Structure and Function of the Chromogranin A Gene

CLUES TO EVOLUTION AND TISSUE-SPECIFIC EXPRESSION*

Hong-Jiang Wu, David J. Rozansky, Robert J. Parmer, Bruce M. Gill, and Daniel T. O’Connor†

From the Department of Medicine and Center for Molecular Genetics, University of California and Veterans Administration Medical Center, San Diego, California 92161

Chromogranin A is the index member of a family of acidic proteins stored and released throughout the neuroendocrine system with peptide hormones and neurotransmitters. To better understand its functional domains, its evolutionary lineage, and the basis of its tissue-specific pattern of expression, we obtained a mouse chromogranin A cDNA and used it to isolate the chromogranin A gene from the mouse genome. We then characterized the gene’s exon/intron structure, and the structure and function of its 5’-regulatory region (promoter/enhancer). The chromogranin A gene was complex, its eight exons and seven introns spanning about 11 kilobase pairs. The eight exons displayed some correspondence to putative functional domains suspected within the cDNA-deduced primary structure of the protein. Three exons also displayed both length and sequence homology to exons in another member of the chromogranin(secretogranin family, chromogranin B, suggesting an evolutionary relationship. A 1.2-kilobase pair genomic fragment just 5’ of the coding region was able to program cell type-specific gene expression in a transfection/reporter system. Both pituitary corticotropes and adrenal chromaffin cells recognized this promoter. The promoter possessed some known consensus transcriptional control elements (TATA box, cyclic AMP-response element, and Sp1 site), but otherwise novel transcriptional control elements seem to be operative. The results suggest that this complex gene is encoded by exon modules with evolutionary links to homologous modules in other chromogranin/secretogranin family members, and that the 5’-flanking region of the gene is sufficient to confer neuroendocrine tissue-specific expression of the chromogranin A gene.

Chromogranin A (1) is the index member of the chromogranin/secretogranin protein family (2), a group of acidic soluble proteins with a widespread if not universal occurrence in the soluble core of amine and peptide hormone and neurotransmitter storage vesicles throughout the neuroendocrine system (3-5).

Chromogranin A is the topic of many unanswered questions. Although it exists as a single copy gene in both the rodent and human genomes (6-10), the organization of its gene’s structure is unknown. The cDNA-deduced primary structure of chromogranin A suggests a number of functional domains in the protein (6-8), but whether such domains correspond to discrete exons has not yet been investigated. Finally, although the chromogranin A gene is selectively expressed at high levels within the neuroendocrine system (3-5), the basis of this tissue-specific expression is unexplored.

To address these issues, and ultimately to be able to manipulate the chromogranin A gene in an intact organism, we have isolated the gene from the mouse genome. We established its exon/intron structure, isolated and sequenced its 5’-regulatory region, and established its cell type-specific promoter activity by transfection and expression after reporter construction. Our results suggest that this complex gene is constructed of exon modules with evolutionary links to homologous modules in other chromogranin/secretogranin family members, and that the region just 5’ of exon 1 is sufficient to confer neuroendocrine tissue-specific expression of the chromogranin A gene.

MATERIALS AND METHODS

Mouse Chromogranin A cDNA Isolation—A mouse brain cDNA library in the vector λ-ZAP (Stratagene, La Jolla, CA) was screened by plaque hybridization with a 32P nick translation (11) labeled full-length rat chromogranin A cDNA (6). 43 positive clones were found in the 200,000 clones screened. Restriction digestion established that 3 of these were greater than 1.8 kb in length, and therefore likely full-length (6). The longest (clone 5-3, 1892 bp) of these clones was sequenced by the dideoxy chain termination (12) method (Sequenase Kit, United States Biochemicals, Cleveland, OH). Synthetic 17-mer oligonucleotide sequencing primers were synthesized to enable progressive sequencing runs spanning each strand of the entire cDNA (Model 381A automated oligonucleotide synthesizer, Applied Biosystems, Foster City, CA).

Sequence comparisons were made with Pustell sequence analysis programs (International Biotechnologies, Inc., New Haven, CT) using an IBM-PC-AT microcomputer.

Mouse Chromogranin A Genomic Clone Isolation—The 32P nick translation (11) labeled full-length mouse chromogranin A was used to screen 160,000 clones of a mouse (strain A/J) genomic DNA library in the vector sCos-1 (14, 15). Colony hybridization (13) yielded 42 positive clones. Since the poly-Gln trinucleotide repeat ((CAG)n) region of rodent chromogranin A recognizes dispersed sequences throughout the genome (6, 7), we rescreened the 42 positive clones with a 32P-end labeled (13) synthetic 20-mer oligonucleotide corresponding to the farthest 5’ portion of the 5’-untranslated region of the mouse chromogranin A cDNA: 5’-GCTGCTGCCGCCACC GC-3’. 3 of the 42 clones remained positive. Regions of overlap among

* The work was supported by the Department of Veterans Affairs, The National Kidney Foundation, and the National Institutes of Health. The costs of publication of this article were defrayed in part to indicate this fact.

† To whom correspondence should be addressed: Nephrology Hypertension (V-111-H), University of California, San Diego, 3350 La Jolla Village Drive, San Diego, CA 92161. Tel.: 619-552-8585 (ext. 7373); Fax: 619-552-7452.

The abbreviations used are: kb, kilobase pair; bp, base pair; ACTH, adrenocorticotropic hormone.
the 3 positive clones were established by restriction digestion, and 2 overlapping clones spanning a total of approximately 52 kb were restriction-mapped at EcoRI, XhoI, and BamHI sites using the phage T5 and T7 promoters flanking the inserts (14, 15; Gene Mapping Kit, Stratagene).

The exon/intron structure of the gene was established by resequencing the mouse chromogranin A coding region, but this time in exons within genomic DNA. As exon/intron borders were crossed, new primers were synthesized corresponding to the next exon, and sequencing continued. Positioning of each exon on the established genomic clone maps (14, 15) was accomplished by restriction digestion of the genomic clones with 9 different enzymes (EcoRI, XhoI, BamHI, EcoRV, HindIII, XbaI, SmaI, AccI, and PstI), followed either by Southern blotting (13), using exon-specific 32P-labeled (13) synthetic oligonucleotides (the previously mentioned sequencing primers) as probes, or by subcloning some of the genomic restriction fragments (n = 7) into pBluescript (Stratagene), after which the inserts were sequenced (12).

Promoter Analysis—The genomic DNA 387 bp 5' of the ATG initiation codon (or 230 bp 5' of the longest cDNA 5' end) was sequenced on both strands by the dideoxy chain termination method (12).

To establish the mRNA cap site, we performed primer extension (13) on an mRNA template, using an oligonucleotide primer corresponding to nucleotides −72 through −103 (numbering 5' of the ATG initiation codon) of the longest cDNA: 5'-GCTACCAAAGCTCT-GCACCCTGGAACACTG-3'. The source of the mRNA (13) was the mouse endocrine cell line AtT-20 (16), while the source of the negative control mRNA was mouse liver.

A 1.2-kb EcoRI-PstI putative promoter fragment was excised from the region 5' of exon 1 (Fig. 4); it spanned the region from about 1.2 kb upstream of the cap site to 54 bp downstream of the cap site (Figs. 4 and 7, Table I). This fragment was blunted (13) and ligated into the blunted HindIII promoter insertion site of the promoterless luciferase expression reporter vector pSV0ALΔ5 (17). Proper orientation of the promoter insert in the resulting construct (pHJAlΔ5) was verified by appropriate asymmetric double restriction digestion (13).

The plasmid (10−30 μl) of each promoter was transfected into both neuroendocrine and nonneuroendocrine cells and control mRNA  was mouse liver.

RESULTS

Mouse Chromogranin A cDNA—1892-bp mouse brain cDNA clone 5-3, when sequenced on both strands (Fig. 1) revealed a full-length 463-amino acid open reading frame bounded by initiation (ATG) and termination (TGA) codons. The cDNA contained a remarkably long 157-bp 5'-untranslated region. The deduced primary structure of mouse chromogranin A (Fig. 2) suggests an 18-residue hydrophobic N-terminal signal peptide, followed by a 445-amino acid mature protein. Its sequence homology to rat chromogranin A (6) is 16/18 residues identical (89% identity) in the signal peptide, followed by 397/445 residues identical (89% identity) in the mature protein. Analysis of the primary structure of mouse chromogranin A suggests several structural domains in the protein (Fig. 3) also shared by rat chromogranin A (6, 7, 23), including the signal peptide, 2 cysteines toward the N terminus (at residues 17 and 38) which form an intramolecular disulfide loop (24), an 11-residue polyglutamate region, one consensus site (NQS, residues 112–114) for N-linked glycosylation (25), three sets of oligoglutamic acid clusters, a 52-residue pancreastatin (26) homology (residues 259–311) with nearby single basic residues, and 8 sets of paired basic residues, likely recognition sites for proteolytic cleavage (27).

Structure of the Mouse Chromogranin A Gene—The chromogranin A gene was restriction mapped onto 2 overlapping mouse cosmids clones spanning about 52 kb (Fig. 4). This complex gene is composed of 8 exons separated by 7 introns, and the coding region spans about 11 kb. Sequence analysis of the exon/intron borders (Fig. 5) revealed good agreement of the borders with customary splice donor/acceptor (GT donor/AG acceptor) consensus sequences (28).

Chromogranin A's exons (Fig. 3) were at least partially congruent with putative functional domains in the primary structure: exon 1 (amino acid residues +18 to +3) encoded 16 of the 18 signal peptide amino acids; exon 2 (residues +3 to +13) encoded the N terminus of the mature protein; exon 3 (residues +14 to +45) spanned the relatively hydrophobic (6) intramolecular disulfide loop; exon 4 (residues +45 to +68) spanned a region without clear distinguishing features; exon 5 (residues +68 to +121) encoded the polyglutamine region as well as the NQS glycosylation site; exon 6 (residues +121 to +259) encoded an oligoglutamic acid cluster and two di- or tribasic sites; exon 7 (the longest coding exon, residues +259 to +418) began at the N terminus of pancreastatin and included not only pancreastatin but also five dibasic sites and two oligoglutamic acid clusters; and exon 8 (residues +419 to +445) encoded the C terminus with one dibasic site.

Analysis of homologies (Fig. 6) between the exons of mouse chromogranin A and mouse chromogranin B (29) reveals intriguing sequence similarities in three exons. Chromogranin A and B exons 2 each encode 16-residue domains at the N-termini of the mature proteins which share the sequence PV at residues +1 and +2. Exons 3 each encode 32-residue domains sharing 41% (13/32 residues) sequence identity, including the 2 cysteine residues that form the intramolecular disulfide loop (24). Chromogranin A exon 8 and chromogranin B exon 5 encode 25–27-residue C-terminal domains with 44% (11/25 residues) sequence identity.

One site of polymorphism differentiated the mouse chromogranin A coding (exon) sequences in the cDNA versus the genomic DNA; while the (CAG), trinucleotide repeat region in the cDNA encoded an 11-residue polyglutamine region (Figs. 2 and 3), the same (CAG), repeat region in the genomic clone encoded a 18-residue polyglutamine region.

Mouse Chromogranin A Promoter—Sequencing the genomic DNA just upstream of exon 1 (Fig. 7) revealed consensus sequences for a TATA box (TATATAA) promoter element (28) at positions −16 to −22, a cyclic AMP-response element (TGACCTAA; 30) at positions −64 to −71; and two
Fig. 2. Nucleotide sequence of the mouse brain chromogranin A cDNA with its deduced primary structure. Nucleotide numbers are in the left column, while amino acid residue numbers are in the right column. The amino acids in the 18-amino acid signal peptide are underlined and numbered negatively. The dash after amino acid residue 445 signifies the position of the TGA termination codon.

[Diagram showing the nucleotide sequence and deduced amino acid sequence of the mouse brain chromogranin A cDNA]

Fig. 3. Putative structural domains in mouse chromogranin A, as well as their correspondence to the eight exons in the mouse chromogranin A gene. The arrows at exons 1 and 8 indicate that these exons extend into noncoding (5' and 3') regions.

Sp1 sites (GGGCCG; 31) at positions −77 to −82 and −112 to −117. A CAAT box homology was not found (28). No other perfect (100%) consensus matches for other known transcriptional control regions were apparent up to nucleotide position −230. The mouse promoters for chromogranins A (Fig. 7) and B (29) shared a TATA box, a cyclic AMP-response element, and an Sp1 site, but were otherwise distinct.

Primer extension of a neuroendocrine (AtT-20 cell) mRNA displayed promoter activity (i.e. programmed luciferase expression) after transfection into neuroendocrine cells such as PC-12 adrenal chromaffin cells and AtT-20 pituitary corticotropes. The chromogranin A promoter was approximately as active as the SV-40 early promoter in PC-12 adrenal chromaffin cells, and approximately 18-fold more active than the SV-40 early promoter in AtT-20 pituitary corticotropes. In nonendocrine cell lines (NIH-3T3, COS), by contrast, the chromogranin A promoter was relatively inactive.

The mouse brain chromogranin A gene was successfully sequenced using the entire coding region of chromogranin A, including the 5' and 3' untranslated regions. The gene was shown to contain eight exons and seven introns, spanning a chromosomal region of about 12 kb. The promoter activity of the chromogranin A gene was tested in various cell lines, including neuroendocrine cells such as PC-12 adrenal chromaffin cells and AtT-20 pituitary corticotropes. The promoter activity of the chromogranin A gene was compared to the SV-40 early promoter and was found to be approximately as active as the SV-40 early promoter in PC-12 adrenal chromaffin cells, and approximately 18-fold more active than the SV-40 early promoter in AtT-20 pituitary corticotropes. In nonendocrine cell lines (NIH-3T3, COS), by contrast, the chromogranin A promoter was relatively inactive.
motorless luciferase negative control plasmid (pSV0ALΔ5) was virtually without expression in these studies.

**DISCUSSION**

Mouse chromogranin A is similar to rat chromogranin A (6) in primary structure (Fig. 6) as well as in the conservation of several structural domains (Fig. 3). However, some differences between mouse and rat chromogranin A were noted: (a) the mature mouse chromogranin A was 1 amino acid residue longer (445 versus 444 residues), (b) mouse chromogranin A, unlike rat chromogranin A (6), contained no RGD integrin sequence (32), (c) although both rat (6, 7, 23) and mouse (Figs. 2 and 3) chromogranin A contain polymorphic polyglutamine domains encoded by (CAG), trinucleotide repeats, the polyglutamine region in the rat (16–20 residues) differs in length from that in the mouse (11–18 residues), and (d) mouse chromogranin A contains 8 dibasic cleavage sites (Figs. 2 and 3), versus 10 dibasic sites in rat chromogranin A (6, 7, 23). Chromogranins A generally display considerable interspecies sequence homology, especially at the termini (6, 23). The pancreastatin domain in mouse chromogranin A has nearby single basic residues, rather than dibasic or tribasic clusters (Fig. 3). However, prohormone processing proteases may selectively recognize certain basic residues in chromogranin A (33).

Chromogranin A’s eight exons showed substantial correspondence to putative functional domains deduced from the cDNA sequence (Fig. 3). In addition, three of chromogranin A’s eight exons (exons 2, 3, and 8) displayed both sequence and length homology to three of chromogranin B’s five exons (29; Fig. 6). Perhaps these exons were common to an evolutionary ancestor of chromogranins A and B, ultimately giving rise to separate chromogranins by gene duplication and subsequent mutation (34). Also of note, the C-terminal homologous chromogranin A exon 8 and chromogranin B exon 5 were each bounded or closely apposed by dibasic sites, suggesting that the ability of this domain to be liberated by proteolytic cleavage may be conserved across the chromogranins. This C-terminal exon displays additional partial sequence homology to other members of the chromogranin/secretogranin family, including secretogranin II (chromogranin C; 55, 36) and secretogranin III (brain clone 1B1075; 37).

The transient expression results (Table 1) suggest that information sufficient to confer both strong and tissue-specific expression of the chromogranin A gene is contained within the genomic DNA immediately (1.2 kb) 5′ of the chromogranin A coding region (Fig. 4). Indeed, this promoter/enhancer region was as potent as the SV-40 early promoter in chromaffin cells, and 18-fold more potent in anterior pituitary corticotropes, but was virtually inactive in nonendocrine cells. Initial sequence analysis of this promoter revealed only a few exact (chromogranin homologies to known consensus promoter/enhancer elements, such as the TATA box (28), the cyclic AMP-response element (30), and Sp1-binding site (31). Thus, there are no doubt novel transcriptional control elements in cis within this promoter, which may be of funda-
mental importance in determining the acquisition of the neuroendocrine phenotype, or patterns of gene expression which unite peptide and neurotransmitter storing/releasing cells.

Acknowledgments—The mouse brain cDNA library and the mouse genomic DNA library were gifts from Dr. Stephen Heineman and Dr. Glen Evans (Salk Institute, San Diego, CA).

REFERENCES

1. Takiyuddin, M. A., Cervenka, J. H., Hsiao, R. J., Barbosa, J. A., Parmer, R. J., and O'Connor, D. T. (1990) Hypertension 15, 237–246
2. Eiden, L. E., Huttner, W. B., Mallet, J., O'Connor, D. T., Winkler, H., and Zanini, A. (1987) Neuroscience 21, 1019–1021
3. O'Connor, D. T. (1983) Regul. Pept. 6, 283–280
4. Cohn, D. V., Elting, J. J., Frick, M., and Elde, R. (1984) Endocrinology 114, 1963–1974
5. Simon, J. P., and Annis, D. (1989) Biochem. J. 262, 1–13
6. Parmer, R. J., Koop, A. H., Handa, M. T., and O'Connor, D. T. (1989) Hypertension 14, 436–444
7. Huttner, J. C., Nielsen, E., and Kasten, W. (1988) FEBS Lett. 236, 269–274
8. Konecki, D. S., Benedum, U. M., Gerdes, H. H., and Huttner, W. B. (1987) J. Biol. Chem. 262, 17026–17030
9. Murray, S., Deaven, L. W., Burton, D. G., O'Connor, D. T., Mellon, P. L., and Defos, L. J. (1987) Biochem. Biophys. Res. Commun. 142, 141–146
10. Modi, W. S., Levine, M. A., Seuanez, H. N., Dean, M., and O'Brien, S. J. (1989) Am. J. Hum. Genet. 45, 814–818
11. Rigby, P. W. J., Dieckman, W., Rhodes, C., and Berg, P. (1977) J. Mol. Biol. 113, 237–251
12. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
13. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed, Cold Spring Harbor, NY
14. Evans, G. A., Lewis, K., and Rothenberg, B. E. (1989) Gene (Amst.) 79, 9–20
15. Evans, G. A., and Wahl, G. M. (1987) Methods Enzymol. 152, 604–610
16. Dickerson, I. M., and Mains, R. E. (1990) Endocrinology 127, 133–40
17. DeWet, J. R., Wood, K. V., DeLuca, M., Helinski, D. R., and Subramaniam, S. (1987) Mol. Cell. Biol. 7, 725–737
18. Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrup, J. P., Ringold, G. M., and Danielson, M. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 7413–7417
19. Greene, L. A., and Tischler, A. S. (1976) Proc. Natl. Acad. Sci. U. S. A. 73, 2424–2428
20. Gluzman, Y. (1981) Cell 23, 175–182
21. Gorman, C. M., Moffat, L. F., and Howard, B. H. (1982) Mol. Cell. Biol. 2, 1044–1051
22. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
23. Iacangelo, A., Okayama, H., and Huttner, W. B. (1988) FEBS Lett. 227, 115–121
24. Benedum, U. M., Lamouroux, A., Konecki, D. S., Rosa, P., Hille, A., Baeuerle, P. A., Frank, R., Lottspeich, F., Mallet, J., and Huttner, W. B. (1987) EMBO J. 6, 1203–1211
25. Hubbard, S. C., and Ivatt, R. J. (1981) Annu. Rev. Biochem. 50, 555–583
26. Tatetomo, K., Efendic, S., Mutt, V., Makk, G., Feistner, G. J., and Parmer, R. (1984) Mol. Endocrinology 128, 174–190
27. Barbosa, J. A., Gill, B. M., Takiyyuddin, M. A., and O'Connor, D. T. (1991) Endocrinology 128, 50–55
28. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1989) Molecular Biology of the Cell, Garland Publishing Co., New York
29. Pohl, T. M., Phillips, E., Song, K. Y., Gerdes, H. H., Huttner, W. B., and Ruther, U. (1990) FEBS Lett. 262, 219–224
30. Montminy, M. R., Sevarino, K. A., Wagner, J. A., Mandel, G., and Goodman, R. H. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 6682–6686
31. Briggs, M. R., Kadonaga, J. T., Bell, S. P., and Tjian, R. (1986) Science 234, 47–52
32. Ruoslahti, E., and Pierschbacher, M. D. (1986) Cell 44, 517–518
33. Leduc, R., Hendy, G. N., Seidah, N. G., Chretien, M., and Lazare, C. (1980) Life Sci. 25, 1427–1432
34. Doolittle, R. F. (1986) Of urs and orfs. A Primer on How to Analyze Derived Amino Acid Sequences, University Science Books, Mill Valley, CA
35. Gerdes, H. H., Phillips, E., and Huttner, W. B. (1988) Nucleic Acids Res. 16, 11811
36. Gerdes, H. H., Rosa, P., Phillips, E., Baeuerle, P. A., Frank, R., Argos, P., and Huttner, W. B. (1989) J. Biol. Chem. 264, 12009–12015
37. Ottiger, H. P., Battenberg, E. F., Tsou, A. P., Bloom, F. E., and Sutcliffe, J. G. (1990) J. Neurosci. 10, 3135–3147
38. Linard, C. G., Miklay, M., Seidah, N. G., and Chretien, M. (1990) Nucleic Acids Res. 18, 1298

| Cell line | Cell type | Mouse CgA | SV-40 early | None |
|-----------|-----------|-----------|-------------|------|
| PC-12     | Chromaffin| 97.44     | 100         | 0.45 |
| AtT-20    | Pituitary (ACTH) | 1840.92 | 100 | 0.31 |
| NIH 3T3   | Fibroblast | 0.41      | 100         | 0    |
| COS      | Monkey kidney | 0.11      | 100         | 0    |

Table I: Activity of the mouse chromogranin A (CgA) promoter/enhancer

Activity of the mouse chromogranin A promoter/enhancer (1.2 kb 5′-fragment in the construct pHJALΔ5) assessed by luciferase reporter assay after lipofection-mediated transfection into neuroendocrine cells (PC-12, AtT-20) or nonneuroendocrine cell lines (NIH-3T3, COS). Each transfection included a control plasmid encoding chloramphenicol acetyltransferase (pHSV-CAT), to correct for differences in transfection efficiency across cell lines. The negative control plasmid, pSV2ALA5, contained the luciferase coding region but no mammalian promoter. After 48 h, transfected cells were extracted and assayed for protein as well as the luciferase and chloramphenicol acetyltransferase reporters. The luciferase reporter results are normalized both to cell protein content and to cell chloramphenicol acetyltransferase content. For each cell type, results are normalized to the activity of the SV-40 early promoter (100%).