Proc. Natl. Acad. Sci. USA
Vol. 76, No. 5, pp. 2153–2157, May 1979
Biochemistry

Corticotropin and β-endorphin: Construction and analysis of recombinant DNA complementary to mRNA for the common precursor

(cDNA clone/mRNA purification/nucleotide sequence)

JAMES L. ROBERTS*†, PETER H. SEEBURG*, JOHN SHINE*, EDWARD HERBERT‡, JOHN D. BAXTER*†, AND HOWARD M. GOODMAN*

*Howard Hughes Medical Institute Laboratories and the Departments of Biochemistry and Biophysics and Medicine, and the †Metabolic Research Unit at the University of California, San Francisco, California 94143; and ‡Department of Chemistry, University of Oregon, Eugene, Oregon 97403

Communicated by Alexander Rich, January 29, 1979

ABSTRACT **A cDNA fragment synthesized from mouse mRNA (ACTH/LPH mRNA) that codes for the precursor polypeptide containing corticotropin (ACTH), β-lipotropin (LPH), and several other peptides has been cloned in bacteria. The mRNA was enriched for ACTH/LPH mRNA translational activity (to about 75%) prior to cDNA synthesis. It appears to contain about 1200 bases, of which approximately 450 bases are not translated. The cloned DNA fragment is complementary to the region of the mRNA coding for the protein fragment β-LPH-(44–90); this contains all of the amino acids of [Met]-enkephalin (residues 61–65 of β-LPH), most of the amino acids of β-melanocyte-stimulating hormone, and all but the carboxy-terminal amino acid of β-endorphin. Based on assignment of the amino acid sequence of mouse β-LPH from the nucleic acid sequence, it appears that there is extensive homology of mouse β-LPH with human and porcine β-LPH. The data also establish the linkage between β-melanocyte-stimulating hormone and β-endorphin as a Lys-Arg sequence. It is hoped that this cloned DNA can be used as a probe to study the expression and structure of the ACTH/LPH gene.**

The peptide hormone adrenocorticotropin (corticotropin, ACTH), with a molecular weight of 4500, is synthesized as part of a precursor protein of molecular weight about 30,000 (1–4). This precursor protein contains the overlapping sequences for β-endorphin, [Met]enkephalin, β-melanocyte-stimulating hormone (β-MSH), α-MSH, β-lipotropin (β-LPH), ACTH-like intermediate lobe peptide, and other peptides with no known function (refs. 5–7 and Fig. 3). These peptides are generated by posttranslational proteolysis and glycosylation of the precursor protein (8, 9). This precursor has been identified in both primary pituitary cultures and tumor cell lines, and the pituitary is clearly one site for production of these peptides (1–3). However, ACTH, β-endorphin, and enkephalin have been identified in other regions of the brain (10–12). For instance, in some regions there is material that reacts with antisera to β-endorphin but not to [Met]enkephalin, whereas in other regions, the converse is true (12). Thus, the question is raised as to whether these polypeptides always come from the same mRNA or whether their gene sequences are also present in other mRNAs. To probe such questions, it is important to isolate nucleic acid sequences coding for the precursor. These could then be used to detect and purify mRNA sequences from other tissues.

The sequences of ACTH (which contains α-MSH and ACTH-like intermediate lobe peptide) and β-LPH (which contains β-MSH and the opiate peptides, β-endorphin and [Met]enkephalin) have been determined for a variety of species (13). However, the sequence of the other portions of the precursor protein is not known, since it is difficult to isolate it in quantity due to its instability. The entire sequence of the precursor is of particular interest because an identification of the structure of the unknown NH_2-terminal region might be made and the linkage between the individual hormones and the possible glycosylation sites would be determined. An alternate approach to identifying the primary sequence of the ACTH/LPH protein precursor is to isolate the mRNA coding for it. DNA complementary to the mRNA could be synthesized and cloned in bacteria and its nucleotide sequence could be determined. This would then allow an assignment of the amino acid sequence of the protein. In this communication, we report the cloning and sequence analysis of a DNA fragment containing bases complementary to a portion of mouse ACTH/LPH mRNA.

MATERIALS AND METHODS

Mouse pituitary tumor cells, AtT20/D_{16v}, were grown in roller bottle cultures to confluency as described (1). The cells were harvested and membrane-associated polysomes were isolated as described (1). RNA was extracted from the polysomes with phenol/chloroform and fractionated by oligo(dT)-cellulose chromatography (1, 14). Poly(A)-enriched RNA was fractionated in a 5–25% sucrose (Schwarz/Mann, ultra pure) gradient containing 70% formamide (Eastman) as denaturant (15). Gradients were eluted; the $A_{257\ nm}$ was monitored with a Gilford spectrophotometer with the gradient elution attachment. Fractions (0.5 ml) were collected, precipitated with ethanol, and dissolved in sterile 10 mM Hepes (pH 7.4). Recovery of RNA was 70–85%. Reticulocyte mRNA (≈60% globin mRNA) and phage MS2 RNA (Miles) were used as size markers for the gradient in addition to the ribosomal RNAs.

The fractionated RNA was translated in the mRNA-dependent reticulocyte lysate protein-synthesizing system, and its ability to code for the ACTH/LPH precursor was measured by indirect immunoprecipitation with purified antisera to α-ACTH-(1–24) (1, 16). Sodium dodecyl sulfate/polyacrylamide (12.5%) slab gel electrophoresis was also used to analyze the synthesized proteins (17). Double-stranded cDNA complementary to both the total and fractionated RNAs was syn-

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: ACTH, adrenocorticotropin (corticotropin); β-LPH, β-lipotropin; β-MSH, β-melanocyte-stimulating hormone; bp, base pairs; poly(A)-RNA, polyadenylylated RNA.

thesized, purified, cleaved with *Hae* III endonuclease, analyzed by polyacrylamide gel electrophoresis, and cloned as described (18). Briefly, 10 μg of enriched RNA was converted into double-stranded cDNA (5% conversion on the first strand, 80% on the second strand) and then digested with *Hae* III. *Hin*dIII decamers were then ligated to the DNA. The ligated material was subsequently cleaved with *Hsu* I (an isoschizomer of *Hin*dIII). This material was then fractionated on an 8% polyacrylamide preparative gel, and the band corresponding to the ≈140-base-pair (bp) *Hae* III fragment (now ≈155 bp) was cut out and electroeluted. This material (≈2 ng) was ligated into the *Hin*dIII site of pBR322 and transformed into the χ1776 strain of *Escherichia coli* in a P3 physical containment facility as specified by the National Institutes of Health guidelines.

Clones containing recombinant DNA were grown on a preparative scale; the plasmids were isolated and used for restriction endonuclease analysis and determination of the nucleotide sequence of the inserted piece of DNA. The plasmid was digested with *Hsu* I; the inserted DNA was isolated and labeled at either the 5′ end with [γ-^{32}P]ATP and bacteriophage T4 polynucleotide kinase or at the 3′ end with [α-^{32}P]NTPs and reverse transcriptase (RNA-dependent DNA nucleotidyltransferase) from avian myeloblastosis virus (19). The chemical degradation method of Maxam and Gilbert (19) was then used to determine the sequence of the DNA after *Hha* I cleavage and separation of the resulting two labeled fragments by polyacrylamide gel electrophoresis.

RESULTS

Isolation and Partial Purification of ACTH/LPH mRNA. mRNA coding for the ACTH/LPH precursor protein was assayed by *in vitro* cell-free translation and immunoprecipitation as described in *Materials and Methods*. Previous translational assays have suggested that this mRNA is almost entirely (96%) localized in the membrane fraction of AtT20 cell homogenates (1). Since mRNA in membrane-associated polysomes represents approximately 65% of the total cellular mRNA (1), an enrichment for ACTH/LPH mRNA translational activity was obtained by isolating the RNA from polysomes of a membrane subcellular fraction. Polyadenylylated RNA [poly(A)-RNA] was then isolated by oligo(dT)-cellulose chromatography. All of the detectable ACTH/LPH mRNA activity was bound to the column, even after a low salt (0.1 M) wash, suggesting that essentially all of the ACTH/LPH mRNA is polyadenylylated. The

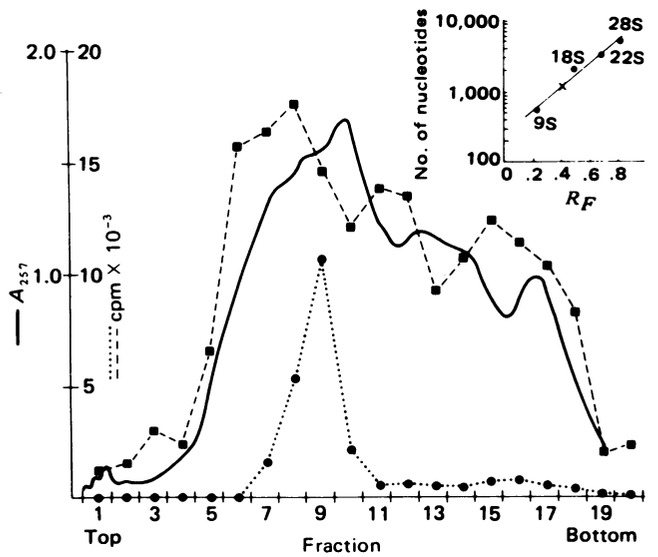

FIG. 1. Sucrose gradient purification of ACTH/LPH mRNA. Poly(A)-mRNA (400 μg) was centrifuged through a 5–25% sucrose gradient containing 70% formamide in an SW 41 rotor for 16 hr at 40,000 rpm at 20°C. Solid line, A_{257} monitored while the gradient was eluted from the top. Fractions (0.5 ml) were collected and the RNA was isolated by ethanol precipitation. The fractionated RNA was translated in an mRNA-dependent reticulocyte protein-synthesizing system and the translation products were analyzed for total protein synthesis (■---■) and for ACTH/LPH specific synthesis (●·····●). Two identical parallel sucrose gradients were run containing either the size markers globin mRNA (9S, 560–600 nucleotides) (20) and MS2 RNA (22S, 3500 nucleotides) (21) or the ribosomal RNAs.

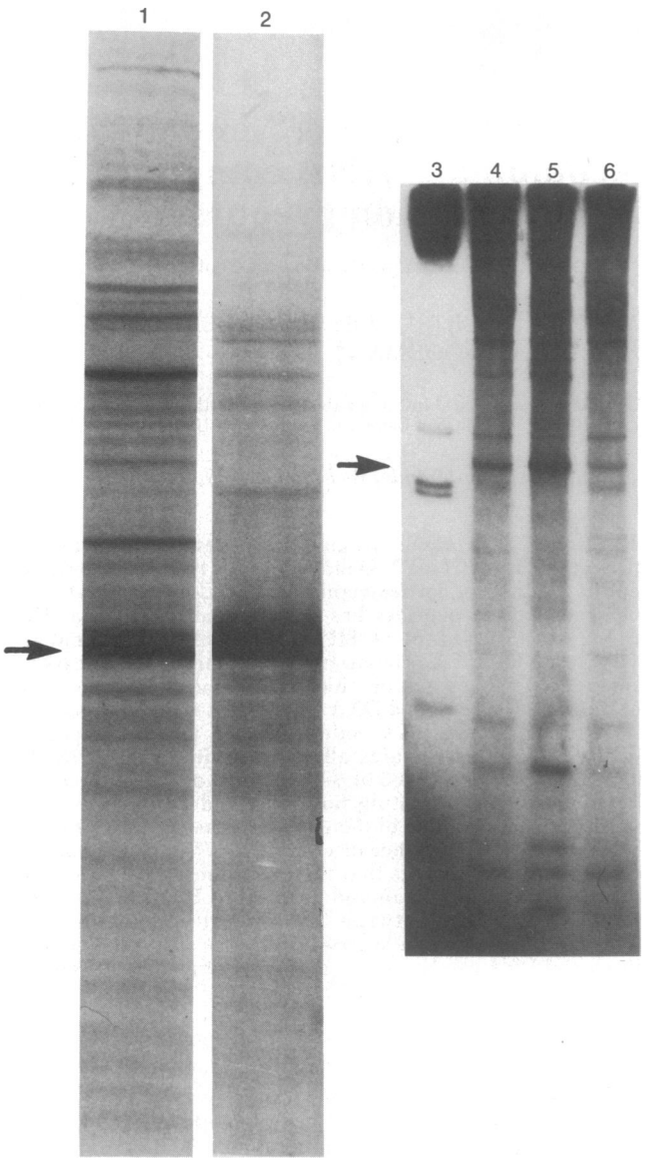

FIG. 2. Protein and cDNA template activity of RNA-enriched ACTH/LPH mRNA. Lanes 1 and 2, autoradiogram of [^{35}S]methionine-labeled proteins synthesized in the reticulocyte system and separated on a 12.5% polyacrylamide gel. Proteins translated from unfractionated RNA are shown in lane 1, whereas lane 2 shows proteins synthesized from fraction 9 of the sucrose gradient. Arrow next to lane 1 denotes the position of the ACTH/LPH precursor protein. Lanes 3–6, autoradiogram showing a separation of [^{32}P]DNAs in an 8% polyacrylamide gel. *Hpa* II-cut fd DNA (lane 3) was used as a double-stranded DNA marker. Double-stranded cDNA was prepared from unfractionated (lane 4), enriched (lane 5, gradient fraction 9), or nonenriched (lane 6, gradient fraction 14) poly(A)-RNA and cut with *Hae* III. Arrow next to lane 3 shows the position of the ≈140-nucleotide fragment discussed in the text.

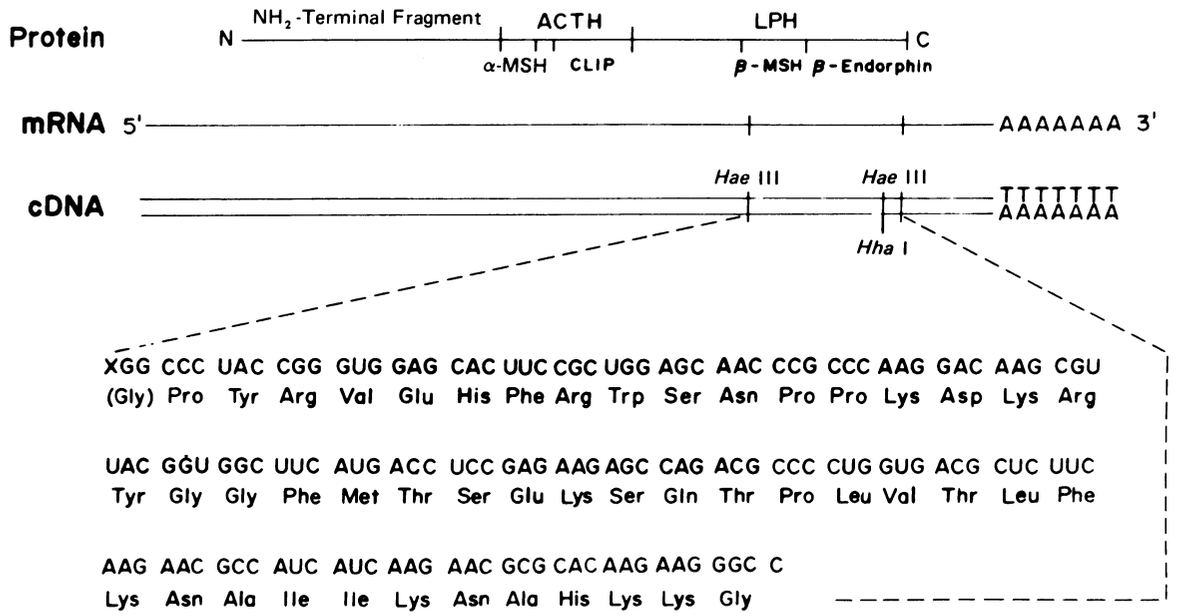

FIG. 3. Sequence of the cloned fragment. The nucleotide sequence of the cloned fragment, as determined by the chemical degradation method of Maxam and Gilbert (19), is shown as an RNA sequence with the corresponding amino acid sequence for one of the reading frames. The protein structure of the ACTH/LPH precursor and its constituent hormones is also shown. CLIP, ACTH-like intermediate lobe peptide.

resulting preparation of poly(A)-RNA was approximately 2-fold enriched for ACTH/LPH cell-free translational activity (data not shown; see Fig. 2).

Further purification of the ACTH/LPH mRNA was obtained by velocity sedimentation of the poly(A)-RNA through a sucrose gradient containing formamide (Fig. 1). Based on the standards used (*inset*, Fig. 1), ACTH/LPH mRNA had a length of approximately 1200 nucleotides. In the peak fraction, ACTH/LPH mRNA was enriched to where it accounted for 75% of the total cell-free translation activity. However, this estimation of purity may be an overestimate of the true proportionate weight of ACTH/LPH mRNA (see below).

Characterization of cDNA to AtT20 Cell mRNA. Double-stranded cDNA was synthesized with avian myeloblastosis virus reverse transcriptase from both total poly(A)-RNA and the RNAs purified by sucrose gradient centrifugation as templates (18). The cDNA was digested with *Hae* III endonuclease and the resulting fragments were analyzed by polyacrylamide gel electrophoresis (Fig. 2, lanes 4–6). The restriction fragment pattern was complex, even when ACTH/LPH mRNA purified by sucrose gradient centrifugation was used as template (lane 5). This is in contrast to the relatively simple pattern of proteins generated by translation of the same RNA (lane 2). However, certain restriction fragments were enriched in those cDNAs synthesized from gradient fractions containing the peak of ACTH/LPH mRNA translation activity (Fig. 2, lane 5) and, hence, might be derived from ACTH/LPH cDNA. One such fragment was a *Hae* III fragment of approximately 140 bp which contains an asymmetric *Hha* I site. To determine whether this fragment is ACTH/LPH related and to obtain a DNA probe for further use in studying ACTH/LPH gene sequences, we cloned this DNA fragment in bacteria and determined its nucleotide sequence.

Cloning and Sequence Determination of a DNA Restriction Fragment from ACTH/LPH cDNA. The double-stranded cDNA (Fig. 2) was digested with *Hae* III. The whole digest was ligated to a double-stranded decanucleotide linker containing the *Hin*dIII (or its isoschizomer *Hsu* I) endonuclease recognition sequence (C-C-A-A-G-C-T-T-G-G), and the ligated products were digested with *Hin*dIII (18). The resulting material was fractionated on a preparative polyacrylamide gel, and the ≈140 bp *Hae* III fragment (now ≈155 bp due to linker addition) was excised from the gel and the DNA was electrophoretically eluted (18). The fragment with cohesive termini was then ligated to *Hin*dIII-cut and alkaline phosphatase-treated pBR322 and transformed into the χ1776 strain of *E. coli* (18). Subsequent analysis of Amp^R Tet^S colonies (only 8 of the 80 colonies were screened) showed that six out of eight plasmids in the recombinant clones contained an inserted ≈155-bp piece of DNA in the *Hin*dIII site. Digestion of the DNA containing the ≈155-bp insertion with *Hha* I resulted in two fragments, suggesting that this DNA might be the same as the predominant *Hae* III fragment seen after purification of the ACTH/LPH mRNA on the sucrose gradient. The nucleotide sequence of this fragment (Fig. 3) was determined by the chemical degradation method of Maxam and Gilbert (19). Fig. 3 shows the fragment and the amino acid sequence corresponding to one of the possible reading frames. To assay whether the DNA sequence obtained for the inserted DNA had been altered during the transformation and the cloning procedure, we pooled DNA from the six positive clones and subsequently determined its sequence. A sequence identical to that shown in Fig. 3 was obtained.

The amino acid sequence of mouse β-LPH has not to our knowledge been determined. However, the sequences of β-endorphin and β-MSH are highly conserved between mammalian species (13). Sabol (22) has found that the amino acid composition of β-endorphin isolated from mouse AtT20 cells is identical to that of porcine endorphin. Other studies of the tryptic peptides derived from β-LPH also support the observation that murine and porcine β-endorphin are similar (5, 6). Thus, it may be anticipated that the amino acid sequence of mouse, human, and ovine β-LPH will show extensive homology. Fig. 3 shows the amino acid sequence corresponding to one of the possible reading frames of the cloned fragment. This sequence from the cloned DNA corresponds to amino acids 44–90 of human β-LPH and does show extensive homology with these hormones. These observations indicate that the cloned DNA consists of sequences complementary to ACTH/LPH mRNA.

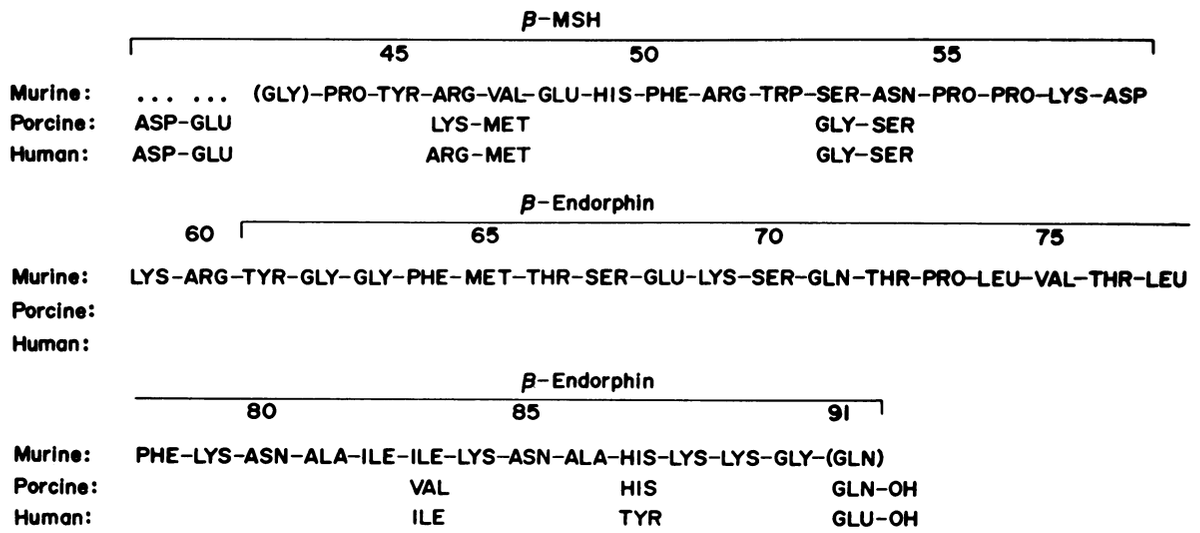

FIG. 4. Amino acid sequence comparison. The amino acid sequence of mouse β-MSH and β-endorphin determined from the cloned DNA sequence are compared to the same sequences from the human and porcine peptides (taken from ref. 24).

DISCUSSION

In the current studies, we have cloned a DNA fragment containing sequences of the ACTH/LPH gene by using RNA enriched for ACTH/LPH mRNA translational activity. Based on studies of the migration of this mRNA in partially denaturing (formamide) sucrose density gradients, it appears that it is about 1200 nucleotides long. This value agrees with that reported earlier by Jones and Grunberger (23), who worked with RNA from AtT20 solid tumors. The ACTH/LPH precursor contains ≈250 amino acid residues, requiring ≈750 nucleotides in the RNA to code for it. Thus, there are approximately 450 bases, including the poly(A) segment in the ACTH/LPH mRNA not involved in coding. These findings are also consistent with the view that the protein of molecular weight ≈28,500 detected in cell-free translation systems, and not a significantly larger precursor, is the actual primary translation product of ACTH/LPH mRNA (1, 3, 4).

About 75% of the amino acid incorporation is into the ACTH/LPH precursor when the sucrose gradient-enriched RNA preparation is translated in a cell-free system. When cDNA was prepared from the same RNA, cleaved with endonuclease *Hae* III, and analyzed on polyacrylamide gels, prominent bands enriched relative to the total digest appeared, indicating that some mRNA species in the preparation (likely ACTH/LPH mRNA) was predominant. One of these *Hae* III fragments was isolated and cloned. Nucleotide sequence analysis of this cloned fragment revealed that it contained the codons for a polypeptide sequence very similar, but not identical, to part of human and porcine LPH. We therefore concluded that the cloned DNA did indeed contain mouse ACTH/LPH gene sequences. The sequence of the fragment is complementary to a region of the mRNA coding for the protein fragment β-LPH-(44–90). This fragment is near the carboxyl terminus of the precursor and consequently corresponds to a portion near the 3′ end of the mRNA. The fragment codes for all of [Met]enkephalin (Fig. 4, residues 61–65 inclusive), most of the β-MSH sequence (it is missing residues 41–43, Fig. 4), and all but the carboxyl-terminal residue (Fig. 4, residue 91) of β-endorphin.

It is of interest to compare the structure of murine β-LPH-(44–90) (determined in this study from the nucleotide sequence of the cloned DNA) with that of other species determined from amino acid sequence analysis (Fig. 4). Murine β-MSH differs in three amino acids from human β-MSH and four amino acids from bovine β-MSH (data are not available for amino acids 1 and 2 of the murine hormone). This is consistent with the findings (unpublished) that antisera to human or bovine β-MSH failed to crossreact with the mouse hormone. Of special note is the substitution of a valine in murine β-MSH for methionine. Methionine has been present in β-MSH in all other species thus far examined. This change occurs in the region of β-MSH (amino acids 47–53 of β-LPH) that is normally homologous with a region in α-MSH (13). Although the assignment of valine could, in principle, be due to an error (e.g., in reverse transcription or sequence analysis), the change is substantiated by analysis of the tryptic peptides of the murine precursor, which failed to reveal the presence of a β-MSH methionine (5, 6).

The DNA sequence shows that a Lys-Arg linkage connects the amino acids of murine β-MSH and β-endorphin. This linkage has also been found in other species (13) and perhaps suggests that a similar type of enzymatic cleavage might be involved in the processing of the precursor in the various species. Murine β-endorphin differs from porcine β-endorphin only in amino acid 83 (Ile for Val) and from human β-endorphin only in amino acid 87 (His for Tyr). [The carboxyl-terminal amino acid of mouse β-endorphin (residue 91 of β-LPH) is not known.] Tryptic peptide analysis of the murine precursor also supports the two amino acid substitutions in β-endorphin noted above (unpublished observations).

Attempts were made to prepare full-length cDNA to ACTH/LPH mRNA. Many clones were obtained by using double-stranded DNA 800–1200 bases long. Unfortunately, none of these contained the sequences for the ACTH/LPH precursor. This was a surprising result, considering the purity of the mRNA used as template. A possible explanation is that there exists an internal *Hsu* I (*Hin*dIII) site in the ACTH/LPH cDNA. This would have resulted in internal cleavage of the cDNA when it was treated with *Hsu* I after addition of the *Hin*dIII linkers used for cloning.

It is hoped that the cloned ≈155-bp fragment will be useful as a probe to detect and isolate full-length cDNA clones as well as clones from the natural DNA. In addition, it should be useful in studying whether the ACTH/LPH gene is expressed by tissues other than the pituitary (detected by identification of ACTH/LPH mRNA) and the extent to which β-endorphin and enkephalins arise from a common precursor.

Note Added in Proof. A clone of bovine ACTH/LPH has recently been described by Nakanishi *et al.* (25).

This work was supported by National Institutes of Health Grants AM19997 (to J.D.B. and H.M.G.), CA 14026 (to H.M.G.), and AM16879 (to E.H.). J.L.R. is a National Institutes of Health Postdoctoral Fellow; Grant 5F32 AM 05774-02. H.M.G. and J.D.B. are Investigators of the Howard Hughes Medical Institute.

1. Roberts, J. L. & Herbert, E. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 4826–4830.
2. Mains, R. E. & Eipper, B. A. (1976) *J. Biol. Chem.* **251,** 4115–4120.
3. Nakanishi, S., Taii, S., Hirata, Y., Matsukura, S., Imura, H. & Numa, S. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 4319–4323.
4. Jones, R. E., Pulkrabek, P. & Grunberger, D. (1977) *Biochem. Biophys. Res. Commun.* **74,** 1490–1495.
5. Roberts, J. L. & Herbert, E. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5300–5304.
6. Mains, R. E., Eipper, B. A. & Ling, N. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 3014–3018.
7. Scott, A. P., Ratcliffe, J. G., Rees, L. H., Landon, J., Bennett, H. P. J., Lowry, P. J. & McMartin, C. (1973) *Nature (London) New Biol.* **244,** 65–67.
8. Roberts, J. L., Phillips, M. A., Rosa, P. A. & Herbert, E. (1978) *Biochemistry* **17,** 3609–3618.
9. Eipper, B. A. & Mains, R. E. (1978) *J. Biol. Chem.* **253,** 5732–5744.
10. Krieger, D. T., Liotta, A. & Brownstein, M. J. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 648–652.
11. Krieger, D. T., Liotta, A., Suda, T., Palkovits, M. & Brownstein, J. J. (1977) *Biochem. Biophys. Res. Commun.* **76,** 930–936.
12. Rossier, J., Vargo, T. M., Minick, S., Ling, N., Bloom, F. E. & Guillemin, R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5162–5165.
13. Dayhoff, M. O. (1972) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, DC), Vol. 5, D194–D197.
14. Palmiter, R. D. (1973) *J. Biol. Chem.* **248,** 2095–2106.
15. Suzuki, Y., Gage, L. P. & Biown, D. D. (1972) *J. Mol. Biol.* **70,** 637–649.
16. Pelham, H. R. B. & Jackson, R. J. (1976) *Eur. J. Biochem.* **67,** 247–256.
17. O'Farrell, P. H. (1975) *J. Biol. Chem.* **250,** 4007–4021.
18. Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. (1977) *Nature (London)* **270,** 486–494.
19. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
20. Efstratiadis, A., Kafatos, F. C. & Maniatis, T. (1977) *Cell* **10,** 571–585.
21. Fiers, W., Contreras, R., Duerinck, F., Haegeman, D., Iserentant, D., Merregaert, J., Minjou, W., Molemans, F., Raeymaekers, A., Van den Berghe, A., Volckaert, G. & Ysebaert, M. (1976) *Nature (London)* **260,** 500–507.
22. Sabol, S. L. (1978) *Biochem. Biophys. Res. Commun.* **82,** 560–569.
23. Jones, R. E. & Grunberger, D. (1978) *Arch. Biochem. Biophys.* **188,** 476–483.
24. Li, C. H. & Chung, D. (1976) *Nature (London)* **260,** 622–624.
25. Nakanishi, S., Inoue, A., Kita, T., Numa, S., Chang, A. C. Y., Cohen, S. N., Nunberg, J. & Schimke, R. T. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 6021–6025.