Proc. Natl. Acad. Sci. USA
Vol. 77, No. 8, pp. 4583–4587, August 1980
Biochemistry

Genomic organization of rat prolactin and growth hormone genes

(gene library screening/restriction mapping/electron microscopic mapping)

YUEH-HSIU CHIEN AND E. BRAD THOMPSON

Laboratory of Biochemistry, Division of Cancer Biology and Diagnosis, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205

Communicated by David R. Davies, May 6, 1980

ABSTRACT **Five overlapping cloned DNAs containing the rat prolactin gene and its flanking sequences, as well as one cloned DNA containing the rat growth hormone gene and its flanking sequences, were isolated from a chromosomal DNA library. They were characterized by restriction enzyme mapping and electron microscopy. In each gene, the structural gene sequence coding for mature mRNA of a length of about 1 kilobase is split into at least five segments by a minimum of four intervening sequences. The two genes are similar in the length and organization of their coding regions, consistent with the suggestion that they are derived from a common ancestral gene. However, the two genes differ greatly in the lengths of their intervening sequences. That leads to a total gene length of 10 kilobase pairs for the prolactin and 2.1 kilobase pairs for the growth hormone gene. At least one intervening sequence appears to be in the 5′ nontranslated regions of the prolactin and growth hormone mRNA coding sequences.**

The peptide hormones prolactin and growth hormone are synthesized and secreted by specialized cells of the pituitary. It has been suggested that they are of common evolutionary origin (1). The physiological regulation of these hormones by neurotransmitters and steroid and peptide hormones has been studied extensively (2–9). Because of these features, prolactin and growth hormone genes provide a good system for comparative structural and functional analysis of related genes and for the study of hormone-regulated gene expression.

The construction of cDNA clones of prolactin and growth hormone mRNA has been reported by other laboratories (6, 10, 11), but further analysis of these systems requires the isolation of genomic DNA fragments that contain the entire coding sequences and the regulating elements that may be involved in the expression of the genes.

This paper describes the isolation of rat prolactin and growth hormone genes from a rat chromosomal DNA library by using probes made from size-selected rat pituitary mRNA, and the analysis of the structural organization of these two genes.

MATERIALS AND METHODS

Probes, Preparation, and Screening the Library. Female Sprague–Dawley rats (≈150 g) were injected twice, at weekly intervals, with 25 μg of 17β-estradiol in sesame oil. Total RNA was isolated from pituitaries of the rats by using the guanidinium thiocyanate/cesium chloride method (12). The RNA was heat denatured and centrifuged on a 10–30% sucrose gradient in the presence of 0.1 M NaCl/0.01 M Tris·HCl/1 mM EDTA, pH 7.5/0.1% Sarkosyl with *Escherichia coli* 5S, 16S, and 23S rRNA as markers. The RNA obtained from the 12S region of the sucrose gradient was pooled and used as template for the synthesis of [^{32}P]cDNA probe. [^{32}P]cDNA was made with $(dT)_{12-18}$ as primer. A 100-μl reaction mixture contained 50 mM Tris·HCl at pH 8.3, 6 mM $MgCl_2$, 60 mM KCl, 1 mM each of dGTP and dTTP, 49 μM [^{32}P]dCTP (1.8 μmol/Ci), 94 μM [^{32}P]dATP (1.7 μmol/Ci), 20 mM dithiothreitol, and RNA at 6 μg/ml (1 Ci = 3.7×10^{10} becquerels). The reaction mixture was incubated at 37°C for 2 hr. The specific activity of the [^{32}P]cDNA synthesized was 10^9 cpm/μg. The rat chromosomal library has been constructed by Sargent *et al.* (13), who ligated fragments from partially *Eco*RI digested DNA from the liver of a single Sprague-Dawley rat to the arms of phage Charon 4A. Screening was with 150-mm plates containing 1.5×10^4 plaque-forming units per plate as described by Benton and Davis (14). Filters were pretreated, hybridized, and washed as described by Wahl *et al.* (15) with 10^6 cpm of probe in 2.5 ml of hybridization solution per filter for 24 hr. Positive plaques were purified by two further screening cycles with plating at much lower density.

Recombinant λ Charon 4A phages were grown and purified and the DNA was prepared as detailed by Maniatis *et al.* (16). Recombinant phages were handled under P2/EK2 containment as described by the National Institutes of Health Guidelines.

Restriction Analysis. Digestions with different restriction endonucleases were done as recommended by the supplier (New England BioLabs).

Cell-Free Translation and Product Analysis. Translations of mRNA in rabbit reticulocyte translation systems using [^{35}S]methionine were done as recommended by the supplier (Bethesda Research Laboratories, Rockville, MD). The translation products were analyzed either by two-dimensional gel electrophoresis (17) or by immunoprecipitation with antisera to prolactin and growth hormone in the presence of *Staphylococcus aureus* (18) and one dimensional $NaDodSO_4$/polyacrylamide gel electrophoresis (17).

Identification of Cloned DNAs by Translation of Hybrid-Selected RNA. One to 2 μg of purified cloned DNA was heat denatured and applied to a nitrocellulose filter of about 0.16 cm^2 as described (19). The DNA-filter was then finely sliced into squares and hybridized with RNA. The hybridization reaction was carried out at 44°C for 16 hr in 20 μl of a solution containing 6 μg of rat total pituitary RNA, 0.4 M NaCl, 0.01 M 1,4-piperazinediethanesulfonic acid (Pipes) at pH 6.4, 1 mM EDTA, and 70% (vol/vol) formamide. The filters were then washed 10 times with 0.15 M NaCl/0.015 M sodium citrate at 65°C and three times with 20 mM Tris·HCl/2 mM EDTA, pH 7.5 at 55°C. The last wash was incubated for 3 min at 55°C. Bound RNA was eluted from the filter by boiling for 1 min in 10 mM EDTA, pH 6, with tRNA at 15 μg/ml. The RNA was then recovered by ethanol precipitation and assayed by translation and immunoprecipitation as described above.

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: Pipes, 1,4-piperazinediethanesulfonic acid; SV40, simian virus 40; kb, kilobase or kilobase pairs.

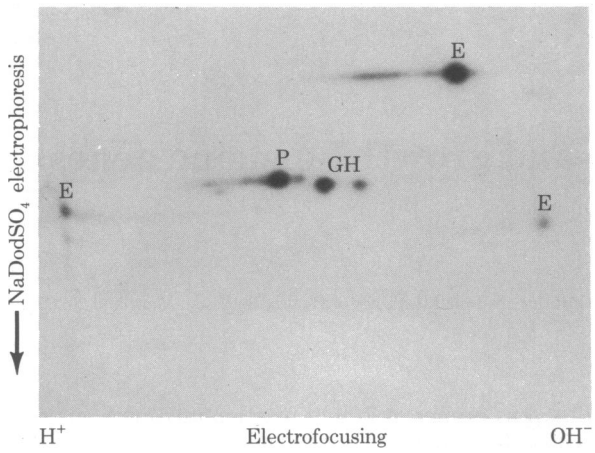

FIG. 1. Two-dimensional polyacrylamide gel analysis of *in vitro* translation products of 12S rat pituitary RNA labeled with [^{35}S]-methionine. The pH range used in the isoelectric focusing dimension was 3.5–10. The second dimension consisted of a 12.5% polyacrylamide gel. E indicates protein synthesized by the translation system in the absence of exogenous mRNA. P and GH indicate the preprolactin and pregrowth hormone spots. They exhibit molecular heterogeneity similar to that reported by Evans *et al.* (4).

Electron Microscopy. DNA·DNA heteroduplexes of various cloned DNAs were formed and spread as described by Davis *et al.* (20). The poly(A)$^+$ fraction of total pituitary RNA was obtained by oligo(dT) column chromatography (21) and used in R-loop formation with cloned DNA or restriction enzyme fragments of cloned DNA. The DNAs were crosslinked with psoralen, and the R-loops were prepared as described by Kaback *et al.* (22) except the glyoxal treatment step was omitted. The 3′ poly(A) of the RNA in the R-loop was mapped by using the simian virus 40 (SV40) poly(BrdUrd) labeling technique (23) except the SV40 poly(BrdUrd) and R-loops were incubated in 40% (vol/vol) formamide/0.08 M NaCl/0.05 M Pipes/5 mM EDTA, pH 6.4, at 4°C before spreading. SV40 and phage ϕX174 DNAs were used as double- and single-stranded standards.

RESULTS

Screening the Library and Restriction Endonuclease Mapping of the Clones. Previous studies have shown that functional preprolactin and pregrowth hormone mRNAs have sedimentation coefficients of about 12S, and that the prolactin and growth hormone sequences are the most abundant components of the 12S fraction of total pituitary RNA (10). The practicality of using size-selected pituitary RNA as template for the synthesis of [^{32}P]cDNA probes for prolactin and growth hormone sequences was tested as follows. Pooled 12S pituitary RNA was translated in the rabbit reticulocyte system. The products were analyzed in a two-dimensional polyacrylamide gel as shown in Fig. 1. The major spots have the same isoelectric points and molecular weights as reported by Evans *et al.* (4) for preprolactin and pregrowth hormone. The translational products were further identified as preprolactin and pregrowth hormone by immunoprecipitation with specific antisera to prolactin and growth hormone (data not shown). Thus, this preparation of mRNA seemed adequate for the preparation of probes.

This RNA fraction was then used as a template to synthesize a [^{32}P]cDNA probe for the screening of the rat chromosomal library. With the probe, 900,000 recombinant plaques were screened and 9 hybridizing clones were identified.

A schematic drawing of the restriction endonuclease maps of six of these clones is shown in Fig. 2. Clones 11, 55, 17, 16, and 6 contain overlapping *Eco*RI fragments as judged by restriction enzyme analysis and heteroduplex mapping in the electron microscope (data not shown). Clone 50 has no restriction fragments in common with the other five clones. One other clone (clone 13, not included in Fig. 2) contains a 16-kilobase pair (kb) insert consisting of three *Eco*RI fragments: 0.4, 2.6, and 13 kb. The 13-kb fragment hybridized weakly with the [^{32}P]cDNA probe. This clone has not been characterized further. The inserts of the remaining two positive recombinant phage contain the

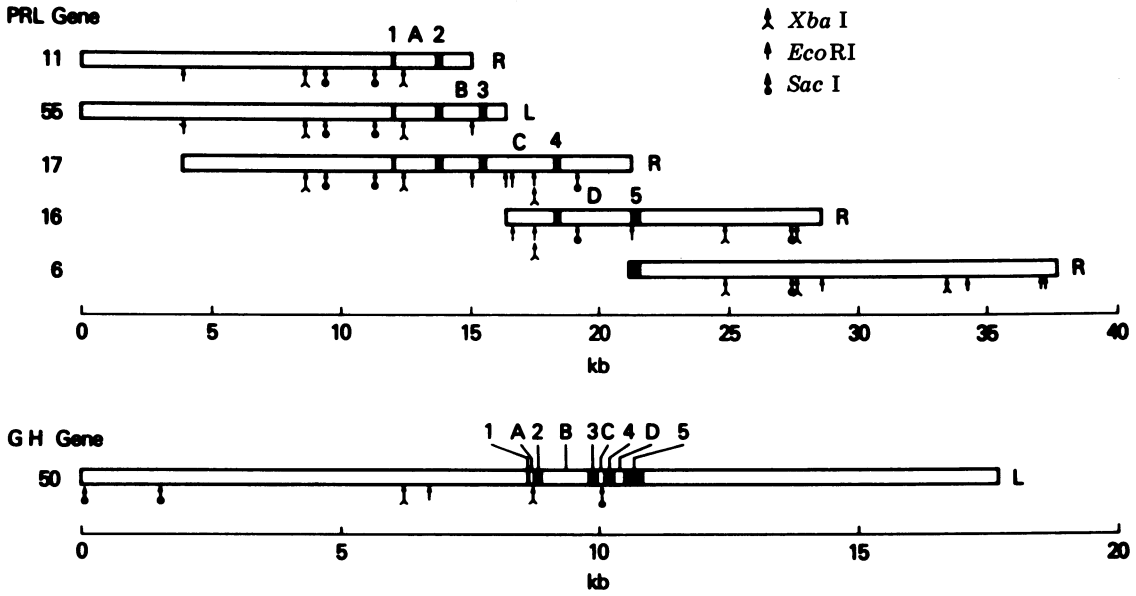

FIG. 2. Restriction endonuclease map and structural organization of prolactin (PRL) and growth hormone (GH) genes. This map was generated by restriction endonuclease, heteroduplex, and R-loop mapping. Black boxes represent prolactin and growth hormone structural gene sequences. They were designated as 1, 2, 3, 4, and 5. The white boxes A, B, C, and D represent intervening sequences. R and L indicate the right and left arms, respectively, of the λ Charon 4A vector. The identities of the clones were determined by the translation of filter-selected pituitary RNA (Fig. 3). The map is drawn so that the 5′-3′ orientation of the RNAs determined in Fig. 4 *b* and *d* is from left to right.

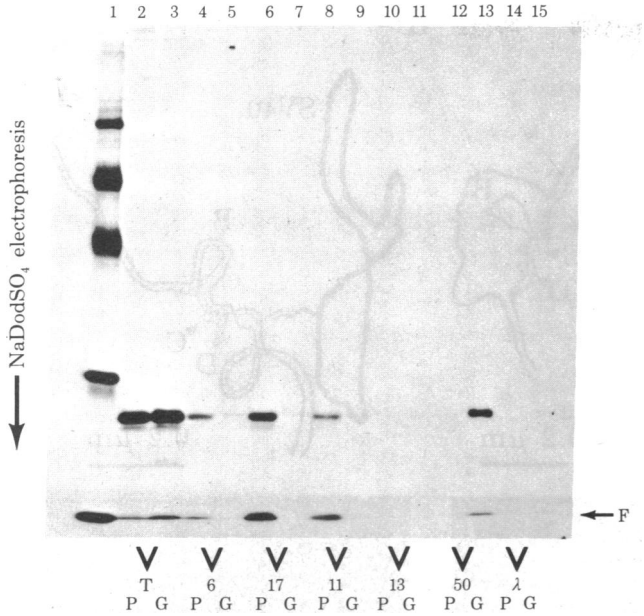

FIG. 3. Characterization of the translation products of RNA hybridized to cloned DNAs. F indicates the buffer front. Lane 1 contains molecular weight markers phosphorylase B (92,500), bovine serum albumin (69,000), ovalbumin (46,000), carbonic anhydrase (30,000), and cytochrome *c* (12,200; migrates with the buffer front). Lanes 2–13 represent pairs of experiments in which total RNA (lanes 2 and 3) and RNA selected by various cloned DNAs (clone numbers shown under the lanes) was translated. Aliquots of the translation mixtures were immunoprecipitated with antiserum to prolactin (indicated by P) and antiserum to growth hormone (indicated by G). Lanes 14 and 15 are control experiments showing the translational product of RNA selected by wild-type λ phage DNA after immunoprecipitation with antiserum to prolactin and growth hormone, respectively.

same *Eco*RI fragments as clones 6 and 17, but in the opposite orientations with respect to the phage arms.

Identification of Cloned Rat DNA Fragments. The coding abilities of the six recombinant DNAs were determined by cell-free translation of hybrid-selected RNA specifically bound to each cloned DNA. DNAs isolated from clones 6, 11, 13, 17, and 50 and wild-type phage λ were bound to filters, hybridized with total pituitary RNA, and washed, and the RNA was eluted. The protein products obtained by translating these RNA samples were analyzed by immunoprecipitation with antisera to prolactin and to growth hormone and gel electrophoresis as shown in Fig. 3. The results indicate that clone 13 and λ do not select RNA coding for either prolactin or growth hormone; the translational product of the RNA that hybridized to clones 6, 11, and 17 DNA is preprolactin; and the translational product of the RNA hybridized to clone 50 is pregrowth hormone. Therefore, clones 6, 11, and 17 contain sequences coding for prolactin and clone 50 contains sequences coding for growth hormone. Because of the overlapping nature of the DNAs in clones 11, 55, 17, 16, and 6, clones 55 and 16 must also contain prolactin sequences.

Sequence Organization of Prolactin and Growth Hormone Genes. The R-loop formed between the poly(A)$^+$ RNA and clone 17 contains four structural gene segments—1, 2, 3, and 4—separated by three intervening sequences—A, B, and C—located within the 7.8-kb *Sac* I fragment (see Fig. 2), as shown in Fig. 4*a*. The 7.3-kb *Xba* I fragment of clone 16 contains the structural gene segments 4 and 5 separated by intervening sequence D (Fig. 4*b*). The lengths are shown in Table 1. The position of structural gene segment 5 was mapped to within 50 nucleotides from the external *Eco*RI site of the 7.4-kb *Eco*RI fragment of clone 6, and the structural gene segment 1 was mapped at 8.1 ± 0.4 kb (± SD, n = 23) from the internal *Eco*RI site of the 11.2-kb *Eco*RI fragment of clone 11 (data not shown).

The direction of transcription was determined by hybridization of the 3′-terminal poly(A) of the RNA in the R-loop with the SV40 poly(BrdUrd) as shown in Fig. 4*b*. The orientation is given in Fig. 2.

The R-loops formed between the pituitary poly(A)$^+$ RNA and the growth hormone clone, clone 50, also show four intervening sequences—A, B, C, and D—separating five structural gene segments—1, 2, 3, 4, and 5 (Fig. 4*c*). The lengths are given in Table 1. The R-loop formed with the 11-kb *Eco*RI fragment of clone 50 shows that the structural gene segment 1 was mapped at 1.96 ± 0.09 kb (n = 22) from the internal *Eco*RI site, and the orientation of the transcription was determined as mentioned above (Fig. 4*d* and Fig. 2).

DISCUSSION

The results described in this paper show that it has been possible to screen a rat chromosomal DNA library with cDNA probes made from size-selected mRNAs to isolate clones containing prolactin and growth hormone sequences. Five overlapping cloned DNAs were isolated and found to contain prolactin sequences. These clones represent a total of 38 kb of DNA from the rat genome, with the prolactin sequence situated approximately in the middle. The identities of the restriction maps in their overlapping parts suggest that the cloned sequences are representative of uncloned genomic DNA.

A single clone of about 17.7 kb of rat genomic DNA containing growth hormone sequence was isolated. The growth hormone gene was situated within one 11-kb *Eco*RI fragment.

Table 1. Length of the structural gene segments and intervening sequences of the prolactin gene and the growth hormone gene

Structural gene segment	Intervening sequence	Mean length ± SD, kb: Prolactin gene	Mean length ± SD, kb: Growth hormone gene
1		<0.05	<0.05
	A	1.65 ± 0.11 (n = 25)	<0.1
2		0.17 ± 0.03 (n = 30)	0.16 ± 0.03 (n = 12)
	B	1.70 ± 0.13 (n = 30)	0.94 ± 0.03 (n = 19)
3		0.15 ± 0.03 (n = 33)	0.16 ± 0.04 (n = 16)
	C	2.67 ± 0.14 (n = 25)	<0.1
4		0.18 ± 0.04 (n = 33)	0.19 ± 0.04 (n = 14)
	D	3.0 ± 0.14 (n = 17)	<0.1
5		0.19 ± 0.05 (n = 20)	0.33 ± 0.06 (n = 17)

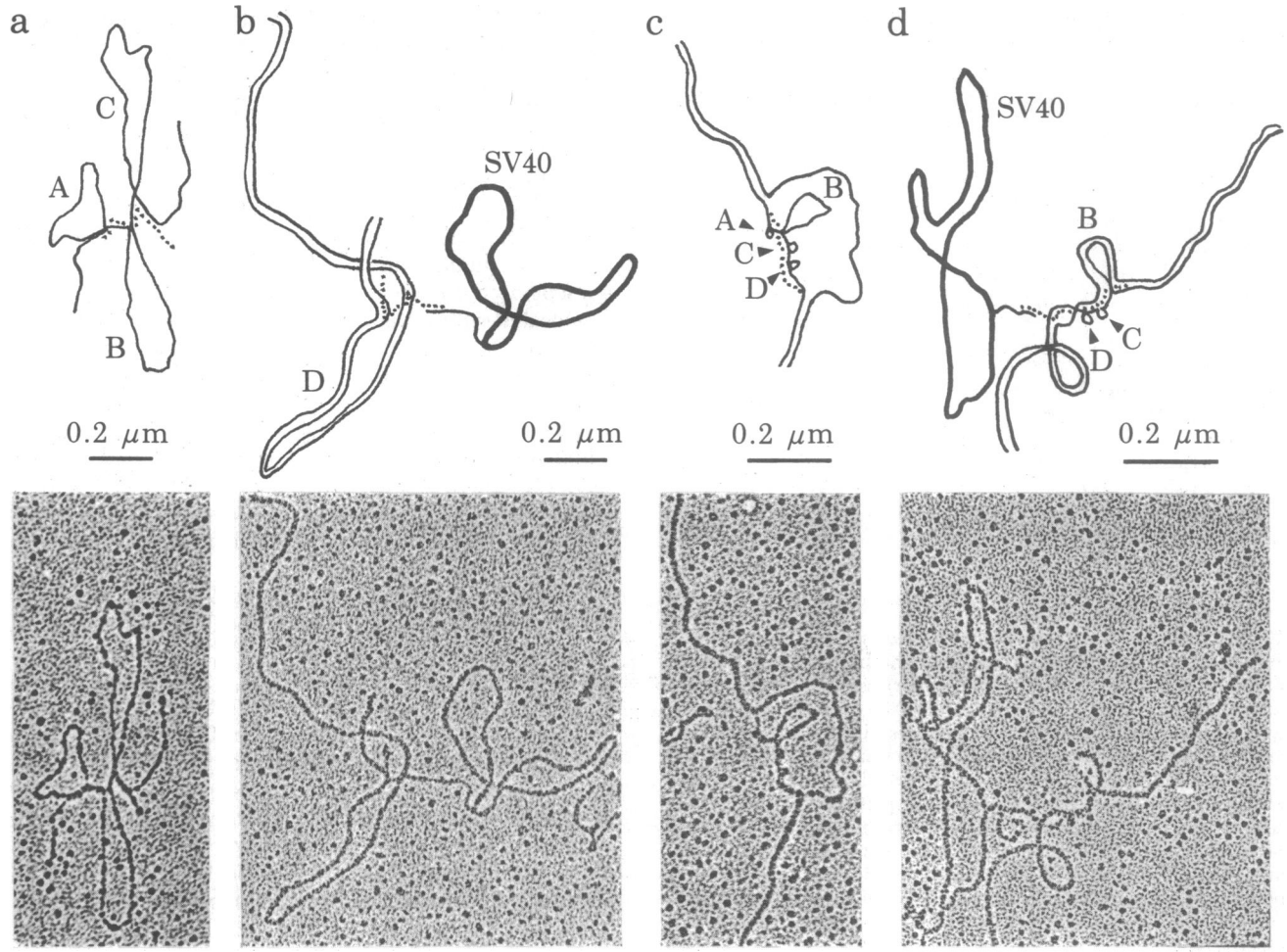

FIG. 4. Electron micrographs of R-loops and hybrids. (*a*) An RNA·DNA heteroduplex formed between pituitary poly(A)$^+$ RNA and the 7.8-kb *Sac* I fragment of clone 17. A, B, and C represent intervening sequences separating structural gene segments 1, 2, 3, and 4. (*b*) R-loop formed between pituitary poly(A)$^+$ RNA and the 7.3-kb *Xba* I fragment of clone 16 DNA, with the 3′ poly(A) of the RNA hybridized to a poly-(BrdUrd)-tailed SV40 DNA molecule. D represents the intervening sequence between structural gene segments 4 and 5. (*c*) R-loop formed between pituitary poly(A)$^+$ RNA and the 11-kb *Eco*RI fragment of clone 50 DNA. A, B, C, and D represent intervening sequences separating structural gene segments 1, 2, 3, 4, and 5. (*d*) R-loop formed between pituitary poly(A)$^+$ RNA and the 11-kb *Eco*RI fragment of clone 50 DNA with the 3′ poly(A) of the RNA hybridized to a poly(BrdUrd)-tailed SV40 DNA molecule. The dotted lines in the drawings represent RNA. The length of SV40 DNA molecules (5.27 kb) is taken as 1.76 μm.

It is not yet known whether these two hormone genes are linked. However, from the present data, we can conclude that if the two genes are linked the intergene distance must be larger than 19 kb.

It has been reported that mature mRNAs coding for prolactin and growth hormone are about 1 kb in length (2, 5). We have established the genomic organization of both genes within the resolution of electron microscopy. It is clear that, in both cases, the structural gene sequence is split into at least five segments by a minimum of four intervening sequences, leading to total gene lengths of 10 kb for the prolactin gene and 2.1 kb for the growth hormone gene. The DNA sequence determined from the cDNA clones showed that there are at least 17 nucleotides (R. Maurer, personal communication) and 22 nucleotides (6) at the 5′ untranslated regions of the prolactin and growth hormone mRNAs, respectively. Judging from the length of the mRNAs and the length of the cDNA clones, the 5′-terminal untranslated region of the mRNAs is likely to be longer than 50 nucleotides for both genes. If this assumption is correct, then our data suggest that the 5′ end (approximately 50 nucleotides or less) of the nontranslated regions is not contiguous with the remaining coding segments in either gene. Such leader sequences have also been seen in viral mRNAs and mammalian cellular mRNAs (24–27). Two limitations of our conclusions must be considered. The R-loops are only accurate to about 50 nucleotides. If a short (about 30 nucleotides) leader sequence exists at the 5′ end of the gene and is separated from the structural gene segment 1 by another intervening sequence, then the R-loop would not reveal such a structure. The same argument holds true for the 3′ end of the gene. Several very small (0.1 kb) intervening sequences were observed. It is quite possible that additional smaller intervening sequences have not been detected. The total length of structural gene segments of the growth hormone gene is 0.9 kb, which is close to the estimated length of the mRNA, excluding the poly(A) stretches at the 3′ end. However, the total length of the structural gene segments of the prolactin gene is 0.7 kb, which is shorter than the estimated length of the mRNA. This is probably due to technical problems, including branch migration of the boundary between RNA·DNA hybrid and duplex DNA. Irrespective of these reservations, it appears that the coding sequences of prolactin and growth hormone genes are similar in size and organization, but the lengths of the intervening sequences differ greatly. If these two genes have arisen by duplication as has been suggested, then the ancestral gene may have had the same organization of coding sequences as the present-day prolactin and growth

hormone genes. Under this assumption, extensive changes in the length of the intervening sequences must have taken place since the divergence of the two genes. Whether the coding sequence segments define discrete functional domains in the proteins as suggested by Gilbert (28) remains to be tested.

We thank T. Sargent and J. Bonner for making their rat chromosomal DNA library available to us; N. Davidson for providing laboratory facilities for screening the DNA library; P. Dannies for providing the antisera to prolactin and growth hormone; N. Davidson, M. Davis, I. Dawid, E. Fyrberg, and J. Mullins for advice; and C. Buckler, M. Davis, I. Dawid, E. Long, M. Singer, and J. Strobl for comments on the manuscript.

1. Wallis, M. (1971) *Biochem. J.* **125,** 54–56.
2. Ryan, R., Shupnik, M. A. & Gorski, J. (1979) *Biochemistry* **18,** 2044–2048.
3. Dannies, P. S. & Tashjian, A. H., Jr. (1976) *Biochem. Biophys. Res. Commun.* **70,** 1180–1189.
4. Evans, G. A., David, D. N. & Rosenfeld, M. G. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 1294–1298.
5. Tushinski, R. J., Sussman, P. M., Yu, L.-Y & Bancroft, F. C. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 2357–2361.
6. Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. (1977) *Nature (London)* **270,** 486–494.
7. Shapiro, L. E., Samuels, H. H. & Yaffe, B. M. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 45–49.
8. Seo, H., Vassart, G., Brocas, H. & Refetoff, S. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 2054–2058.
9. MacLeod, R. M. (1976) in *Frontiers in Neuroendocrinology,* eds. Martial, L. & Ganong, W. F. (Raven, New York), Vol. 4, p. 159.
10. Gubbins, E. J., Maurer, R. A., Hartley, J. L. & Donelson, J. E. (1979) *Nucleic Acids Res.* **6,** 915–930.
11. Harpold, M. M., Dobner, P. R., Evans, R. M. & Bancroft, F. C. (1978) *Nucleic Acids Res.* **5,** 2039–2053.
12. Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tiseher, E., Rutter, W. J. & Goodman, H. M. (1977) *Science* **196,** 1313–1319.
13. Sargent, T. D., Wu, J., Sela-Trepat, J. M., Wallace, R. B., Reyes, A. A. & Bonner, J. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3256–3260.
14. Benton, W. D. & Davis, R. W. (1977) *Science* **196,** 180–182.
15. Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3683–3687.
16. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978) *Cell* **15,** 687–701.
17. O'Farrell, P. H. (1975) *J. Biol. Chem.* **250,** 4007–4021.
18. Kessler, S. W. (1975) *J. Immunol.* **115,** 1617–1624.
19. Ricciardi, R. P., Miller, J. G. & Roberts, B. E. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 4927–4931.
20. Davis, R. W., Simon, M. & Davidson, N. (1971) *Methods Enzymol.* **21,** 413–428.
21. Rosbash, M. & Ford, P. J. (1974) *J. Mol. Biol.* **83,** 87–101.
22. Kaback, D. B., Angerer, L. M. & Davidson, N. (1979) *Nucleic Acids Res.* **6,** 2499–2517.
23. Bender, W. & Davidson, N. (1976) *Cell* **7,** 595–607.
24. Chow, L. T., Gelinas, R. Z., Broker, T. R. & Roberts, R. J. (1977) *Cell* **12,** 1–2.
25. Klessig, D. F. (1977) *Cell* **12,** 9–21.
26. Dugaiczyk, A., Woo, S. L. C., Cobert, D. A., Lai, E. C., Mace, M. L. & O'Malley, B. W. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 2253–2259.
27. Gannon, F., O'Hare, K., Perrin, F., Lepennec, J. P., Benoist, C., Cochet, M., Roeathnach, R., Royal, A., Garapin, A., Cami, B. & Chambon, P. (1979) *Nature (London)* **278,** 428–434.
28. Gilbert, W. (1978) *Nature (London)* **271,** 501.