T. Shenk, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1983), p. 45.
17. G. Levinson, G. Khoury, G. Vande Woude, P. Gruss, *Nature (London)* **295**, 568 (1982).
18. H. Scholer and P. Gruss, in *Enhancers and Eukaryotic Gene Expression*, Y. Gluzman and T. Shenk, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1983), p. 16.
19. S. Gillies and S. Tonegawa, *Nucl. Acids Res.* **11**, 7981 (1983); S. Gillies, S. Morrison, V. Oi, S. Tonegawa, in *Enhancers and Eukaryotic Gene Expression*, Y. Gluzman and T. Shenk, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1983), p. 121; J. Baneji, L. Olson, W. Schaffner, *ibid.*, p. 136.
20. A. R. Kimmel, C. Lai, R. Firtel, in *ICN-UCLA Symposium on Molecular and Cellular Biology*, R. Axel, T. Maniatis, C. F. Fox, Eds. (Academic Press, New York, 1979), vol. 14, p. 195; A. R. Kimmel and R. Firtel, *Cell* **16**, 787 (1979).
21. J. G. Sutcliffe, R. J. Milner, F. E. Bloom, R. A. Lerner, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 4942 (1982).
22. J. G. Sutcliffe, R. J. Milner, J. M. Gottesfeld, R. A. Lerner, *Nature (London)* **308**, 237 (1984).
23. R. J. Milner, F. E. Bloom, C. Lai, R. A. Lerner, J. G. Sutcliffe, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 713 (1984).
24. P. Brickell, D. Latchman, D. Murphy, K. Willison, P. Rigby, *Nature (London)* **306**, 756 (1983).
25. W. Klein, T. Thomas, C. Lai, R. Scheller, R. Britten, E. Davidson, *Cell* **14**, 889 (1978); J. W. Posakony, G. Flytzanis, R. Britten, E. Davidson, *J. Mol. Biol.* **167**, 361 (1983).
26. R. J. Milner and J. G. Sutcliffe, *Nucl. Acids Res.* **11**, 5497 (1983).
27. J. G. Sutcliffe, R. J. Milner, T. M. Shinnick, F. E. Bloom, *Cell* **33**, 671 (1983); J. G. Sutcliffe and R. J. Milner, *Trends Biochem. Sci.* **9**, 95 (1984).
28. J. O. Thomas and R. J. Thompson, *Cell* **10**, 633 (1977).
29. P. D. Greenwood and I. R. Brown, *Neurochem. Res.* **7**, 965 (1982); I. R. Brown, *Biochim. Biophys. Acta* **698**, 307 (1982); C. C. Kuenzle *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **47**, 493 (1983).
30. D. Lenoir, E. L. F. Battenberg, F. E. Bloom, P. Danielson, J. G. Sutcliffe, R. J. Milner, unpublished observations.
31. P. L. Deininger, D. J. Jolly, C. M. Rubin, T. Friedmann, C. W. Schmid, *J. Mol. Biol.* **151**, 17 (1981).
32. J. Arriza, A.-P. Chen, P. Danielson, C. Lai, R. J. Milner, J. G. Sutcliffe, unpublished observations.
33. P. A. Weil, J. Segall, B. Harris, S.-Y. Ng, R. G. Roeder, *J. Biol. Chem.* **254**, 6163 (1979).
34. J. L. Manley, A. Fire, A. Cano, P. A. Sharp, M. L. Gefter, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3855 (1980).
35. D. P. Carlson and J. Ross, *Cell* **34**, 357 (1983).
36. W. Hahn, personal communication.
37. N. Chaudhari and W. E. Hahn, *Science* **220**, 924 (1983); W. E. Hahn, N. Chaudhari, L. Beck, K. Wilber, D. Peffley, *Cold Spring Harbor Symp. Quant. Biol.* **48**, 465 (1983).
38. I. Lemischka and P. Sharp, *Nature (London)* **300**, 330 (1982); L. Schuler, J. Weber, J. Gorski, *ibid.* **305**, 149 (1983); F. G. Gonzalez and C. B. Kasper, *J. Biol. Chem.* **258**, 1363 (1983).
39. R. McKinnon, T. M. Shinnick, J. G. Sutcliffe, unpublished observations.
40. J. Banerji, L. Olson, and W. Schaffner, in *Enhancers and Eukaryotic Gene Expression*, Y. Gluzman and T. Shenk, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1983), p. 136.
41. W. Reynolds and J. Gottesfeld, in preparation.
42. J. Gottesfeld and L. S. Bloomer, *Cell* **28**, 781 (1982); D. F. Bogenhagen *et al.*, *ibid.*, p. 413.
43. S. Saragosti, G. Moyne, M. Yaniv, *ibid.* **20**, 65 (1980); E. B. Jakobovitz, S. Bratosin, Y. Aloni, *Nature (London)* **285**, 263 (1980); A. J. Varshavsky, O. H. Sundin, M. J. Bohn, *Cell* **16**, 453 (1979); J. Nickol, M. Behe, G. Felsenfeld, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1771 (1982).
44. We thank R. McKinnon, T. Shinnick, D. Pallotta, F. Bloom, R. Lerner, J. Arriza, A.-P. Chen, P. Danielson, and C. Lai for numerous discussions and collaborations on unpublished experiments; B. Hahn for communicating unpublished results; and L. Eddy for manuscript assistance. Supported in part by a grant from the National Institutes of Health (GM32355) and McNeil Laboratories. Publication 3529-MB from the Research Institute of Scripps Clinic.

# Alternative RNA Processing: Determining Neuronal Phenotype

Michael G. Rosenfeld, Susan G. Amara, Ronald M. Evans

*Summary.* On the basis of an analysis of the human and rat calcitonin genes and of a related gene, alternative RNA processing represents a developmental strategy of the brain to dictate tissue-specific patterns of polypeptide synthesis. This regulation allows the calcitonin gene to generate two messenger RNA's, one encoding the precursor of a novel neuropeptide, referred to as CGRP, which predominates in the brain, and the second encoding the precursor to the hormone calcitonin which predominates in thyroid C cells. The distribution of CGRP in the central and peripheral nervous system and in endocrine and other organ systems suggests potential functions in nociception, ingestive behavior, cardiovascular homeostasis, and mineral metabolism.

Complex regulatory mechanisms restrict the expression in neural tissues of genes encoding neuroendocrine peptides to precise groups of neurons. Understanding the developmental mechanisms responsible for such regulation is likely to provide general insights into the molecular strategies critical for brain development and function. Regulation of gene expression in the brain, as in other tissues, could occur during gene transcription, RNA processing and transport, messenger RNA (mRNA) stability, and posttranslational events (*1*). Because RNA processing regulation has the potential to qualitatively, as well as quantitatively, alter the nature of the gene product, it would present specific advantages in expression of certain genes. Analysis of calcitonin gene expression in neural tissues suggests that differential RNA processing events are one type of developmental regulation specifying the pattern of neuroendocrine gene expression (*2–6*), and could represent a common strategy in expression of certain genes dictating neural development. Alternative RNA processing occurs in a tissue-specific fashion to produce alternative polypeptide products (*5, 6*) and, therefore, serves to increase the diversity of neuropeptides generated from a single genomic locus.

## A Single Neuroendocrine Gene Generates Multiple RNA Products

The preparation and DNA sequence analysis of plasmids containing DNA complementary to rat calcitonin mRNA predicted the structure of the protein precursor to calcitonin, a 32–amino acid calcium-regulating hormone, produced in thyroid C cells (Fig. 1) (*2, 7, 8*). In addition to calcitonin, proteolytic processing of the precursor generates in thyroid C cells an 82–amino acid $NH_2$-terminal peptide and a 16–amino acid COOH-terminal calcitonin cleavage product (CCP) (Fig. 1). The production of multiple calcitonin-related mRNA's was first noted during the spontaneous and permanent "switching" of serially transplanted rat medullary thyroid carcinomas (MTC's) from states of high to low or absent calcitonin production (*3, 9*). The unexpected explanation for the "switch" was that calcitonin gene transcription continued but generated a series of new, structurally distinct tran-

M. G. Rosenfeld is a professor and S. G. Amara is a Schering-Plough fellow of the Life Science Research Foundation in the Eukaryotic Regulatory Biology Program, University of California San Diego School of Medicine, La Jolla 92093. R. M. Evans is a senior member of the Molecular Biology and Tumor Virology Laboratory, Salk Institute, San Diego, California 92138.

scripts, referred to as calcitonin-gene-related peptide mRNA (CGRP mRNA) (*4, 5*). These new mRNA's encode a 16,000-dalton (16K) protein product containing no immunoreactive calcitonin (*4*). On the basis of genomic DNA and RNA hybridization (blotting) analyses, it was suggested that both CGRP and calcitonin mRNA's were generated by differential RNA processing from a single genomic locus (*4, 5*). That this model is appropriate has been shown by a complete analysis of the two mRNA products and of the calcitonin gene.

Examination of the sequences of the two genes products reveal that CGRP and calcitonin mRNA's share sequence identity through nucleotide 227 of the coding region, predicting that the initial 72 $NH_2$-terminal amino acids of each precursor are identical (Fig. 2). Each mRNA subsequently diverges entirely in nucleotide sequence, encoding distinct COOH-terminal domains (*5*). Protein processing signals within the COOH-terminal region of CGRP predict the excision of a 37–amino acid polypeptide containing an amidated phenylalanine residue at the COOH-terminal domain (*5*) (Fig. 1). On the basis of the structure of the calcitonin-CGRP gene (Fig. 1), production of calcitonin mRNA involves splicing of the first three exons, present in both mRNA's, to the fourth exon, which encodes the entire calcitonin-CCP sequence. Alternative splicing of the first three exons to the fifth and sixth exons, which contain the entire CGRP coding sequence and the 3′ noncoding sequences, respectively, results in production of CGRP mRNA. In this case the fourth exon is excised along with the flanking intervening sequences.

The discovery of CGRP mRNA and its encoded products necessitated a plan to determine if CGRP is physiologically expressed and, if so, to ascertain its potential functions. An approach that could resolve these questions regarding CGRP biology would then be a prototype for studying the large number of unknown neuropeptides that are likely to be discovered with the use of current molecular biological techniques.

The initial evidence of CGRP mRNA expression in the brain was provided by hybridization of CGRP exon-specific probes hybridized to poly(A)-rich (polyadenylated) RNA from rat trigeminal ganglia and hypothalamus (*6*). To provide evidence for the production of the predicted peptide in the brain and to determine the precise sites of synthesis, a strategy that combined histochemical and molecular biological approaches was used. Antibody to a synthetic peptide corresponding to the 14 COOH-terminal amino acids of CGRP was generated and used to determine the distribution of CGRP in neural tissues (*6*). Immunoreactive CGRP was present in a characteristic distribution in a large number of cell groups and pathways in the central nervous system distinct from that of any known neuropeptide (Fig. 3). The distribution of CGRP in pathways and neurons believed to serve specific sensory, integrative, and motor systems (*6, 10*) suggests several possible physiological roles for the peptide. The localization of CGRP immunoreactivity in the olfactory and gustatory systems, including taste buds, the hypoglossal, facial, and vagal nuclei, and in the hypothalamic and limbic regions strongly suggests that it may have a functional effect in ingestive behavior (*6*). Additional studies have revealed the presence of CGRP at the neuromuscular junctions in the striated muscle of upper esophagus (*11*) and in certain skeletal muscles, the first peptide identified at neuromuscular junctions in mammalian species. CGRP is present in small trigeminal and spinal sensory ganglion cells which relay thermal and nociceptive information to the brainstem and spinal cord (*6*). In the spinal ganglia CGRP-positive cells represent 30 to 50 percent of the total population of small

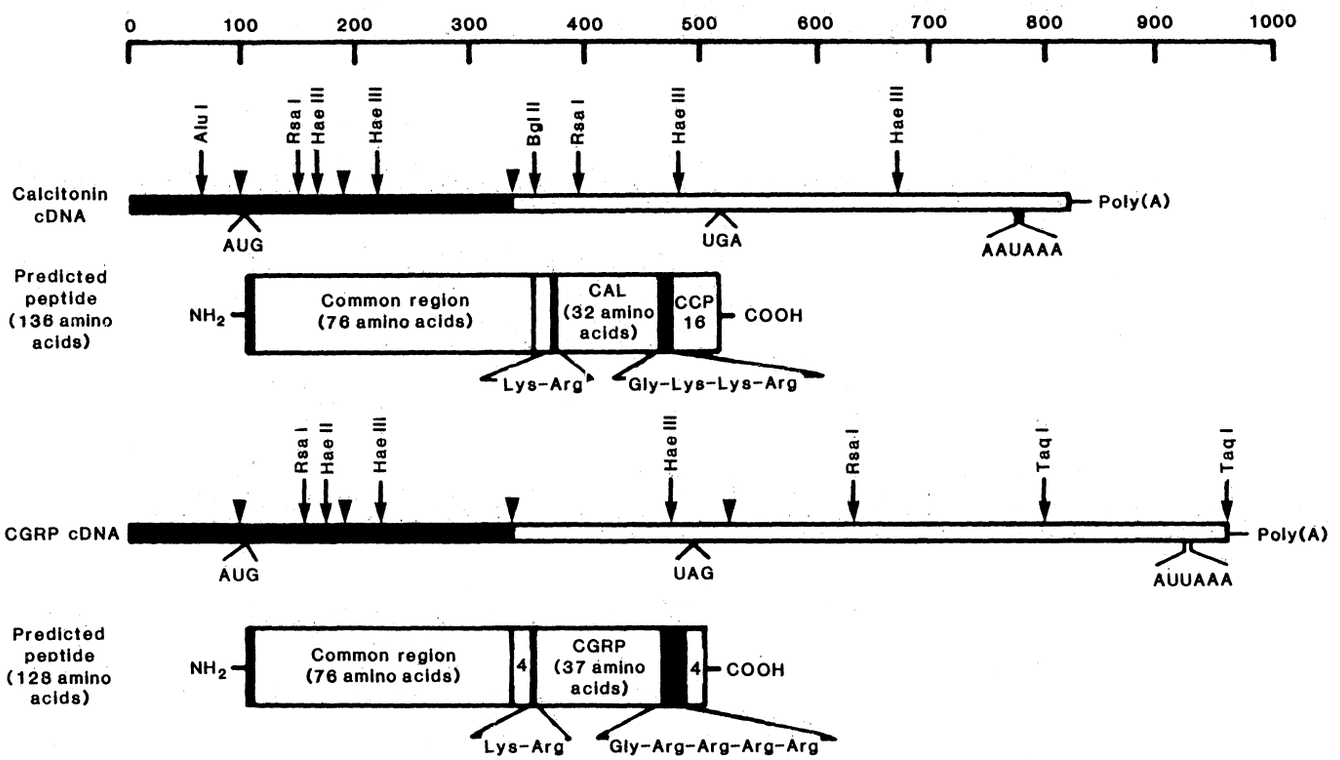


Fig. 1. Schematic diagram of rat calcitonin and CGRP cDNA's (mRNA) and their encoded peptides. The triangle indicates the point of sequence divergence between the two cDNA's. The two mRNA's share a 5′ region of identity (indicated by the dark bar); the initiator codon, stop codon, and canonical poly(A) signal sequence is indicated for each RNA. The structural organization of the encoded peptides, and the proteolytic processing signals are indicated in the boxed schematic diagrams of the precursor proteins.

ganglion cells, a percentage significantly greater than that for any other neuropeptide; for example, neurons stained with antiserum to substance P co-localize only to a subset of CGRP-containing small cells (*11*). CGRP is present in a subset of cells in one of the vagal motor nuclei (nucleus ambiguus) and is widely distributed in fibers in vascular musculature and in veins of virtually all organs, suggesting a role in cardiovascular homeostasis. Finally, CGRP-containing nerve fibers are widely distributed to most other organ systems, located mainly in the adventitia around arteries and veins, and in sensory fibers derived from the sensory ganglia or dorsal root ganglia. Consistent with certain features of its anatomical distribution, administration of synthetic rat CGRP produces profound effects on blood pressure, a characteristic pattern of catecholamine release in dogs and rats, and gastric hypoacidity (*12*). CGRP is also widely distributed in the endocrine system, in a subset of adrenal medullary cells, in bronchiolar cells, intestinal cells, and in fiber baskets which innervate the pancreatic islets and, interestingly, in thyroid C cells (*6, 11*).

S1 nuclease protection assay, mRNA-directed cell-free translation, and complementary DNA (cDNA) clonal sequence analyses confirmed the production of bona fide CGRP mRNA in the brain, and identified the sites of its biosynthesis. Gel filtration analysis of brain immunoreactive peptide suggests that this precursor is processed in brain to generate the predicted peptide product (CGRP) (*6*), and primary cultures of rat trigeminal ganglia appear to secrete authentic CGRP peptide (*13*). Tissue specificity of the RNA processing events is suggested because calcitonin mRNA is present at less than 0.5 percent of the level of CGRP mRNA in the brain (Table 1), while calcitonin and CGRP mRNA's and their encoded peptides are present in a ratio of 95:1 to 98:1, respectively, in thyroid C cells (*14, 15*). Histochemical analysis and radioimmunoassay indicate that in thyroid C cells small amounts of CGRP are present and that both calcitonin and CGRP can be coproduced within the identical cell (*16*).

Additional analyses revealed a distinct RNA bearing homology to CGRP mRNA. This mRNA encodes a 134–amino acid precursor peptide which, on the basis of the presence of excision signals, should generate five separate polypeptide products, one of which dif-

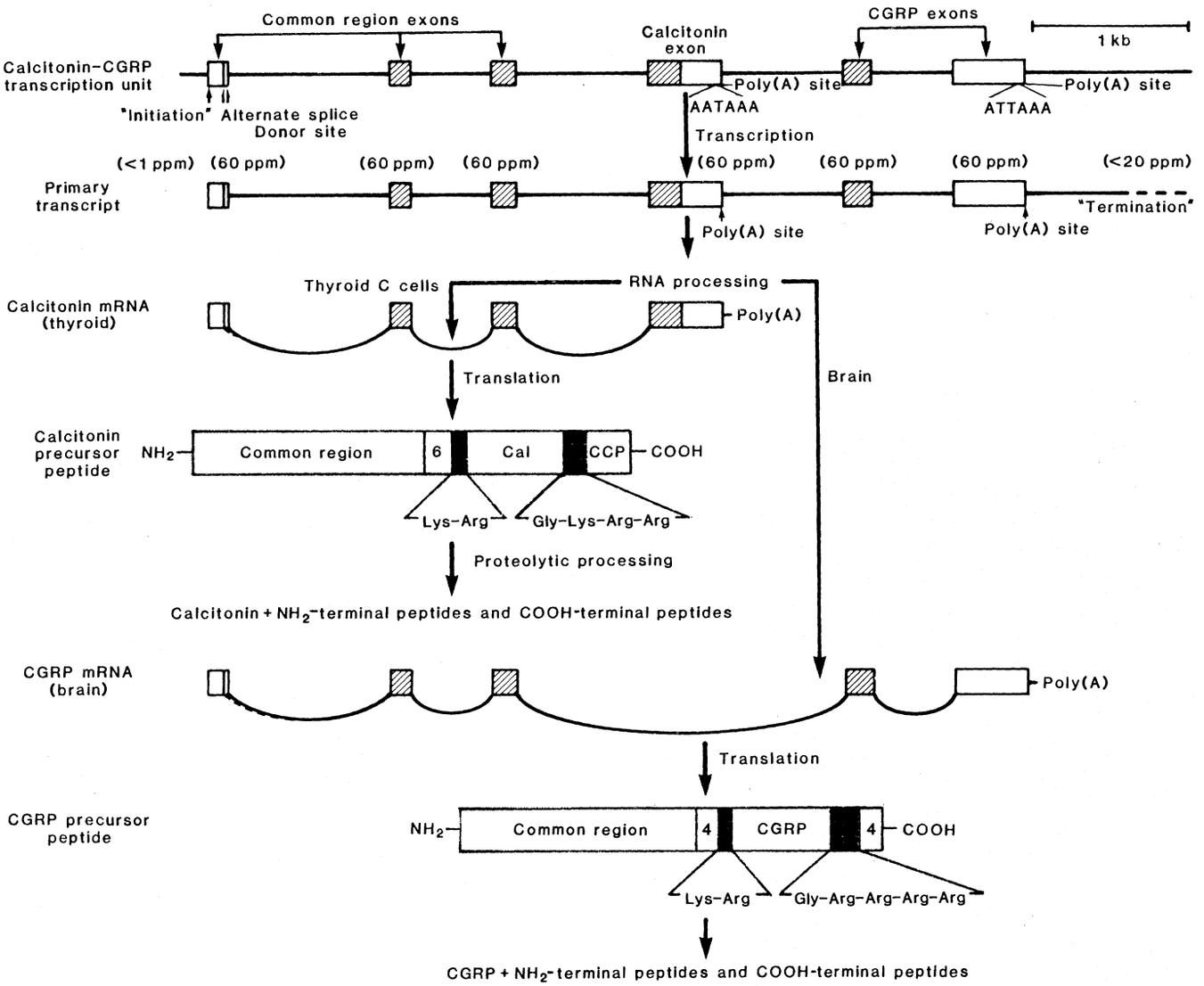


Fig. 2. Diagram of the rat calcitonin-CGRP gene and the alternative pattern of RNA processing of the primary transcript. The relative transcription rates in a calcitonin mRNA–producing tissue are indicated in parentheses (counts per minute per kilobase per $10^6$ counts); analysis was performed with clonal probes under conditions of DNA excess. The thyroid actually contains calcitonin and CGRP mRNA's in a ratio of 95:1 to 98:1. In medullary thyroid tumors, the relative production of CGRP increases dramatically. In brain most, or all, of the mature transcript is present as CGRP mRNA.

fers in primary sequence from CGRP by only a single internal amino acid residue (Lys instead of a Glu at residue 35) (Fig. 4) (*16*). This mRNA is therefore referred to as β-CGRP mRNA. In contrast, the RNA is quite divergent from CGRP mRNA in both 5′ and 3′ noncoding regions, and these regions do not cross-hybridize to CGRP and calcitonin mRNA's. When labeled restriction fragments from these 5′ and 3′ noncoding regions are used as hybridization probes, β-CGRP mRNA is detected in poly(A)-rich RNA prepared from trigeminal ganglia and midbrain areas. Because of structural differences, S1 nuclease mapping readily distinguishes CGRP and β-CGRP mRNA's.

In order to examine the potential operation of RNA processing regulation occurring in other species, the human calcitonin gene and gene products were cloned and sequenced. Human calcitonin mRNA (*17–19*) and CGRP mRNA share sequence identity through nucleotide 222 of the coding region, then diverge entirely to encode either calcitonin or CGRP. The predicted sequence of human CGRP differs from that of the rat in four amino acids. Interestingly, one of the alterations (the Lys for Glu in position 35) is identical to the single amino acid substitution between rat CGRP and β-CGRP. The structure of human CGRP predicted by the cDNA analysis concurs with the sequence of the peptide recently purified from human MTC tumors and determined by a novel fast atom bombardment mass spectrometry (*20*). Analysis of human genomic DNA clones confirmed that all of the coding information for both calcitonin and CGRP mRNA's are present within a single genomic locus, and that the point of divergence of these two mRNA's corresponds precisely to intervening sequence, exon boundaries. A second human gene which contains exons related to both common region, calcitonin, and CGRP exons has been identified, but its potential expression is, as yet, unproved.

Table 1. Relative distribution of CGRP mRNA in the rat brain, based on S1 nuclease analysis (*6*). The relative levels of CGRP are proportional to the percentage of cells that contain histochemically detectable material reactive with antiserum to CGRP. No calcitonin mRNA or immunohistochemically staining peptide was detected.

| Tissue | Relative mRNA (%) | |
|---|---|---|
| | CGRP | Calcitonin |
| Terminal ganglion | 100 | 0.1 |
| Trigeminal tract nucleus | 5 | 0.005 |
| Midbrain | 1 | ~0 |
| Hypothalamus | 1 | ~0 |
| Amygdala | <0.02 | ~0 |
| MTC clonal cell line | 4 | 3 |

Analysis of the sequence of human and rat CGRP shows a statistically significant relation to salmon calcitonin (*19*), suggesting that the calcitonin and CGRP exons arose from a common primordial genomic region. It is suggested that the complex calcitonin-CGRP gene arose either by duplication and sequence divergence of either the primordial calcitonin-like exon itself or via gene duplication and rearrangement. Because of this homology it is possible that a subset of the known actions of salmon calcitonin in humans results from its interaction with CGRP receptors.

Human and monkey calcitonin and CGRP mRNA's and their encoded peptides are expressed in the pituitary gland, and in the monkey it was shown to be localized to pars intermedia on the basis of immunocytochemical (*19*) analysis and radioimmunoassay (*15*). These data suggest that observations of histochemically detected material in rat pituitaries with antiserum to calcitonin (*21–23*), despite the very low or undetectable levels of calcitonin mRNA (*2, 24*), are likely to reflect actual expression of the calcitonin gene, even though there may be wide variation between species. We have determined that the content of CGRP mRNA in human pituitary is approximately 3 percent that of thyroid gland; therefore, the contribution of the pituitary to circulating CGRP would be expected to be proportionally small. On this basis, it would appear likely that, if the production of calcitonin and CGRP in the pituitary has functional significance, it might act within the pituitary gland itself.

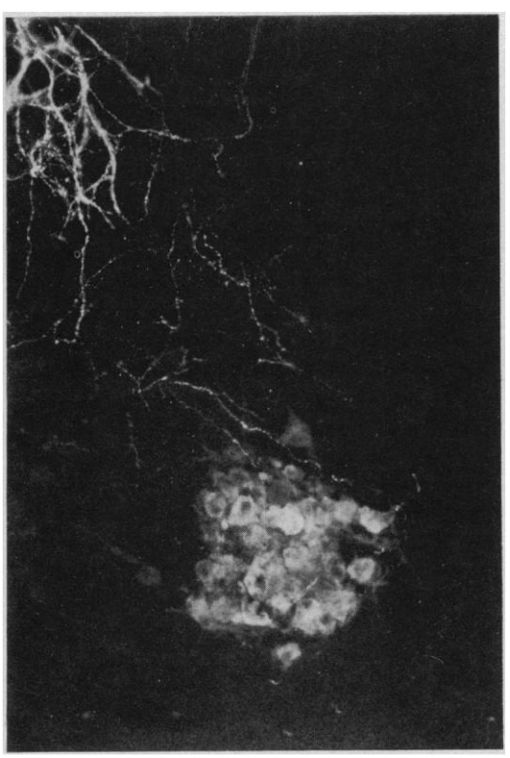

Fig. 3. Immunohistochemical demonstration of CGRP-containing neurons and fibers in the nucleus ambiguus and the spinal tract of cranial nerve V (*6*).

### Possible Mechanisms of Alternative RNA Processing

Several possible mechanisms could account for the production of multiple mRNA's from a single genomic locus. First, there could be two overlapping transcription units with alternative promoters that would generate mRNA's with different 5′ ends which might in turn dictate differential splicing pathways. Similarly, multiple mRNA's could derive from one transcription unit having a single site for initiation of RNA synthesis, but two or more poly(A) addition sites. The use of a particular poly(A) signal could be determined by alternative splicing events or by selective endonucleolytic cleavage adjacent to a poly(A) site. Finally, there could be differential regulation of splicing of the product (or products) of a single transcription unit, with both products having common initiation and poly(A) sites.

Several strategies, including primer extension and S1 nuclease resistance analyses, were used to demonstrate that calcitonin and CGRP mRNA's shared an identical transcriptional start (cap) site.

Therefore, both RNA's are products of a single transcription unit (*25*). As part of this analysis, a cryptic splice site was identified that generated a 24-nucleotide extension of the first exon in the case of both calcitonin and CGRP mRNA's. The utilization of this site appears to be a nonregulated event in all tissue in which the calcitonin-CGRP gene is expressed; it occurs in untranslated region of the message, and thus does not alter the encoded protein products (Fig. 2). Similar variations in RNA splicing patterns have been observed in the case of other transcripts, including those of the genes encoding human growth hormone, γ-fibrinogen, α-A crystallin (*26–28*), and possibly substance P (*29*).

The 3′ end of calcitonin and CGRP mRNA's were determined by S1 nuclease mapping experiments and confirmed by identifying the appropriate polyadenylation sequences in rat calcitonin cDNA and genomic clones (*25*). The polyadenylation site of calcitonin mRNA appears to be 18 or 19 nucleotides 3′ to a sequence AATAAA (A, adenine; T, thymine) located 226 nucleotides downstream of the calcitonin termination codon. CGRP mRNA utilizes a recognition sequence, ATTAAA, and the CGRP poly(A) addition signal and site are situated 1.9 kilobases downstream from the analogous region for calcitonin mRNA, defining the terminus of the large 3′-CGRP noncoding exon (the sixth genomic exon). Thus, production of calcitonin and CGRP mRNA's is associated with the selective polyadenylation of transcripts at one of two alternative poly(A) sites. The two mRNA's appear to have similar cytoplasmic stabilities so that their relative amounts are apparently reflective of their production, rather than differential turnover.

Two mechanisms may be considered for the selective utilization. First, transcripts for calcitonin mRNA could terminate at or adjacent to the calcitonin poly(A) site, while antitermination or readthrough would allow polymerase to continue to the next poly(A) site generating a primary transcript for CGRP. Second, transcripts could always terminate at or beyond the second poly(A) site, and specific RNA production could be determined by selective poly(A) addition and splicing patterns. To distinguish between these two possibilities, the distribution of nascent nuclear RNA transcripts across the calcitonin-CGRP transcription unit was quantitated. The apparent transcription rate across all regions from the cap site through to a fragment 0.64 kilobase downstream of the CGRP poly(A) site were essentially equivalent with marked decrease in transcription occurring approximately 1 kilobase downstream of the CGRP poly(A) site (see Fig. 1) (*25*). The results of these experiments indicate that selective RNA splicing and polyadenylation, and not alternative RNA transcriptional termination, are the regulated events.

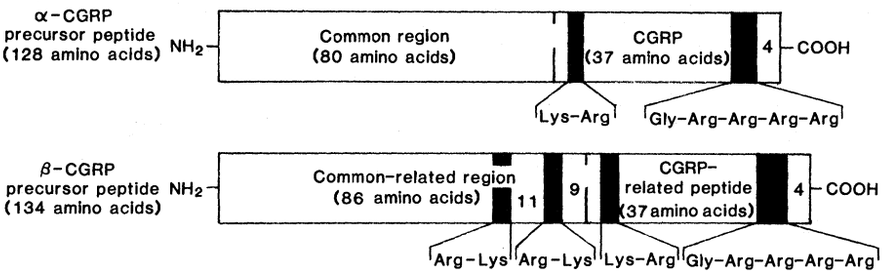


Fig. 4. Schematic diagram comparing the structures of the precursor of rat CGRP and CGRP-related peptide (β-CGRP) based on sequence analysis of cDNA clones.

The ability to introduce purified genes into heterologous eukaryotic cells provides an approach to further characterize the molecular events underlying alternative splicing and poly(A) site selection. Transfer of the calcitonin gene into several cell lines results in the appearance of both calcitonin and CGRP mRNA's in approximately equal ratios. In contrast, transfer of the rat calcitonin gene into terminally differentiated plasma cells resulted in clear patterns of selective mRNA production, suggesting that the machinery necessary to direct alternative RNA processing is operative in specific cell types, and is probably mediated by a *trans*-acting factor.

## Alternative RNA Processing Generality in Eukaryotic Gene Expression

Multiple mRNA's can be generated from single transcription units of several viral and eukaryotic genes (*30–41*). Both adenovirus and simian virus 40 (SV40) utilize alternative RNA splicing to generate a larger number of gene products than might be predicted from the length of their genomic DNA and the small number of transcription promoters, thereby maximizing the functional utilization of the limited genetic information. For adenovirus, this RNA polymorphism results, in part, from the use of 3′ polyadenylation sites that can be used to direct alternative splicing choices (*32*). Analysis of the expression of immunoglobulin heavy chain genes suggests that the biological demands of the immune system are partially met by differential posttranscriptional events in addition to the developmentally determined DNA rearrangements. During development of B cells, it appears that alternative poly(A) site selection, in association with alternative RNA processing dictate production of mRNA's encoding the membrane form ($\mu_M$) or secreted form ($\mu_S$), respectively, of immunoglobulin M (*34–36*) or immunoglobulin D (*38*) and the simultaneous synthesis of immunoglobulin heavy chain μ and δ (*37*). Similar events may occur in expression of several developmental genomic loci in *Drosophila* (*40–41*). Furthermore, the use of multiple transcriptional promoters is proposed to lead to alternative splicing patterns of the yeast invertase gene (*42*), the α-amylase gene (*43–46*), the myosin light chain gene (*47*), and the kininogen gene (*39*). The similarity of the alternative RNA processing events in genes of the neuroendocrine system (calcitonin-CGRP gene) and of the immune system (immunoglobulin heavy chain gene) suggest that the underlying mechanisms may operate in expression of other eukaryotic transcription units.

It is likely that multiple levels of regulation will be utilized in the developmental regulation of gene expression in the brain, as in other tissues. The calcitonin-CGRP gene is an example of a complex transcription unit able to generate two distinct gene products in a tissue-specific fashion. Our observations indicate that the RNA processing dictating poly(A) site selection may be an important mechanism for determining specificity of neuropeptide production and thus in establishing the phenotype of specific groups of neurons. Production of CGRP mRNA in the brain appears to result from the effects on a *trans*-acting regulatory factor. The control of poly(A) site selection in calcitonin-CGRP gene expression might be analogous to the mechanism regulating the formation of the 3′ termini of sea urchin histone H3 mRNA (*48*), which appears to involve the action of a 60-nucleotide RNA present as a small ribonucleoprotein particle (*49*). Therefore a nonabundant species of ribonucleoprotein expressed in neurons is a potential candidate for the putative

regulatory factor dictating the observed pattern of the processing of the calcitonin-CGRP transcript.

Identification of the factor (or factors) which determines the tissue-specific pattern of splicing and poly(A) site selection should allow for its isolation and detailed characterization, a prerequisite for understanding the mechanisms of its actions. On the basis of the potential utilization of this type of control in certain developmental loci of *Drosophila*, it is possible that alternative RNA processing is, in fact, a critical mechanism for control of neuronal development. While one consequence of such developmental regulation is that a single neuroendocrine gene can generate distinct, tissue-specific mRNA's, each encoding a different neuropeptide, there could be additional, widespread consequences with regard to productive or nonproductive processing of the transcripts of a large number of active brain transcription units. In such a case, one might envision that, in certain transcription units, selection of alternative poly(A) sites might result in (or from) a splicing pattern generating an unstable or an unproductive transcript. Only by evaluating the expression of genes encoding the factor (or factors) responsible for alternative RNA processing events in the brain, such as that required in the case of calcitonin-CGRP gene expression, can these questions be systematically addressed.

**References and Notes**


1. J. E. Darnell, *Nature (London)* **297**, 365 (1982).
2. S. G. Amara, D. N. David, M. G. Rosenfeld, B. A. Roos, R. M. Evans, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 4444 (1980).
3. M. G. Rosenfeld, S. G. Amara, B. A. Roos, E. S. Ong, R. M. Evans, *Nature (London)* **290**, 53 (1982).
4. M. G. Rosenfeld, C. R. Lin, S. G. Amara, L. S. Stolarsky, E. S. Ong, R. M. Evans (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1717 (1982).
5. S. G. Amara, V. Jonas, M. G. Rosenfeld, E. S. Ong, R. M. Evans, *Nature (London)* **298**, 240 (1982).
6. M. G. Rosenfeld *et al.*, *ibid.* **304**, 129 (1983).
7. J. W. Jacobs *et al.*, *Science* **213**, 457 (1981).
8. S. G. Amara *et al.*, *J. Biol. Chem.* **257**, 2129 (1982).
9. B. A. Roos *et al.*, *Endocrinology* **105**, 27 (1979).
10. L. W. Swanson and P. E. Sawchenko, in preparation.
11. S. J. Gibson *et al.*, *J. Neurosci.*, in press; J. M. Polak *et al.*, unpublished data.
12. L. A. Fisher *et al.*, *Nature (London)* **305**, 534 (1983); H. J. Lenz, M. T. Mortand, J. E. Rivier, M. R. Brown, in preparation.
13. R. T. Mason, R. A. Peterfreund, P. E. Sawchenko, A. Z. Corrigan, J. E. Rivier, W. W. Vale, *Nature (London)* **308**, 653 (1984).
14. M. I. Sabate *et al.*, in preparation.
15. F. A. Tschopp, P. H. Tobler, J. A. Fischer, *Mol. Cell. Endocrinol.* **36**, 53 (1984).
16. S. G. Amara, R. M. Evans, M. G. Rosenfeld, in preparation.
17. R. H. Craig, L. Hall, M. R. Edbrooke, J. Allison, I. MacIntyre, *Nature (London)* **295**, 345 (1982).
18. J. M. LeMoullec, A. Jullienne, J. Chenais, F. Lesmoles, J. M. Guliana, G. Milhaud, M. S. Monkhtar, *FEBS Lett.* **167**, 93 (1984).
19. V. Jonas, C. R. Lin, O. Kawashima, D. Semon, L. W. Swanson, J.-J. Mermod, R. M. Evans, M. G. Rosenfeld, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
20. H. R. Morris, M. Panico, T. Etienne, J. Tippins, S. I. Girgis, J. MacIntyre, *Nature (London)* **308**, 746 (1984).
21. P. Braga, S. Terri, A. Santagostino, V. R. Olgiatio, A. Perle, *Life Sci.* **22**, 971 (1978).
22. L. J. Deftos *et al.*, *J. Clin. Endocrinol. Metab.* **452**, 452 (1978); Mendelsohn, G. R. D'Agostino, J. C. Eggleston, S. B. Baylin, *J. Clin. Invest.* **63**, 1297 (1979).
23. M. Yamamoto, W. Tachukawa, H. Maeno, *Neuropharmacology* **20**, 83 (1981).
24. J. W. Jacobs, D. Gottzman, J. F. Habener, *Endocrinology* **11**, 2014 (1982).
25. S. G. Amara, R. M. Evans, M. G. Rosenfeld, *J. Mol. Cell. Biol.*, in press.
26. F. M. DeNoto, D. D. Moore, H. M. Goodman, *Nucl. Acids Res.* **9**, 3719; M. Wallis, *Nature (London)* **284**, 512 (1980).
27. G. R. Crabtree and J. Kant, *Cell* **31**, 159 (1982).
28. C. R. King and J. Piatigorsky, *ibid.* **32**, 707 (1983).
29. H. Nawa, T. Hirose, H. Takashima, S. Inayama, S. Nakanishi, *Nature (London)* **306**, 32 (1983).
30. J. Rogers, P. Early, C. Carter, K. Calame, M. Bond, L. Hood, R. Wall, *Cell* **20**, 303 (1980).
31. J. R. Nevins and J. E. Darnell, Jr., *ibid.* **15**, 1477 (1978).
32. E. B. Ziff, *Nature (London)* **287**, 491 (1980).
33. L. T. Chow and T. R. Broker, *Cell* **15**, 497 (1978).
34. J. Rogers *et al.*, *ibid.* **20**, 303 (1980).
35. F. W. Alt *et al.*, *ibid.*, p. 293.
36. P. Early, J. Rogers, M. Davis, K. Calame, M. Bond, R. Wall, L. Flow, *ibid.*, p. 313.
37. R. Maki *et al.*, *J. Mol. Biol.* **129**, 547 (1981).
38. H. L. Cheng, F. R. Blattner, L. Fitzmaurice, J. F. Mushinski, P. W. Tucker, *Nature (London)* **296**, 410 (1982).
39. N. Kitamura, Y. Takagaki, S. Furuto, T. Tanaka, H. Nawa, S. Nakanishi, *ibid.* **305**, 545 (1983).
40. C. E. Rozek and N. Davidson, *Cell* **32**, 23 (1983).
41. M. Carlson and D. Botstein, *ibid.* **28**, 145 (1982).
42. S. Heinkoff, J. S. Sloan, J. D. Kelly, *ibid.* **34**, 405 (1983).
43. M. Tosi, R. A. Young, O. Hagenbuchle, U. Schibler, *Nucl. Acids Res.* **9**, 2313 (1981).
44. O. Hagenbuchle, M. Tosi, U. Schibler, R. Bovey, P. K. Wellauer, R. M. Young, *Nature (London)* **289**, 643 (1981).
45. R. A. Young, O. Hagenbuchle, U. Schibler, *Cell* **23**, 451 (1981).
46. U. Schibler, O. Hagenbuchle, P. K. Wellauer, A. C. Pittet, *ibid.* **33**, 501 (1983).
47. Y. Nabeshima, Y. F. Kuriyama, M. Muramatsu, K. Ogata, *Nature (London)* **308**, 333 (1984).
48. H. G. Stunnenberg and M. L. Birnstel, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6201 (1982).
49. G. Galli, H. Hofstetter, G. Henrik, G. Stunnenberg, M. L. Birnstel, *Cell* **34**, 823 (1983).


50. We thank our colleagues C. Lin, L. Stolarsky, V. Jonas, J.-J. Mermod, E. S. Ong, L. Swanson, P. Sawchenko, W. Vale, M. Brown, and J. Rivier for their contributions to these investigations.


# DNA Markers for Nervous System Diseases


James F. Gusella, Rudolph E. Tanzi, Mary Anne Anderson
Wendy Hobbs, Kerin Gibbons, Roya Raschtchian
T. Conrad Gilliam, Margaret R. Wallace
Nancy S. Wexler, P. Michael Conneally


One of the major challenges in modern biology is to unravel the mystery of how the human brain functions. A method that has proved fruitful in other systems is analyzing mutations in order to define the normal mechanisms they disrupt. In man, genetic alterations affecting brain function are commonly expressed as inherited neurological or behavioral diseases. The development of recombinant DNA techniques has advanced our understanding of many human genetic diseases by permitting the isolation and sequencing of specific defective genes. It is now relatively straightforward to clone and characterize any gene whose product is known. Unfortunately, in most genetic diseases, the primary protein defect has yet to be identified. The task of finding a specific protein defect among the 40,000 to 200,000 possible proteins encoded in the human genome is enormous unless additional information is available to aid in the search. We have therefore embarked on the development of an alternative approach to the investigation of human genetic diseases. We intend to base our strategy for isolating a defective gene on its location in the human genome. The identification of an altered protein by analysis of the gene


James F. Gusella, Rudolph E. Tanzi, Mary Anne Anderson, Wendy Hobbs, Kerin Gibbons, Roya Raschtchian, and T. Conrad Gilliam are associated with the Neurology Service and Genetics Unit, Massachusetts General Hospital and Department of Genetics, Harvard Medical School, Boston 02114. Margaret R. Wallace and P. Michael Conneally are with the Department of Medical Genetics, Indiana University Medical Center, Indianapolis 46223. Nancy Wexler is president of the Hereditary Disease Foundation, 606 Wilshire Boulevard, Suite 504, Santa Monica, California 90401.