid residue forms a salt linkage to the Mg^{2+} that is required for the GTPase activity of EF-Tu. Region 3 was shown to be intimately involved with the guanine moiety itself, again involving well-conserved residues; in this case the asparagine and aspartic acid residues (see Table 1) were important in direct interactions with the guanine group.

It is clear from Table 1 that not only the *E. coli* GTP-binding proteins but all of the *ras* gene products share these properties (37). There are, however, important differences that distinguish *ras*-like gene products from other GTP-binding proteins. Region 1, for instance, contains the highly conserved sequence -Gly(or Ala)-Xaa-Zaa-Xaa-Xaa-Gly-Lys-Ser(or Thr)-. The Xaa residues are variable, and it has been well established that Zaa is a glycine residue in all human *ras* protooncogene products as well as yeast *RAS1* and *RAS2* products (see Table 1 and refs. 8–10, 15, 38–43). It is important to note that the only amino acid which can replace

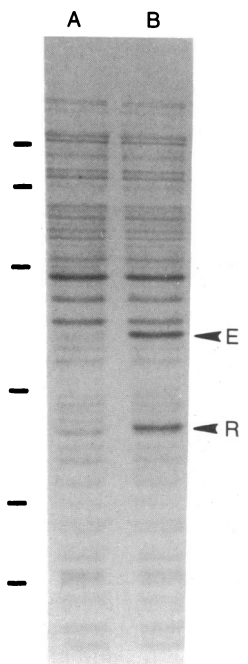

FIG. 4. Expression of the proteins of the *rnc* locus. *E. coli* cells harboring the plasmid pJR002 were grown as described in the text. The sample in lane A was prepared from cells that were incubated in the absence of isopropyl thiogalactoside, and the sample in lane B was prepared from cells that were incubated in the presence of this inducer. Equal culture equivalents were loaded in each lane and after electrophoresis the gel was stained with Coomassie blue. The production of bands E (molecular weight = 36,000) and R (molecular weight = 26,000) is clearly induced by isopropyl thiogalactoside. The positions of molecular weight markers are indicated by the dashes and represent molecular weights of 92,000, 66,000, 45,000, 31,000, 21,000, and 14,000.

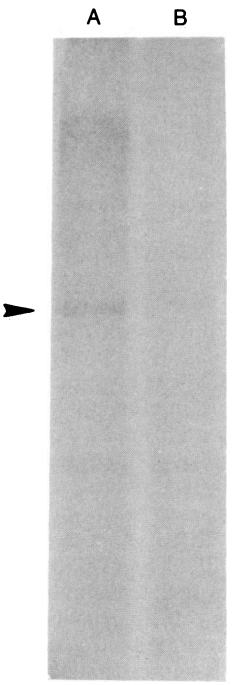

FIG. 5. GTP cross-linking experiments. Cells harboring either pJR001 or pJR002 were induced with isopropyl thiogalactoside as in Fig. 4. Whole cell lysates were prepared, mixed with [α-^{32}P]rGTP, and photolyzed as described in *Materials and Methods*. After NaDodSO$_4$/polyacrylamide gel electrophoresis of the cell lysates, the gel was dried and subjected to autoradiography for 6 hr. Lane A, lysate from cells harboring pJR002; lane B, lysate from cells harboring pJR001. The arrow marks the position of migration of the band marked E in Fig. 4.

this glycine residue and not lead to oncogenic activity is proline (44), which is the residue at position Zaa in Era. The other *E. coli* GTP-binding proteins possess either valine or isoleucine at position Zaa (Table 1).

A major difference that separates Era from all the other *E. coli* GTP-binding proteins can be found in the sequence surrounding region 3. It has been pointed out previously that *E. coli* GTP-binding proteins possess similar sequences in the positions about 4 to 8 amino acid residues immediately before region 3, and no similarity after region 3 (25, 31). This is true when Era is compared with the other *E. coli* GTP-binding proteins as well. However, when Era is compared with the *ras*-related proteins, or *ras*-related proteins are compared among themselves, the similarity around region 3 is found to be much more extensive. It can be seen upon comparison of Era to the yeast RAS1 protein (Fig. 3) that region 3 is embedded in a very extensive region of similarity, which includes residues 109–197 of Era and 107–196 of yeast RAS1. Within this domain of 88 residues there are a total of 52 residues, including conservative replacements, which can be aligned (overall similarity equals 59%). The spacing of this alignment is nearly perfect in the sense that only one deletion and two insertions are required to achieve the best fit. These observations, coupled with those described above—namely, the similar overall size, the domain structure, and position of the cysteine near the carboxyl terminus—make it most likely that Era is a member of the family of *ras*-related proteins with significant similarity to the yeast *RAS1* protein.

In yeast the function of the *RAS* gene products is not completely understood. However, they have been implicated in the regulation of adenylate cyclase (45, 46) and been shown to be required for gluconeogenic growth (47). Experiments designed to investigate whether or not Era has similar functions will definitively establish the relatedness of these proteins, and the importance of such proteins in the growth of cells. In this regard it will be particularly interesting to examine the effects on cell physiology of substitution of proline at position 17 (corresponding to glycine 12 encoded by mammalian *ras* genes and glycine 19 encoded by yeast *RAS* genes) with other amino acids, and to investigate whether Era is able to complement yeast *RAS* gene mutations.

We gratefully acknowledge Dr. T. Smith for his assistance at the Molecular Biology Computer Research Resource. We also thank Dr. M. Perucho and Dr. D. Comeau for critical reading of the manuscript. This work was supported by Grant GM19043 from the National Institute of General Medical Sciences and by Grant NB3871 from the American Cancer Society. P.E.M. is the recipient of a National Institutes of Health Postdoctoral Fellowship. H.E.T. is a recipient of a Burroughs–Wellcome Senior Research Fellowship.

1. Land, H., Parada, L. F. & Weinberg, R. A. (1983) *Science* **222**, 771–778.
2. Varmus, H. E. (1984) *Annu. Rev. Genet.* **18**, 553–612.
3. Der, C. J., Krontiris, T. G. & Cooper, G. M. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 3637–3640.
4. Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. (1982) *Nature (London)* **297**, 474–478.
5. Santos, E., Tronick, S. R., Aaronson, S. A., Pulciani, S. & Barbacid, M. (1982) *Nature (London)* **298**, 343–347.
6. Shimizu, K., Goldfarb, M., Suard, Y., Perucho, M., Li, Y., Kamata, T., Feramisco, J., Stavnezer, E., Fogh, J. & Wigler, M. H. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2112–2116.
7. Hall, A., Marshall, C. J., Spurr, N. K. & Weiss, R. A. (1983) *Nature (London)* **303**, 396–400.
8. Taparowsky, E., Shimizu, K., Goldfarb, M. & Wigler, M. (1983) *Cell* **34**, 581–586.
9. Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H. & Goeddel, D. V. (1983) *Nature (London)* **302**, 33–37.
10. McGrath, J. P., Capon, D. J., Smith, D. H., Chen, E. Y., Seeburg, P. H., Goeddel, D. V. & Levinson, A. D. (1983) *Nature (London)* **304**, 501–506.
11. Shilo, B. & Weinberg, R. A. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 6789–6792.
12. Neuman-Silberberg, F. S., Schejter, E., Hoffman, F. M. & Shilo, B. (1984) *Cell* **37**, 1027–1033.
13. Reymond, C. D., Gomer, R. H., Mehdy, M. C. & Firtel, R. A. (1984) *Cell* **39**, 141–148.
14. DeFeo-Jones, D., Scolnick, E. M., Koller, R. & Dhar, R. (1983) *Nature (London)* **306**, 707–709.
15. Powers, S., Kataoka, T., Fasano, O., Goldfarb, M., Strathern, J., Broach, J. & Wigler, M. (1984) *Cell* **36**, 607–612.
16. Fukui, Y. & Kaziro, Y. (1985) *EMBO J.* **4**, 687–691.
17. Gallwitz, D., Donath, C. & Sander, C. (1983) *Nature (London)* **306**, 704–707.
18. Shih, T. Y., Papageorge, A. G., Stokes, P. E., Weeks, M. O. & Scolnick, E. M. (1980) *Nature (London)* **287**, 686–691.
19. Temeles, G. L., Gibbs, J. B., D'Alonzo, J. S., Sigal, I. S. & Scolnick, E. M. (1985) *Nature (London)* **313**, 700–703.
20. March, P. E., Ahnn, J. & Inouye, M. (1985) *Nucleic Acids Res.* **13**, 4677–4685.
21. March, P. E. & Inouye, M. (1985) *J. Biol. Chem.* **260**, 7206–7213.
22. Wolfe, P. B., Wickner, W. & Goodman, J. M. (1983) *J. Biol. Chem.* **258**, 12073–12080.
23. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
24. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
25. March, P. E. & Inouye, M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7500–7504.
26. Stein, R. B., Robinson, P. S. & Scolnick, E. M. (1984) *J. Virol.* **50**, 343–351.
27. Studier, F. W. (1975) *J. Bacteriol.* **124**, 307–316.
28. Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1342–1346.
29. Willumsen, B. M., Norris, K., Papageorge, A. G., Hubbert, N. L. & Lowy, D. R. (1984) *EMBO J.* **3**, 2581–2585.
30. Chou, P. Y. & Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47**, 251–276.
31. Sacerdot, C., Dessen, P., Hershey, J. W. B., Plumbridge, J. A. & Grunberg-Manago, M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7787–7791.
32. la Cour, T. F. M., Nyborg, J., Thirup, S. & Clark, B. F. C. (1985) *EMBO J.* **4**, 2385–2388.
33. Jurnak, F. (1985) *Science* **230**, 32–36.
34. Yokota, T., Sugisaki, H., Takanami, M. & Kaziro, Y. (1980) *Gene* **12**, 25–31.
35. Zengel, J. M., Archer, R. H. & Lindahl, L. (1984) *Nucleic Acids Res.* **12**, 2181–2192.
36. Inouye, S. & Inouye, M. (1985) *Nucleic Acids Res.* **13**, 3101–3110.
37. McCormick, F., Clark, B. F. C., la Cour, T. F. M., Kjeldgaard, M., Norskov-Lauritsen, L. & Nyborg, J. (1985) *Science* **230**, 78–82.
38. Tabin, C. J., Bradley, S. M., Bargmann, C. I., Weinberg, R. A., Papageorge, A. G., Scolnick, E. M., Dhar, R., Lowy, D. R. & Chang, E. H. (1982) *Nature (London)* **300**, 143–149.
39. Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. (1982) *Nature (London)* **300**, 149–152.
40. Taparowsky, E., Suard, Y., Fasano, O., Shimizu, K., Goldfarb, M. & Wigler, M. (1982) *Nature (London)* **300**, 762–765.
41. Yuasa, Y., Srivastava, S. K., Dunn, C. Y., Rhim, J. S., Reddy, E. P. & Aaronson, S. A. (1983) *Nature (London)* **303**, 755–779.
42. Shimizu, K., Birnbaum, D., Ruley, M. A., Fasano, O., Suard, Y., Edlund, L., Taparowsky, E., Goldfarb, M. & Wigler, M. (1983) *Nature (London)* **304**, 497–500.
43. Capon, D. J., Seeburg, P. H., McGrath, J. P., Hayflick, J. S., Edman, U., Levinson, A. D. & Goeddel, D. V. (1983) *Nature (London)* **304**, 507–513.
44. Seeburg, P. H., Colby, W. W., Capon, D. J., Goeddel, D. V. & Levinson, A. D. (1984) *Nature (London)* **312**, 71–65.
45. Uno, I., Mitsuzawa, H., Matsumoto, K., Tanaka, K., Oshima, T. & Ishikawa, T. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7855–7859.
46. Broek, D., Samiy, N., Fasano, O., Fujiyama, A., Tamanoi, F., Northup, J. & Wigler, M. (1985) *Cell* **41**, 763–769.
47. Tatchell, K., Robinson, L. C. & Breitenbach, M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3785–3789.