Proc. Natl. Acad. Sci. USA
Vol. 77, No. 8, pp. 4890–4894, August 1980
Genetics

Structural organization of human genomic DNA encoding the pro-opiomelanocortin peptide

(DNA cloning/corticotropin/β-endorphin/plasmid/intervening sequence)

ANNIE C. Y. CHANG, MADELEINE COCHET, AND STANLEY N. COHEN

Departments of Genetics and Medicine, Stanford University School of Medicine, Stanford, California 94305

Contributed by Stanley N. Cohen, May 27, 1980

ABSTRACT We have isolated a human genomic DNA segment encoding the corticotropin-β-lipotropin precursor peptide from a fetal DNA library, using previously cloned bovine cDNA for this peptide as a probe. The human genomic DNA was studied by electron microscope heteroduplex analysis and gel blotting methods, and its nucleotide sequence was determined and compared with that of cDNA corresponding to bovine pro-opiomelanocortin mRNA. From this sequence, segments of interspecies conservation and divergence, punctuated by pairs of the basic amino acid residues lysine and arginine, were identified. No noncoding intervening sequence was observed over an 830-base-pair DNA segment extending from a position near the 5′ end of the structural pro-opiomelanocortin gene through the 3′ terminus of the cDNA and including sequences for the component peptide hormones corticotropin and β-lipotropin.

The pituitary hormones corticotropin (ACTH) and β-lipotropin (β-LPH) are derived from a large precursor peptide of 29,500 M_r (1–5). These two hormones are each known to include smaller peptides having distinct biological activities: α-melanotropin (α-MSH) and corticotropin-like intermediate lobe peptide (CLIP) are formed from ACTH; γ-LPH and β-endorphin are peptide components of β-LPH. β-MSH is contained within γ-LPH (4–9). γ-MSH and possibly other discrete peptides are located in the previously "cryptic" pre-ACTH part of the precursor peptide (5), which has been called pro-opiomelanocortin (POMC) (10). Recently, the cloning in *Escherichia coli* of cDNA corresponding to full-length bovine POMC mRNA has been described and the nucleotide sequence of a 1091-base-pair (bp) cDNA insert has been reported (5); from this sequence, the amino acid sequence of the pre-ACTH part of the peptide was predicted, and the structural organization of the biologically active component peptides of the POMC molecule was shown.

The various structural and functional domains of POMC provide an opportunity to study the relationship of noncoding intervening sequences to amino-acid-coding sequences of a multicomponent peptide containing a series of repetitive DNA units. In order to elucidate the structure of the chromosomal DNA segment encoding POMC, and to investigate interspecies variation, we have used the bovine cDNA clone as a probe to isolate and characterize a POMC-encoding human genomic DNA segment.

MATERIALS AND METHODS

A human fetal DNA library of λ Charon 4A bacteriophage containing genomic DNA fragments was a gift from T. Maniatis (11); the phage were grown and DNA was purified as described (12). The pSNAC20 plasmid, containing a cDNA insert encoding bovine POMC was propagated in *E. coli* strain SR1592 $r_k^- m_k^+$ (13) or PM191 *recA* (P. Meacock and S. N. Cohen, unpublished data). DNA isolation and transformation have been described (14). DNA fragments generated as described in the text were fractionated by electrophoresis on 3% or 4.5% acrylamide gel in TBE buffer (15) and were eluted from the gel electrophoretically (16). Samples (50–200 ng) of the DNA were "nick-translated" with DNase I (1 ng/ml) and *E. coli* DNA polymerase (Boehringer Mannheim) in 50 mM Tris·HCl (pH 7.8)/5 mM $MgCl_2$/10 mM mercaptoethanol/bovine serum albumin (50 μg/ml) buffer containing 5 μM [α-^{32}P]dCTP, 5 μM [α-^{32}P]dTTP (both >300 Ci/mmol; 1 Ci = 3.7×10^{10} becquerels) 12.5 μM dGTP, and 12.5 μM dATP and incubated for 1 hr at 15°C (17), yielding DNA with a specific activity of 8×10^7 cpm/μg. Phage plaque hybridization with radioactively labeled probe was carried out essentially as described by Benton and Davis (18).

5′-End-labeled DNA fragments were prepared, and their sequences were determined by using the procedure of Maxam and Gilbert (19); electrophoresis was on 0.35-mm-thick 8% and 20% polyacrylamide gels (20). Formation and analysis of heteroduplexes were carried out as described (21).

Containment conditions were as prescribed in *NIH Guidelines for Recombinant DNA Research.*

RESULTS

Isolation of a Human Chromosomal DNA Sequence Encoding POMC. The pSNAC20 plasmid, which contains the 1091-bp cDNA insert complementary to bovine POMC mRNA, was used (22) to identify corresponding sequences in clones from a human fetal DNA library. The library used (11) was constructed by partial digestion of fetal liver DNA with *Hae* III and *Alu* I restriction endonucleases, yielding fragments that were cloned in *E. coli*, using the λ Charon 4A bacteriophage vector (23).

In total, 1.5×10^5 λ plaques were screened (18) by aqueous hybridization (24) at 60°C with ^{32}P-labeled DNA produced by nick-translation of the bovine POMC cDNA insert removed from the pSNAC20 plasmid by digestion with *Pst* I endonuclease and separated from the vector DNA by gel electrophoresis. This temperature, rather than 65°C, was chosen to allow hybridization to occur in the presence of possible base pair mismatch between the bovine cDNA sequence and the human genomic DNA sequence. Phage derived from each of 16 positively reacting plaques were tested with four separate probes synthesized by nick-translation of bovine POMC cDNA fragments that were generated by using combinations of *Pst* I, *Sst* II, and *Eco*RI enzymes, as shown in Fig. 1*a*. Only one plaque, designated λ8A, reacted strongly with all four probes; the remaining 15 clones reacted strongly with fragment D of the pSNAC20 DNA insert, which contains the 3′ noncoding region

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, corticotropin; β-LPH, β-lipotropin; α-MSH, α-melanotropin; POMC, pro-opiomelanocortin; bp, base pair(s); kb, kilobase(s).

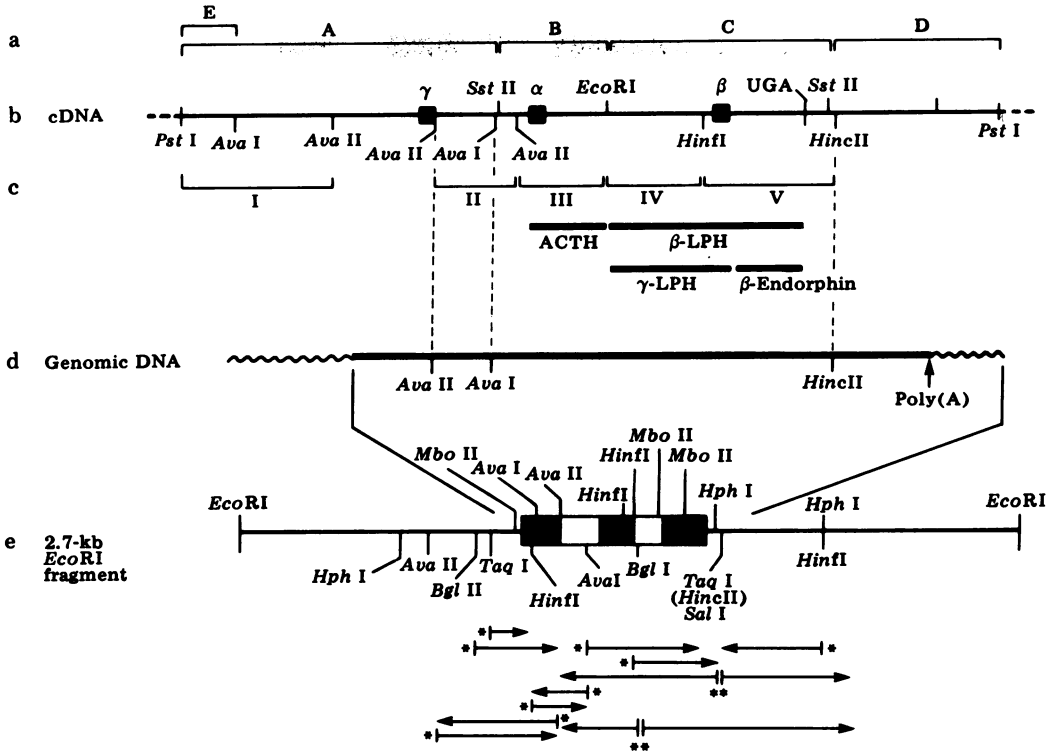

FIG. 1. Analysis of the structural organization of a 2.7-kilobase (kb) human genomic DNA fragment encoding POMC. (*b*) Certain restriction endonuclease cleavage sites on cDNA corresponding to bovine POMC mRNA. The boxes show the positions of α-, β-, and γ-MSH sequences. Broken lines flanking the *Pst* I site represent plasmid vector sequences. (*a* and *c*) Extent of bovine cDNA fragments that were nick-translated into probes for localization of human genomic POMC sequences on λ8A DNA and on the 2.7-kb fragment. Thickened lines show positions of the sequences encoding ACTH, β-LPH, γ-LPH, and β-endorphin on the bovine cDNA. (*d*) Thickened line represents human genomic sequences that correspond to the peptide-coding sequence of bovine cDNA; vertical lines represent endonuclease cleavage sites common to both species. The arrow indicates the site at which poly(A) is added in bovine mRNA. (*e*) Endonuclease cleavage map of the 2.7-kb *Eco*RI fragment containing the human POMC sequences. The block shows the extent of the POMC structural gene. Darkened segments indicate the conserved regions, and white boxes indicate the variable regions of the peptide-coding segment. The arrows below *e* show the DNA fragments whose sequences were determined and the direction of sequencing; asterisks indicate the 5′-^{32}P-labeled ends.

of the bovine POMC cDNA, and also reacted to a much lesser extent with fragment A. No clone other than λ8A reacted with probe fragments B or C.

DNA derived from the λ8A clone was characterized further by the DNA blotting technique described by Southern (Fig. 2*a*) (26). The DNA was cleaved with *Eco*RI, *Bam*HI, *Pst* I, or *Hin*dIII endonuclease and analyzed with each of the pSNAC20-derived nick-translated cDNA probes described above; a 2.7-kb *Eco*RI fragment that contains a DNA sequence common to all four probes was identified (Fig. 2*a*), suggesting that this *Eco*RI fragment contains sequences that span most or all of the pSNAC20 insert, including the 3′ untranslated region. However, when a *Pst* I/*Ava* I-generated fragment of pSNAC20 (fragment E, Fig. 1*a*) that corresponds to the 5′ noncoding region of the bovine cDNA insert was used as a probe, no hybridization was observed.

The 2.7-kb *Eco*RI-generated DNA fragment was isolated from λ8A phage, ligated at the *Eco*RI site of the pBR322 plasmid, and introduced by transformation into strain SR1592. The resulting ampicillin-resistant transformants contained chimeric plasmids; one of these (termed pACYC401), which included a DNA fragment the same length as the 2.7-kb λ8A insert, was selected for study by electron microscopy and gel analysis.

Localization of the Human POMC Gene in the Cloned 2.7-kb *Eco*RI Fragment. To determine the extent of homology between the bovine POMC cDNA and sequences contained in the 2.7-kb *Eco*RI fragment cDNA, we formed heteroduplexes between linearized pBR322 and pSNAC20 DNA, and then hybridized the isolated 2.7-kb fragment (24). Visualization of the resulting molecules by electron microscopy shows a single duplex region that extends for 850 ± 25 base pairs without interruption, and that is bracketed by single-strand DNA arms of 862 ± 45 and of 935 ± 37 nucleotides (Fig. 3, a and b, respectively). These findings strongly suggest that the majority of the bovine POMC cDNA sequence is contained in the 2.7-kb *Eco*RI human genomic DNA fragment as an uninterrupted segment lacking an intervening sequence.

Restriction Endonuclease Analysis of the Human Genomic DNA Fragment Encoding POMC. A restriction endonuclease cleavage map was constructed for the pACYC401 plasmid by cleaving the DNA with combinations of the restriction endonucleases shown in Fig. 1*e*; the component sequences of POMC were assigned to individual DNA fragments by filter paper blotting techniques, again using fragments of the nick-translated bovine POMC cDNA as probes. Because the DNA sequences encoding the α-, β-, and γ-MSHs on the precursor peptide are similar (5), the presence of one of these sequences on a probe fragment was thought likely to produce cross-hybridization with DNA fragments that include other MSH sequences. Therefore, to enable accurate mapping, the *Pst* I-generated cDNA probe was further cleaved as shown in Fig. 1*c*; only fragments III and V contained any overlapping MSH sequences.

Hybridizations of the human genomic DNA fragment with the separate bovine POMC cDNA fragments (Fig. 2 *b* and *c*) were carried out under stringent annealing conditions (50% formamide, 52°C). Fragments I and IV of the bovine cDNA did not hybridize to the 2.7-kb *Eco*RI genomic DNA fragment; fragment II, which showed faint reactivity, had the same pat-

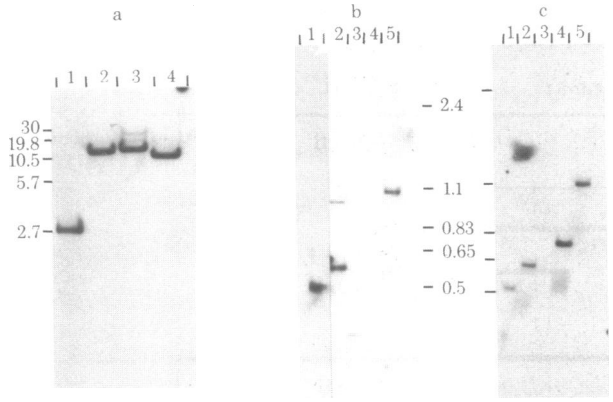

FIG. 2. (*a*) Southern filter hybridization of λ8A hybrid phage with probes nick-translated from fragments of bovine cDNA insert as shown in Fig. 1*a*. Endonuclease-generated DNA fragments were fractionated on 0.5% agarose gels in Tris/acetate/EDTA buffer. DNA fragments were depurinated in gels in 0.25 M HCl, denatured as described (25), and transferred to nitrocellulose filters (26). Filters were pretreated and hybridizations were carried out for 24 hr either at 60°C in aqueous solution or at 52°C in 50% (vol/vol) formamide, and filters were washed as described (27). λ8A DNA was treated with endonucleases *Eco*RI (lane 1), *Bam*HI (lane 2), *Hin*dIII (lane 3), and *Pst* I (lane 4). The horizontal bars represent the λ8A fragments obtained by *Eco*RI digestion; lengths (in kb) have been calibrated with a *Hin*dIII digest of a bacteriophage T5 DNA standard. (*b* and *c*) Southern filter hybridization of endonuclease-treated pACYC401 plasmid DNA, using probes made from fragments III (gel *b*) and V (gel *c*) (Fig. 1*c*). Conditions were as described for *a* except that 1.8% agarose was used. Endonuclease treatment was: lane 1, *Ava* I, *Eco*RI, and *Hin*cII; lane 2, *Ava* II, *Eco*RI, and *Hin*cII; lane 3, *Mbo* II; lane 4, *Hin*fI; lane 5, *Hph* I. The top band in lane 2 of gel *b* results from incomplete digestion of the DNA sample.

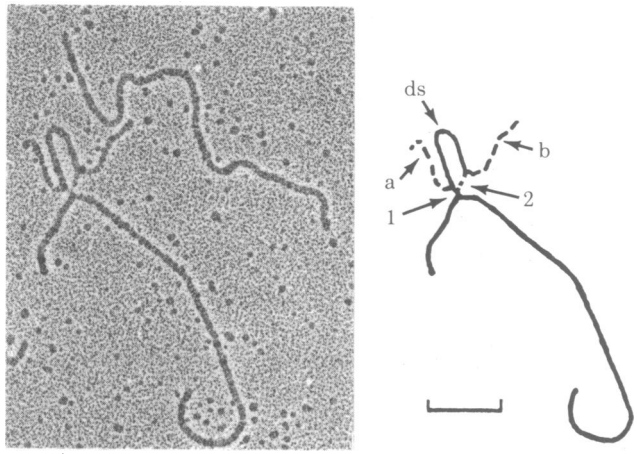

FIG. 3. Heteroduplex analysis. Plasmid DNAs of pBR322 and pSNAC20 (3 μg/ml each) were linearized by *Hin*dIII restriction enzyme, denatured by 0.1 M NaOH, and allowed to reassociate for 4 hr at room temperature as described (21). The purified 2.7-kb *Eco*RI genomic DNA fragment (previously denatured by the same technique) was added at a final concentration of 1.5 μg/ml, and the reassociation was allowed to proceed in the dark overnight at room temperature. The sample was diluted 1:20 and spread as described (21); the film was picked onto Parlodion-coated grids, stained with uranyl acetate, and rotary-shadowed with platinum. The samples were visualized and photographed, using a Philips 201 microscope. pBR322 and single-stranded bacteriophage φX174 DNAs were used as internal length standards. The broken lines and the solid line represent single-stranded and double-stranded DNA, respectively. The 2.7-kb fragment hybridizes to the bovine cDNA for a distance of 850 ± 25 base pairs (ds), leaving two single-stranded DNA segments of 862 ± 45 (a) and 935 ± 37 (b) nucleotides. The nonhybridizing sequences of the bovine cDNA are (1) about 100 nucleotides and (2) 320 ± 35 nucleotides. The bar represents 0.5 kb of double-stranded DNA.

tern as fragment III (Fig. 2*b*), while fragments III and V hybridized strongly (Fig. 2 *b* and *c*, respectively). Hybridization of the radioactively labeled probe fragments III and V was essentially the same with pACYC401 fragments generated by triple digestion with *Ava* I/*Eco*RI/*Hin*cII or *Ava* II/*Eco*RI/*Hin*cII endonucleases, or by single digestion with the *Hph* I endonuclease (Fig. 2 *b* and *c*, lanes 1, 2, and 5). However, the *Mbo* II fragment, which reacted with probe fragment III, showed no hybridization with fragment V of bovine POMC cDNA (Fig. 2 *b* and *c*, lanes 3); the converse was true for the *Hin*fI genomic DNA fragment, which hybridized with fragment V but not with fragment III of the cDNA (Fig. 2 *b* and *c*, lanes 4). Under our experimental conditions, fragments shorter than 250 bp are not detected.

We conclude from the above results that most of the human genomic DNA sequence corresponding to the bovine POMC DNA is located between the *Ava* II and *Hin*cII restriction sites in the 2.7-kb *Eco*RI fragment (Fig. 1*d*). The lack of detectable hybridization of the bovine fragment IV with the human genomic DNA is consistent with known divergence of the human and bovine β-LPH sequences within a 40 amino acid stretch of β-LPH composing the substituent peptide γ-LPH (28). Although fragments III and V of the cDNA probe contain MSH sequences, they have different and distinct hybridization patterns with the genomic DNA (i.e., they correspond to different *Mbo* II and *Hin*fI fragments), indicating that the α- and β-MSH sequences on these fragments do not show cross-hybridization under the stringent annealing conditions used.

Structure and Nucleotide Sequence of Human Gene Encoding POMC. The strategy used for DNA sequence-determination is shown in Fig. 1*e*. The nucleotide sequence determined for the human DNA, the differences found for the corresponding segment of bovine POMC cDNA, and the predicted amino acid sequences are shown in Fig. 4.

Comparison of the genomic DNA and cDNA sequences indicates continuity from a point at the 3′ noncoding end of the bovine mRNA to a position corresponding to the Cys residue 92 amino acids before the beginning of ACTH; there the two sequences diverge. The dinucleotide A-G, which is characteristic of the 3′ terminus of an intervening sequence (29), occurs at this point, and a pyrimidine-rich region (29, 30) is seen just upstream from the putative intron–exon junction.

Despite the potential for evolutionary drift in codon usage without alteration of the amino acid sequence, the nucleotide sequences of three regions are seen to be well conserved within the human and bovine POMC genes: the segment encoding ACTH, the region extending from β-MSH to the end of β-LPH and including the sequences for [Met]enkephalin and β-endorphin, and the segment extending from the end of γ-MSH towards the amino terminus of the POMC protein. In these regions, all amino acids are identical except at seven locations where one-base substitutions result in amino-acid differences (amino-acid positions −90, −78, +33, +80, +84, +125, and +129 of the POMC DNA sequence). Two regions are different in the two species: the segments encoding the amino acids immediately preceding the Lys-Arg pair that separates ACTH from the preceding segment of POMC and the segment immediately following ACTH (i.e., the first 120 bp of β-LPH). Interspecies differences in the latter segment were previously known from the corresponding amino acid sequences, which have been determined for bovine and human β-LPH (5, 28).

A 24-bp nucleotide segment beginning 63 bp (i.e., 21 amino acids) before ACTH and encoding the amino acids Pro-Glu-Gly-Gly-Pro-Glu-Pro-Arg in the human gene, and the triplet encoding the Leu at position 72 (in β-LPH), are absent in the bovine sequence. However, a stretch of 18 bp encoding the amino acids Ala-Glu-Ala-Glu-Ala-Glu of β-LPH (position 74)

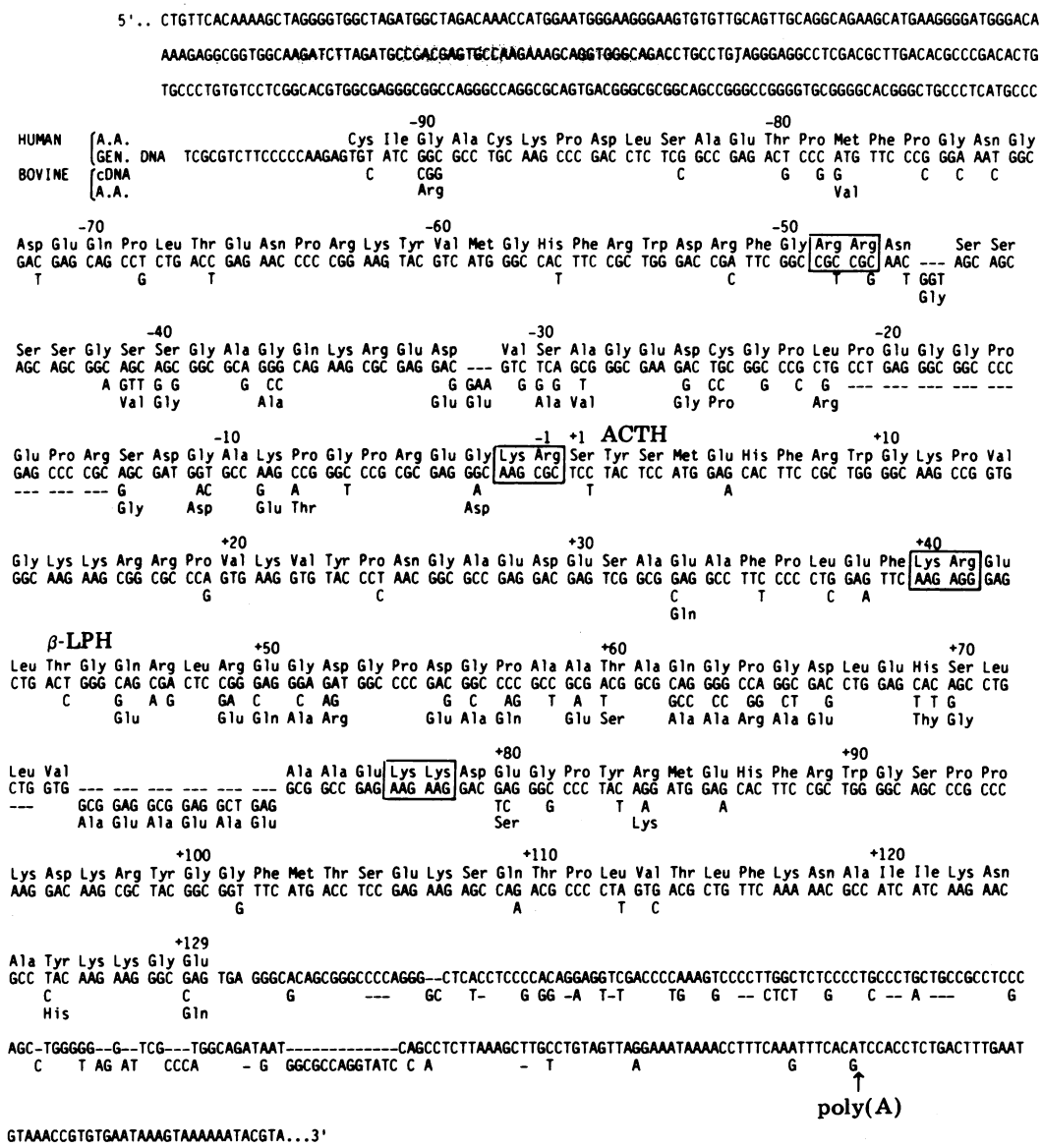

FIG. 4. Nucleotide sequence of the human POMC genome DNA strand corresponding to POMC mRNA. Above the nucleotide sequence is the predicted amino acid sequence; nucleotide and amino acid differences in the human vs. bovine sequences are indicated. Broken lines represent nucleotides absent in the cDNA or genomic sequences. The pairs of basic amono acids (Lys and Arg) that separate conserved and nonconserved regions of the POMC coding sequence are indicated by boxes. The site of poly(A) addition is indicated in the 3′ noncoding region of the sequence. Amino acids are not assigned for the region that preceeds the Cys at amino acid position −92. Amino acids are numbered with plus numbers in the direction of the carboxy terminus of POMC from the first amino acid of ACTH, and with minus numbers in the direction of the amino terminus of the peptide for the pre-ACTH segment, as in ref. 5.

and two triplets at positions −46 to −47 and −31 to −32 are present in the bovine cDNA, but are lacking in the human genomic DNA sequence; thus, the overall lengths for the bovine and human peptides predicted from the nucleotide sequence differ by only one amino acid. More than 50% of the changes in amino acid sequence of the nonconserved region are caused by single nucleotide substitutions in codons.

The 3′ noncoding region of the bovine cDNA corresponds closely to the human genomic DNA segment that follows the translational stop codon terminating the POMC protein. The distance to the poly(A) addition is approximately the same for the two species (168 bp for the bovine, 164 bp for the human) and the nucleotide sequence throughout the region is highly conserved despite its noncoding nature. Twenty-one of the 30 bp immediately following the translational stop codon match in two species, and there is greater than 95% homology in a 50-bp segment immediately preceding the poly(A) addition site.

DISCUSSION

The isolation and characterization of the human genomic DNA sequence encoding the POMC peptides has enabled identification of interspecies evolutionary divergence and conservation within the pre-ACTH segment of the precursor molecule, has provided the genetic basis (i.e., single bp substitutions) for known differences in amino acid sequence within human and bovine β-LPH, and has established the absence of intervening noncoding sequences in the DNA region encoding the bulk of the POMC peptide. The lack of intervening sequences at the sites of the basic amino acid pairs that punctuate and divide the precursor protein into a series of structural and functional domains is of special interest; such sequences have been identified at junctions of the structural domains of the immunoglobulin genes (31).

The nucleotide sequences of human genomic DNA and bovine cDNA diverge at a position 276 base pairs before the beginning of ACTH; the point of divergence has been identified

by electron microscope heteroduplex analysis, Southern blotting gels, and DNA sequence analysis. Immediately 5′ to this locus are characteristic structural features that make it highly likely that the point of divergence between the human genomic and bovine cDNA sequences represents an intron–exon junction rather than a site of interspecies dissimilarity.

The regions of interspecies similarity and divergence in POMC are largely delineated by the pairs of basic amino acid residues that separate the domains of the POMC protein. The two segments of DNA sequence dissimilarity are located immediately preceding and immediately following the nucleotide sequence encoding ACTH, which is highly conserved. The pre-ACTH sequence divergence ends near an Arg-Arg pair that separates this segment from the carboxy-terminus of γ-MSH (position −47), and the post-ACTH divergence ends at a Lys-Lys pair at the amino terminus of β-MSH (position +79).

Despite interspecies differences involving base substitutions and possible DNA sequence rearrangements, the length of the peptide encoded by the POMC DNA segment extending from the 3′ end of the intervening sequence to the poly(A) addition site is similar in the bovine and human species. The nucleotide sequence of the 50 nucleotides preceeding the poly(A) addition site is highly conserved, suggesting that this region may have a specific functional relationship to the POMC gene that precedes it. Shorter segments of interspecies conservation of 3′ noncoding sequences have been observed for the globin (32, 33) and insulin (34–36) genes. In ACTH itself, 95% base sequence similarity is maintained within the two species. The few base substitutions that do occur fail to alter the amino acid sequence of the protein except at position +33. Similar conservation of DNA sequence in segments of the functionally cryptic pre-ACTH region suggests that this region may also encode peptides of considerable biological importance.

The amino acid sequence for human β-LPH predicted from the nucleotide sequence shows some differences from the amino acid sequence determined directly by Li and Chung (28). Because the human genomic DNA we have analyzed represents the progeny of a single DNA molecule isolated from one individual, this difference may reflect possible polymorphism within the POMC gene [e.g., as in globin (37)]. Potentially, spontaneously occurring nucleotide base substitutions during replication of λ or plasmid DNA might also yield such differences.

The availability of a cloned segment of human genomic DNA encoding POMC enables the analysis of adjacent DNA segments and the assignment of the POMC gene to a particular chromosomal location by *in situ* hybridization. Furthermore, the absence of noncoding intervening sequences in component peptide hormones of human POMC will allow the functional expression of the genomic DNA sequence in bacterial cells.

Note Added in Proof. Seidah *et al.* (38) have determined the amino acid sequence of the amino-terminal segment of human POMC. Their data confirm our interpretation that a noncoding intervening sequence exists in this region of the human gene. S. Nakanishi *et al.* (personal communication) recently have isolated genomic DNA encoding *bovine* POMC and have found an intron in the same vicinity.

We thank L. Polder and M. Bibb for their helpful advice. These studies were supported by Grants GM27241 and AI08619 from the National Institutes of Health and in part by a grant to S.N.C. from the American Cancer Society. M.C. is a postdoctoral research fellow of the Centre National de la Recherche Scientifique.

1. Nakanishi, S., Taii, S., Hirata, Y., Matsukura, S., Imura, H. & Numa, S. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 4319–4323.
2. Mains, R. E., Eipper, B. A. & Ling, N. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3014–3018.
3. Roberts, J. L. & Herbert, E. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4826–4830.
4. Roberts, J. L. & Herbert, E. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5300–5304.
5. Nakanishi, S., Inoue, A., Kita, T., Nakamura, M., Chang, A. C. Y., Cohen, S. N. & Numa, S. (1979) *Nature (London)* **278**, 423–427.
6. 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.
7. Li, C. H. & Chung, D. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 1145–1148.
8. Ling, N., Burgus, R. & Guillemin, R. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3942–3946.
9. Chrétien, M., Seidah, N. G., Benjannet, S., Dragon, N., Routhier, R., Motomatsu, T., Crine, P. & Lis, M. (1977) *Ann. N.Y. Acad. Sci.* **297**, 84–107.
10. Chrétien, M., Benjannet, S., Gossard, F., Gianoulakis, C., Crine, P., Lis, M. & Seidah, N. G. (1979) *Can. J. Biochem.* **57**, 1111–1121.
11. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157–1174.
12. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C. & Quon, D. (1978) *Cell* **15**, 687–701.
13. Kushner, S. R. (1978) in *Proceedings of the International Symposium on Genetic Engineering*, eds. Boyer, H. W. & Nicosia, S. (Elsevier, Amsterdam), pp. 17–23.
14. Chang, A. C. Y., Nunberg, J. H., Kaufman, R. J., Erlich, H. A., Schimke, R. T. & Cohen, S. N. (1978) *Nature (London)* **275**, 617–624.
15. Maniatis, T., Jeffrey, A. & van de Sande, H. (1975) *Biochemistry* **14**, 3787–3794.
16. McDonell, M. W., Simon, M. N. & Studier, F. W. (1977) *J. Mol. Biol.* **110**, 119–146.
17. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237–251.
18. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
19. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
20. Sanger, F. & Coulson, A. R. (1978) *FEBS Lett.* **87**, 107–110.
21. Davis, R. W., Simon, M. & Davidson, N. (1971) *Methods Enzymol.* **21D**, 413–428.
22. 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.
23. Blattner, F. R., Williams, B. G., Blechl, A. E., Thompson, K. D., Faber, H. E., Furlong, L. A., Grunwald, D. J., Kiefer, D. O., Moore, D. D., Schumm, J. W., Sheldon, E. L. & Smithies, O. (1977) *Science* **196**, 161–169.
24. Nunberg, J. H., Kaufman, R. J., Chang, A. C. Y., Cohen, S. N. & Schimke, R. T. (1980) *Cell* **19**, 355–364.
25. Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3683–3687.
26. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
27. Bibb, M., Schottel, J. L. & Cohen, S. N. (1980) *Nature (London)* **284**, 526–531.
28. Li, C. H. & Chung, D. (1976) *Nature (London)* **260**, 622–624.
29. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4853–4857.
30. Rogers, J. & Wall, R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1877–1879.
31. Sakano, H., Rogers, J. H., Huppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R. & Tonegawa, S.(1979) *Nature (London)* **277**, 627–633.
32. Proudfoot, N. J., Gillam, S., Smith, M. & Longley, J. I. (1977) *Cell* **11**, 807–818.
33. Nishioka, Y. & Leder, P. (1979) *Cell* **18**, 875–882.
34. Bell, G. I., Pictet, R. & Rutter, W. J. (1980) *Nature (London)* **284**, 26–32.
35. Cordell, B., Bell, G., Tischer, E., DeNoto, F. M., Ullrich, A., Pictet, R., Rutter, W. J. & Goodman, H. M. (1979) *Cell* **18**, 533–543.
36. Lomedico, P., Rosenthal, N., Efstratiadis, A., Gilbert, W., Kolodner, R. & Tizard, R. (1979) *Cell* **18**, 545–558.
37. Kan, Y. W. & Dozy, A. M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5631–5635.
38. Seidah, N. G., Benjannet, S., Routhier, R., DeSerres, G., Rochemont, J., Lis, M. & Chrétien, M. (1980) *Biochem. Biophys. Res. Commun.*, in press.