Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 4978–4982, July 1986
Biochemistry

Amino acid sequence of mammalian elongation factor 2 deduced from the cDNA sequence: Homology with GTP-binding proteins

(cloning/diphtheria toxin/ADP-ribosylation/ras proteins/ribosome)

KENJI KOHNO*, TSUYOSHI UCHIDA*†, HIROAKI OHKUBO‡, SHIGETADA NAKANISHI‡, TOHRU NAKANISHI*, TOSHIKAZU FUKUI§, EIKO OHTSUKA§, MORIO IKEHARA§, AND YOSHIO OKADA†

*Department of Cell Biology, National Institute for Basic Biology, Okazaki, Aichi 444, Japan; ‡Kyoto University Faculty of Medicine, Kyoto 606, Japan; §Faculty of Pharmaceutical Sciences, Osaka University, Suita, Osaka 565, Japan; and †Institute for Molecular and Cellular Biology, Osaka University, Suita, Osaka 565, Japan

Communicated by A. M. Pappenheimer, February 26, 1986

ABSTRACT **Complementary DNA clones, pHEW1 and pRE2, coding for hamster and rat polypeptide chain elongation factor 2 (EF-2), respectively, were isolated and sequenced. It was shown that the cDNA insert in pHEW1 contains a 2574-base-pair open reading frame coding for an 857-amino acid polypeptide with M_r 95,192, excluding the initiation methionine. Comparative studies of sequence homology among EF-2 and several GTP-binding proteins show that five regions in the amino-terminal position of EF-2, corresponding to about 160 amino acids, show homology with GTP-binding proteins, including protein synthesis elongation and initiation factors, mammalian ras proteins, and transducin. The carboxyl-terminal half of EF-2 contains several regions that have 34–75% homology with bacterial elongation factor G. These results suggest that the amino-terminal region of EF-2 participates in the GTP-binding and GTPase activity whereas the carboxyl-terminal region interacts with ribosomes. Finally, the sequence provides direct evidence that diphthamide {2-[3-carboxyamido-3-(trimethylammonio)propyl]histidine}, the site of ADP-ribosylation by diphtheria toxin, is produced by post-translational modification of a histidine residue in the primary translational product.**

Elongation factor 2 (EF-2), a single protein of M_r 100,000, is essential for the polypeptide chain elongation steps in eukaryotic protein synthesis (1). EF-2 catalyzes the GTP hydrolysis-dependent translocation of peptidyl-tRNA from the aminoacyl site to the peptidyl site on the ribosome. EF-2, which is a GTP-binding protein, forms a ternary complex with GTP and ribosomes. Therefore, EF-2 has at least two functional sites, one interacting with guanine nucleotides and one with ribosomes. Eukaryotic EF-2 is specifically ADP-ribosylated (and inactivated) in the presence of NAD by fragment A of diphtheria toxin (DT) (2, 3) or by *Pseudomonas aeruginosa* exotoxin A (PA) (4). To examine the relationships between the structure and the functions of EF-2, the following several approaches have been used: the purification of EF-2 and analysis of its properties (5–7, 12) and the isolation of DT- and PA-resistant cells containing non-ADP-ribosylatable EF-2 was (8, 9). We have also isolated and characterized DT-resistant CHO mutant cells, and purified variant EF-2 was identified by two-dimensional gel electrophoresis from these cells (10, 11). As the next step in these studies, we wished to isolate a cDNA clone encompassing hamster EF-2; however, since the amino acid sequence of hamster EF-2 has not yet been determined, rat EF-2 cDNA was cloned first using a mixture of synthetic oligonucleotides.

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.

Here we report the cDNA sequence encoding the entire hamster EF-2 region and the deduced 857 amino acid sequence. A comparison of the sequences among EF-2 and other GTP-binding proteins shows striking homologies.

MATERIALS AND METHODS

Two mixtures of 16 tetradecamer oligodeoxyribonucleotides constructed from the amino acid sequence of a peptide released by tryptic digestion of rat EF-2 specifically ADP-ribosylated by DT (12) were used to screen a rat liver cDNA library, kindly provided by S. Furuto. These tetradecamers represent all the possible complementary sequences, 5′-TC$^{A}_{G}$TG$^{A}_{T}$AC$^{A}_{G}$TC$^{A}_{G}$AA-3′ and 5′-TC$^{A}_{G}$TG$^{G}_{C}$AC$^{A}_{G}$TC$^{A}_{G}$AA-3′, encoding the pentapeptide sequence Phe-Asp-Val-His-Asp of rat EF-2 (excluding the third nucleotide residue of the terminal aspartic acid). Transformed *Escherichia coli* HB 101 clones replicated on nitrocellulose filters (HATF, Millipore) were screened by hybidization at 36°C with 5′-^{32}P-labeled synthetic oligonucleotides (specific activity, $>10^8$ cpm/μg) (13). The hybridized filters were washed at 36°C in 4× SSC containing 0.1% $NaDodSO_4$ (1× SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0). Positive clones were isolated and analyzed.

To obtain a full-sized hamster cDNA, a hamster cDNA library was constructed following the Okayama and Berg method (14) using a new vector pcD that permits the expression of cDNA in mammalian cells. Total RNA was isolated by the guanidinium thiocyanate method from logarithmically growing CHO-K1 cells (15), and purified once by oligo(dT)-cellulose chromatography (P-L Biochemicals, Type 7). Ten micrograms of poly(A)$^+$ RNA and 5.6 μg of vector-primer were used to prepare the cDNA library. *E. coli* HB101 was transformed, and clones were selected for ampicillin resistance. About 70,000 transformants were screened by colony hybridization with two probes prepared from the rat plasmid pRE2—the 576-base-pair (bp) *Rsa* I fragment A and the 161-bp *Cla* I–*Ava* II fragment B (Fig. 1). The filters were washed at 60°C in 6× SSC/0.1% $NaDodSO_4$, and 25 positive clones hybridizing with the two probes were isolated and analyzed. The appropriate restriction fragments were inserted into M13mp 18 and 19 vectors, and the DNA was sequenced using the chain-termination method of Sanger *et al.* (16). DNA sequences were analyzed by the DNASIS system (Hitachi Software Engineering). All of the cloning procedures were conducted in accordance with the guidelines for research involving recombinant DNA molecules issued by the Ministry of Education, Science and Culture of Japan.

Abbreviations: EF, eukaryotic elongation factor; bp, base pair(s); DT, diphtheria toxin; PA, *Pseudomonas aeruginosa* exotoxin A.

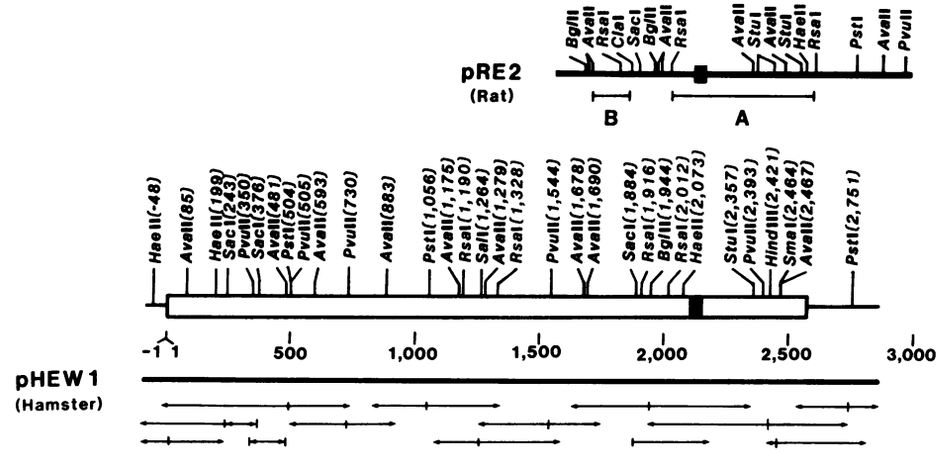

FIG. 1. Restriction map of EF-2 cDNA from rat (pRE2) and hamster (pHEW1). The numbers at each restriction site in hamster cDNA indicate which 5'-terminal nucleotide is generated by cleavage. The scale under the map is in bp. The poly(dA)·poly(dT) and poly(dG)·poly(dC) tails are not included in these maps. The black box indicates the tryptic peptide ADP-ribosylated by DT. Fragments A and B in pRE2 represent respectively the 576-bp *Rsa* I fragment and the 161-bp *Ava* II–*Cla* I fragment, which were used as probes to screen for pHEW1. The horizontal arrows indicate the direction and the extent of sequence determination using Sanger's method of dideoxy chain termination with the M13 universal primer.

RESULTS AND DISCUSSION

Isolation of Rat and Hamster EF-2 cDNAs. Two mixtures of 16 tetradecamer synthetic oligonucleotides were used to screen a rat liver cDNA library. About 50,000 transformants were screened, and 5 positive clones were obtained. Southern blot and partial sequence analyses showed that one clone, designated pRE2, contained DNA sequences complementary to rat EF-2 mRNA and encoding the 15 amino acids corresponding to residues 702–716, the peptide produced by trypsin digestion of EF-2 (12) (Fig. 2). Fig. 1 shows the restriction map of the cDNA insert in pRE2, which is approximately 1430 bp excluding the poly(dA)·poly(dT) and the poly(dG)·poly(dC) tails. To examine the length of the EF-2 mRNA derived from rat and hamster cells, poly(A)$^+$ RNA blot hybridization analysis was performed using the rat 576-bp *Rsa* I fragment A as a probe (Fig. 1). The length of the EF-2 mRNA was approximately 3 kilobases in each case, which is sufficient to encode a protein of about 100 kDa (data not shown).

Both the *Rsa* I fragment A and the *Ava* II–*Cla* I fragment B from rat EF-2 cDNA were used to screen a CHO cDNA library to isolate hamster EF-2 cDNA. Twenty-five positive clones hybridizing to probes A and B were isolated from about 70,000 transformants, and they all showed similar restriction maps. The longest insert, found in clone pHEW1, was sequenced.

Nucleotide Sequence Analysis. Fig. 1 shows the restriction map of hamster EF-2 cDNA derived from pHEW1 and the sequencing strategy used in this case. Analysis of the complete nucleotide sequence of the cloned cDNA revealed that it contained one long open reading frame coding for 858-amino acid residues followed by a TAG termination codon and that the other frames were interrupted by multitermination codons. As shown in Fig. 2, the cDNA insert in pHEW1 contains a 77-bp 5'-untranslated region, a 2574-bp coding sequence, and a 282-bp 3'-untranslated region. The translational initiation site was assigned to the first methionine codon ATG at position 1–3, which is in agreement with the following results: (*i*) the sequence CCACTATG (positions −5 to +3) is a well conserved region containing the functional initiation sequence proposed by Kozak (18), and (*ii*) the 19 amino-terminal amino acids of rat EF-2 sequenced from partial tryptic fragments were identical to those of hamster EF-2 between positions 2 and 20 (17). The 5'-untranslated region of EF-2 cDNA is characterized by the following: (*i*) a high G+C content (70%) compared with that of other regions (56%); (*ii*) the unique sequence 5'-CCAUCCGC-3' (−12 to −5) in the 5'-proximal AUG region of EF-2 mRNA, which is complementary except for a single mismatch to the sequence 5'-GCGGAAGG-3' in the 3' end of 18S rRNA (19, 20); and (*iii*) a 6-bp tandem repeat between positions −74 and −60 (GCCGCAGCCGCAGCC). On the other hand, the 3'-untranslated region is characterized by an additional poly(A)$^+$ tail signal, 5'-AATAAA-3', located 12 nucleotides upstream from the poly(A)$^+$ tail.

Protein Primary Structure of Hamster EF-2. The primary structure of hamster EF-2 deduced from the cDNA pHEW1 nucleotide sequence is shown in Fig. 2. Hamster EF-2 contains 857 amino acids (excluding the initiation methionine) and the calculated M_r is 95,192. The amino acid sequence of hamster EF-2 has not been determined, but partial sequences of EF-2 from other species have been reported. As in rat and chicken EF-2 (5, 6), the amino-terminal residue of hamster is valine. Nineteen residues of the amino-terminal region and 1 residue of the carboxyl-terminal region of rat EF-2 have been identified by Takamatsu *et al.* (17). Furthermore, a 15-amino acid peptide produced by tryptic digestion of ADP-ribosylated rat and bovine EF-2 has been sequenced (12, 21). All these sequences are present in hamster EF-2. Van Ness *et al.* (22) reported that the amino acid specifically ADP-ribosylated by DT is 2-[3-carboxyamido-3-(trimethylammonio)propyl]histidine, designated diphthamide, which is an unusual amino acid modified posttranslationally. The hamster EF-2 sequence analysis provides direct evidence that the modified amino acid is a histidine residue. Furthermore, the rat cDNA insert in pRE2 encodes a sequence of 343 amino acids, located in the carboxyl-terminal region of EF-2 (residues 516–858), all of which are conserved in hamster EF-2. However, the nucleotide sequences of rat and hamster EF-2 cDNAs showed only 90% homology, the differences being mostly due to base replacements at the third position in the codons (Fig. 2). Finally, the amino acid composition of hamster EF-2 is similar to that reported for rat liver, pig liver, rabbit reticulocyte, and hen oviduct EF-2 (5–7).

Homology with GTP-Binding Proteins. It has been reported that mammalian ras proteins and bacterial elongation factors share similarities of sequence and of GTP-binding and GTPase activities (23). The amino acid sequence of EF-2 was compared with the following GTP-binding proteins: bacterial elongation factors EF-Tu (24), EF-G (25), yeast elongation

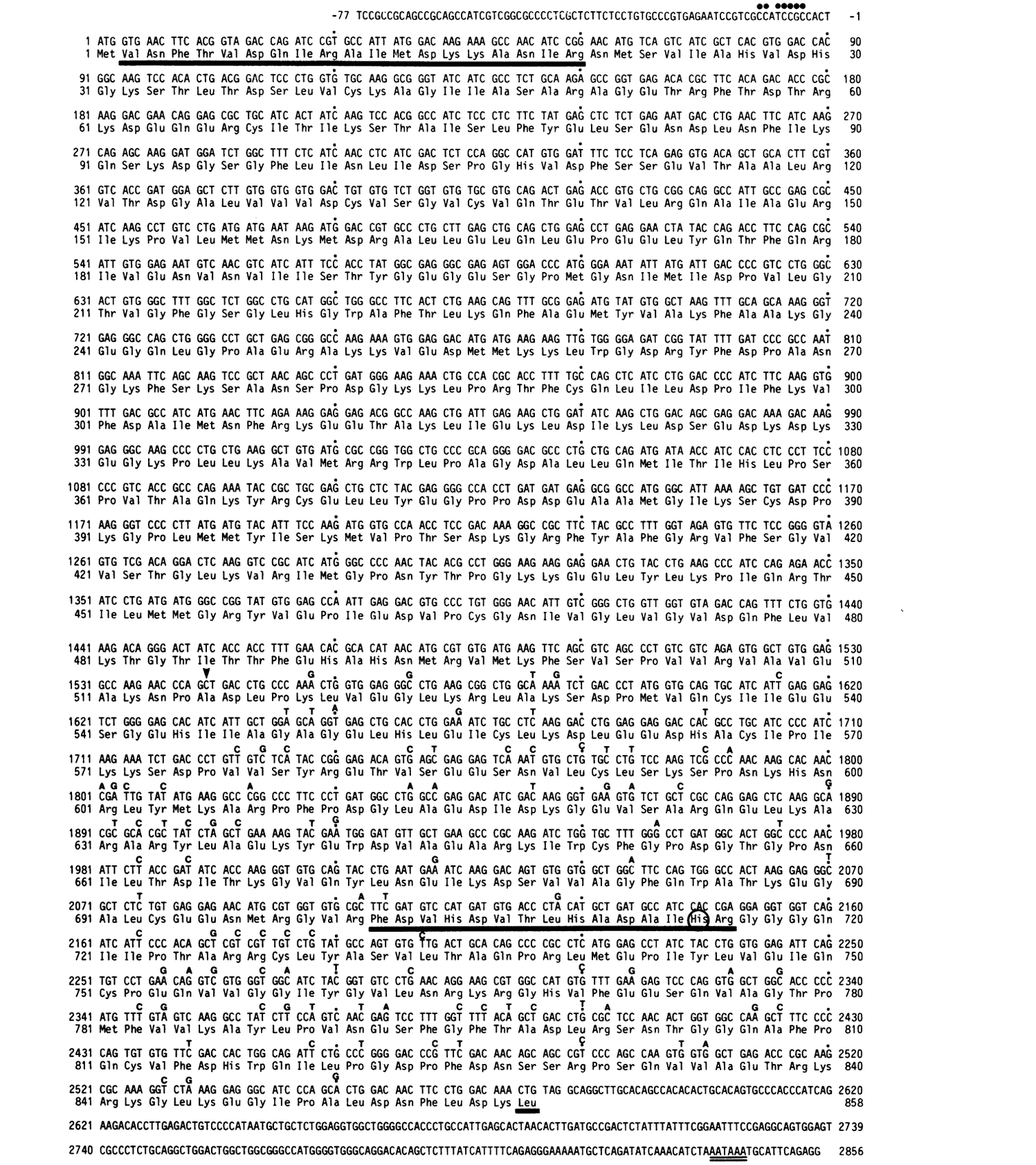

FIG. 2. Complete nucleotide and deduced amino acid sequence of the cDNA insert in pHEW1 encoding hamster EF-2 cDNA. The nucleotide and the amino acid residues are numbered from the beginning of the initiation codon and from the initiation methionine, respectively. The underlined amino acids (positions 2–20, 702–716, and 858) are the ones that completely matched those of rat EF-2 identified by peptide analysis of purified EF-2 (12, 17). The arrowhead denotes the 5′ end of the rat cDNA insert in pRE2. The bases indicated above the coding region of hamster cDNA show where the nucleotide sequence of rat EF-2 cDNA differs from that of hamster EF-2 cDNA. The sequence marked with solid circles in the 5′-proximal ATG region shows the sequence complementary to that of the 3′ end of 18S rRNA. Histidine-715 marked with a circle is the target amino acid ADP-ribosylated by DT. The AATAAA box close to the polyadenylylated 3′ end of the mRNA is doubly underlined.

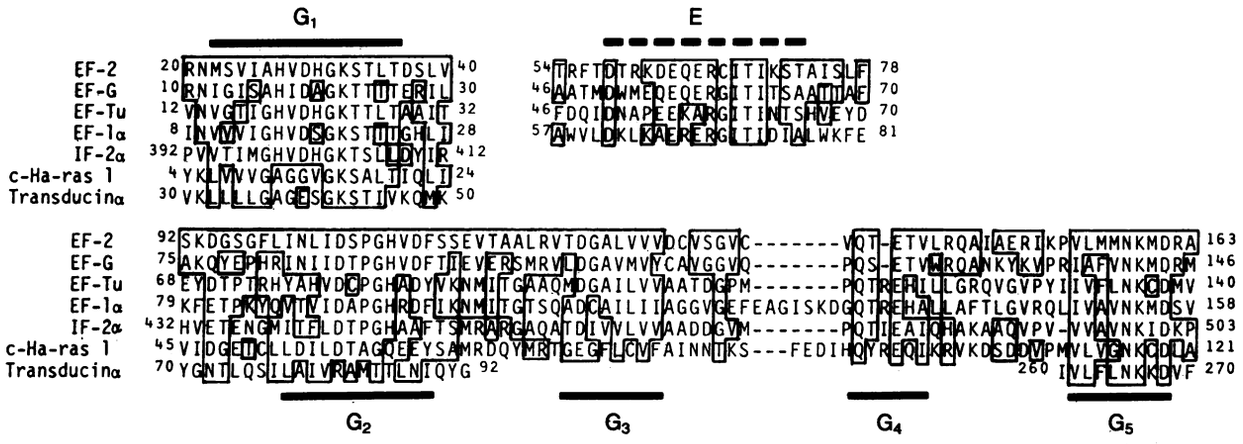

FIG. 3. Comparison of amino acid sequences of EF-2 with other GTP-binding proteins, including bacterial elongation factors EF-G (25) and EF-Tu (24), yeast elongation factor EF-1α (26), bacterial initiation factor IF-2α (27), human ras protein c-Ha-ras 1 (28), and bovine transducin (29). The regions G_1–G_5 indicated by horizontal solid lines contain common homologous regions among the various GTP-binding proteins, whereas the region E indicated by a dashed line contains homologous regions with elongation factors only. The following Dayhoff conservative categories (30) were used: C; A, G, P, S, and T; R, H, and K; N, D, Q, and E; I, L, M, and V; F, W, and Y.

factor EF-1α (26), bacterial initiation factor IF-2α (27), human ras protein (28), and bovine transducin (29). Fig. 3 shows that six highly homologous regions are present in all these proteins. In region G_1, amino acids between positions methionine-22 and threonine-36 of EF-2 align with the sequence Gly/Ala-His-Val-Asp-Xaa-Gly-Lys-Ser/Thr, which is a consensus sequence present in elongation and initiation factors and which is partly conserved as Gly-Xaa-Xaa-Xaa-Xaa-Gly-Lys-Ser in ras proteins and in transducin. Glycine-12, an important amino acid in ras proteins related to malignant transformation (31), is present in this region. Amino acid substitution of glycine by valine in ras proteins caused no major alterations in the binding of guanine nucleotides, although it reduced markedly GTPase activity (31). This result suggests that region G_1 is more important for GTP hydrolysis than for GTP binding. It is interesting to observe that the amino acid corresponding to glycine-12 in ras proteins is replaced by valine or isoleucine in elongation and initiation factors (Fig. 3) and that this substitution in ras proteins strongly increases their potent transforming ability (31). As in G_1, there is a conserved sequence in G_2 that contains an amino acid related to the malignant properties of ras proteins. Substitution of alanine-59 by threonine in ras proteins greatly reduces their GTP-binding activity but activates their transforming ability (32). Studies on elongation factors, especially EF-Tu, indicated that regions G_3, G_4, and G_5 are domains involved in the binding of guanine nucleotides. Using a thiol group modification technique for cysteine residues, it was found that cysteine-137 in region G_5 of EF-Tu is located near the guanine nucleotide binding site (24). Also, in EF-G, the GTP-binding site has been localized to the tryptic peptide from glycine-59 to arginine-127 containing cysteine-113, by photoaffinity labeling (33). Among all of the GTP-binding proteins compared in Fig. 3, the most conserved sequence, Asn-Lys-Xaa-Asp (residues 158–161) in EF-2, is in the region G_5. These amino acids are considered to play important roles in GTP and GDP binding, which is supported by the x-ray crystallographic data of la Cour *et al.* (34).

Bacterial elongation factor EF-G, whose function corresponds to eukaryotic elongation factor EF-2, shares additional homologous sequences with EF-2 as shown in Fig. 4. Five regions are homologous, the homology ranging from 34 to 75%. As expected, since EF-G is not ADP-ribosylated by DT, the region containing histidine-715 in EF-2 shows no homology in EF-G (Fig. 4). This region is required for DT-dependent ADP-ribosylation of EF-2. ADP-ribosylation of EF-2, catalyzed by DT in the presence of NAD, reduces the affinity for the ribosome, leading to an inhibition of protein synthesis (35). Therefore, the carboxyl-terminal second region of EF-2 appears to interact with ribosomes during the peptide chain elongation step in eukaryotes. Further, since EF-2 can weakly bind with fragment A of DT (17), the carboxyl-terminal domain of EF-2 must interact with the fragment A molecules to ADP-ribosylate. Region E shown in Fig. 3 is located in the GTP-binding region of EF-2, but it shows homology only among elongation factors. This region might interact with peptidyl tRNA and/or ribosomes rather than guanine nucleotides.

Fig. 5 proposes a diagram for the functional regions of the primary structure of EF-2. The amino-terminal region of EF-2 encompassing the first 160 amino acids excluding the region E is considered to contain the domains essential for GTP-binding and GTPase activities, whereas the carboxyl-terminal half is considered to contain the domain involved in

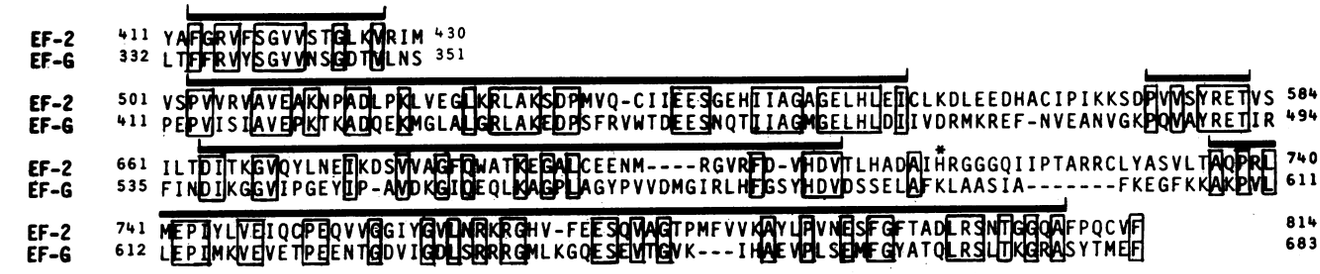

FIG. 4. Amino acid sequences of EF-2 and EF-G from the homologous regions aligned to obtain maximum matching. Identical amino acid residues are enclosed within boxes. The sections indicated by horizontal solid lines contain homologous regions. The histidine residue ADP-ribosylated by DT is indicated by an asterisk.

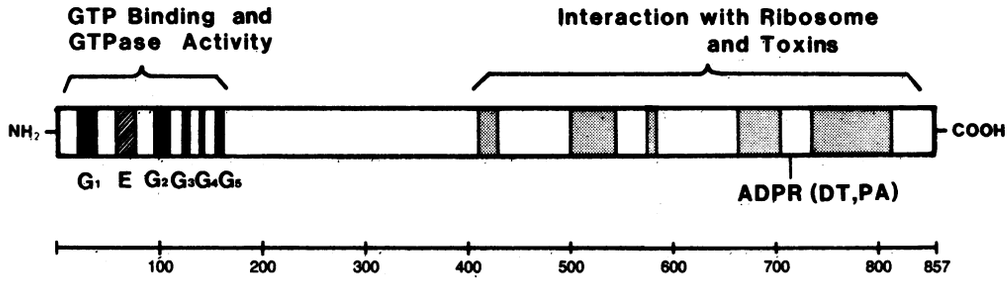

FIG. 5. Schematic representation of the structure of EF-2. The black boxes (G_1–G_5) and the hatched box E indicate the homologous regions of EF-2 with GTP-binding proteins or with only elongation factors, respectively. The shaded boxes correspond to the homologous regions between EF-2 and EF-G. ADPR indicates the site specifically ADP-ribosylated by DT and PA. Amino acid numbers are given below the structures.

the interaction of EF-2 with the ribosome and the toxins. The sites of GTP binding and of ADP-ribosylation shown in Fig. 5 agree with the locations suggested by Nilsson and Nygard (35) from experiments using limited proteolysis of EF-2.

K.K. thanks Ryoichiro Kageyama for his support and helpful advice in this work. We thank Sumiko Furuto for providing the rat liver cDNA library, Hiroto Okayama for providing the plasmids and the information on the pcD vector sequence, Hirokazu Kotani for the preparation of competent cells, and Yukiko Arakawa for excellent technical assistance. This work was supported by a Grant-in-Aid from the Ministry of Science, Education and Culture of Japan.

1. Weissbach, H. & Ochoa, S. (1976) *Annu. Rev. Biochem.* **45**, 191–216.
2. Honjo, T., Nishizuka, Y., Hayaishi, O. & Kato, I. (1968) *J. Biol. Chem.* **243**, 3553–3555.
3. Gill, D. M., Pappenheimer, A. M., Jr., Brown, R. & Kurnick, J. J. (1969) *J. Exp. Med.* **129**, 1–21.
4. Iglewski, B. H. & Kabat, D. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2284–2288.
5. Comstock, J. P. & Van, N. T. (1977) *Biochim. Biophys. Acta* **477**, 199–220.
6. Mizumoto, K., Iwasaki, K., Tanaka, M. & Kaziro, Y. (1974) *J. Biochem.* **75**, 1047–1061.
7. Merrick, W. C., Kemper, W. M. & Anderson, W. F. (1975) *J. Biol. Chem.* **250**, 2620–2625.
8. Moehring, T. J. & Moehring, J. M. (1977) *Cell* **11**, 447–454.
9. Gupta, R. S. & Siminovitch, L. (1978) *Somat. Cell Genet.* **4**, 553–571.
10. Kohno, K., Uchida, T., Mekada, E. & Okada, Y. (1985) *Somat. Cell Mol. Genet.* **11**, 421–431.
11. Kaneda, Y., Yoshida, M. C., Kohno, K., Uchida, T. & Okada, Y. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3158–3162.
12. Robinson, E. A., Henriksen, O. & Maxwell, E. S. (1974) *J. Biol. Chem.* **249**, 5088–5093.
13. Hanahan, K. & Meselson, M. (1983) *Methods Enzymol.* **100**, 333–342.
14. Okayama, H. & Berg, P. (1983) *Mol. Cell. Biol.* **3**, 280–289.
15. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
17. Takamatsu, K., Uchida, Y. & Okada, Y. (1986) *Biochem. Biophys. Res. Commun.* **134**, 1015–1021.
18. Kozak, M. (1984) *Nucleic Acids Res.* **12**, 857–872.
19. Hagenbuchle, O., Santer, M., Steitz, J. A. & Mans, R. J. (1978) *Cell* **13**, 551–563.
20. Chen, Y.-L., Gutell, R., Noller, H. F. & Wool, I. G. (1984) *J. Biol. Chem.* **259**, 224–230.
21. Brown, B. A. & Bodley, J. W. (1979) *FEBS Lett.* **103**, 253–255.
22. Van Ness, B. G., Howard, J. B. & Bodley, J. W. (1980) *J. Biol. Chem.* **255**, 10710–10716.
23. Halliday, K. R. (1984) *J. Cyclic Nucleotide Protein Phosphorylation Res.* **9**, 435–448.
24. Arai, K., Clark, B. F. C., Duffy, L., Jones, M. D., Kaziro, Y., Laursen, R. A., L'Italien, J., Miller, D. L., Nagarkatti, S., Nakamura, S., Nielsen, K. M., Petersen, T. E., Takahashi, K. & Wade, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1326–1330.
25. Zengel, J. M., Archer, R. H. & Lindahl, L. (1984) *Nucleic Acids Res.* **12**, 2181–2192.
26. Nagata, S., Nagashima, K., Tsunetsugu-Yokota, Y., Fujimura, K., Miyazaki, M. & Kaziro, Y. (1984) *EMBO J.* **3**, 1825–1830.
27. Sacerkot, C., Dessen, P., Hershey, J. W. B., Plumbridge, J. A. & Grunberg-Manago, M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7787–7791.
28. Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H. & Goeddel, D. B. (1983) *Nature (London)* **302**, 33–37.
29. Tanabe, T., Nukada, T., Nishikawa, Y., Sugimoto, K., Suzuki, H., Takahashi, H., Noda, M., Haga, Y., Ichiyama, A., Kangawa, K., Minamino, N., Matsuo, H. & Numa, S. (1985) *Nature (London)* **315**, 242–245.
30. Dayhoff, M. O., ed. (1978) *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Washington, DC).
31. Seeburg, P. H., Colby, W. W., Capon, D. J., Goeddel, D. V. & Levinson, A. D. (1984) *Nature (London)* **312**, 71–75.
32. McGraph, J. P., Capon, D. J., Goeddel, D. V. & Levinson, A. D. (1984) *Nature (London)* **310**, 644–649.
33. Girshovich, A. S., Bochkareva, E. S., Kurtskhalia, T. V., Pazdnyakov, V. A. & Ovchinnikov, Y. A. (1979) *Methods Enzymol.* **60**, 726–745.
34. la Cour, Y. F. M., Nyborg, J., Thirup, S. & Clark, B. F. C. (1985) *EMBO J.* **4**, 2385–2388.
35. Nilsson, L. & Nygard, O. (1985) *Eur. J. Biochem.* **148**, 299–304.