Cloning, mRNA Expression, and Chromosomal Mapping of Mouse and Human Preprocortistatin

Luis de Lecea,* Pilar Ruiz-Lozano,† Patria E. Danielson,* Jessica Peelle-Kirley,‡ Pamela E. Foye,* Wayne N. Frankel,‡ and J. Gregor Sutcliffe*1

*Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037; †Department of Medicine, University of California, San Diego, San Diego, California 92093; and ‡The Jackson Laboratories, Bar Harbor, Maine 04609

Received January 24, 1997; accepted April 4, 1997

Cortistatin is a 14-residue putative neuropeptide with strong structural similarity to somatostatin and is expressed predominantly in cortical GABAergic interneurons of rats. Administration of cortistatin into the brain ventricles specifically enhances slow-wave sleep, presumably by antagonizing the effects of acetylcholine on cortical excitability. Here we report the identification of cDNAs corresponding to mouse and human preprocortistatin and the mRNA distribution and gene mapping of mouse cortistatin. Analysis of the nucleotide and predicted amino acid sequences from rat and mouse reveals that the 14 C-terminal residues of preprocortistatin, which make up the sequence that is most similar to somatostatin, are conserved between species. Lack of conservation of other dibasic amino acid residues whose cleavage by prohormone convertases would give rise to additional peptides suggests that cortistatin-14 is the only active peptide derived from the precursor. As in the rat, mouse preprocortistatin mRNA is present in GABAergic interneurons in the cerebral cortex and hippocampus. The preprocortistatin gene maps to mouse chromosome 4, in a region showing conserved synteny with human 1p36. The human putative cortistatin peptide has an arginine for lysine substitution, compared to the rat and mouse products, and is N-terminally extended by 3 amino acids. © 1997 Academic Press

INTRODUCTION

We recently isolated a cDNA clone of the mRNA encoding rat preprocortistatin, a 112-residue protein, the amino acid sequence of which suggests that it is the putative precursor of a novel secreted neuropeptide (de Lecea et al., 1996). Maturation of the rat preprosequences to proprocortistatin would produce a protein that could be processed at a Lys-Arg site to generate a 29-residue peptide (rCST29), at a Lys-Lys site to give rise to a 14-amino-acid peptide (rCST14), also called cortistatin, or at both sites to generate both CST14 and a 13-residue peptide (Fig. 1). rCST14 shares 11 of its 14 residues with somatostatin, including those that are known to be responsible for somatostatin binding to its receptors (Veber et al., 1979) and the cysteines that are likely to render the peptide cyclic. The 13-residue species is unrelated to known proteins. Preprocortistatin mRNA is expressed in a distinct subset of interneurons in the rat cerebral cortex and hippocampus, areas of the brain thought to be important for high cognitive functions (de Lecea et al., 1996). The cDNA sequences of preprocortistatin and preprosomatostatin indicate clearly that they are the products of separate genes.

Synthetic rCST14 was shown to share several biological properties with somatostatin: it bound to somatostatin receptors on GH4 pituitary cells, inhibited the VIP- and TRH-induced accumulation of cAMP in those cells, and depressed neuronal activity in hippocampal neurons, probably by enhancing the potassium M current (de Lecea et al., 1996). However, the effects of cortistatin on cortical electrical activity and sleep were distinct from those found for somatostatin. Intracerebroventricular administration of synthetic rCST14 specifically enhanced the amount of time that the animals spent in slow-wave sleep but did not affect significantly their paradoxical (REM) sleep. Moreover, rCST14 was shown to antagonize the effects of acetylcholine on cortical measures of excitability (de Lecea et al., 1996), whereas somatostatin is known to enhance acetylcholine release and to potentiate acetylcholine responses (Mancillas et al., 1986). These observations demonstrated that cortistatin is functionally distinct from somatostatin and raised the possibility that cortistatin exerts its activities through an uncharacterized cortistatin-selective receptor, although other explanations of different functionalities can be considered.

To gain information on the conservation of the putative processed neuropeptides, we have isolated cDNA clones encoding mouse and human preprocortistatin. We demonstrate that the amino acid sequence of the active cortistatin-14 peptide is fully...
grams of poly(A) + RNA from rat, mouse, and human (Clontech) samples was run on formaldehyde agarose gels and transferred to nylon filters as described (Danielson et al., 1994). Mouse or human cDNA probes were labeled with 32P and random primers.

In situ hybridization. C57BL/6J mice were perfused with 4% paraformaldehyde and processed for in situ hybridization as described (de Lecea et al., 1994). Free-floating sections were hybridized with 10 6 cpm/ml labeled cortistatin probe and washed at 60°C in 0.5× SSC/50% formamide for 3 h. After being mounted, slides were dipped in Ilford K5 emulsion diluted in water and exposed for 3 weeks at 4°C. Slides were developed in Kodak D19, counterstained, and mounted in Permount.

Chromosomal mapping. The oligonucleotides for mapping Cort by SSCP were 5'-AAAAAGCCCTGCAAGAACTT-3' (Cort1F) and 5'-ATTCAGGTCTCGTTGGCATC-3' (Cort1R). The PCR conditions have been described previously (Dietrich et al., 1992) except that [α-32P]dCTP was incorporated into the reaction. The PCR product was denatured, then quick cooled on ice and electrophoresed for 4 h in a 0.5% MDE gel (AT Biochem, Inc.) at 4°C. The gel was exposed to X-ray film overnight. Linkage data were analyzed using the latest version of the computer program MapManager (Manly and Elliott, 1991), which can be obtained on the web (http://mcbio.med.buffalo.edu/mapmgr.html). The sequence of the Cort PCR product was determined by the dideoxy method using Cort1F and Cort1R primers.

RESULTS

Analysis of Mouse Preprocortistatin DNA Sequence

We previously reported the isolation of a rat cDNA clone, the nucleotide sequence of which suggested that it encoded the precursor of cortistatin, a peptide with sequence similarity to somatostatin. Preprocortistatin begins with a 27-residue apparent secretion signal sequence. Interestingly, this region contains six iterations of the trinucleotide CTG, whose triplet expansion in other genes has been implicated as causal in neurological diseases (e.g., myotonic dystrophy) (Brook et al., 1992). The rat preprocortistatin deduced amino acid sequence contains several pairs of basic residues that are possible substrates of prohormone convertases. Cleavage at all basic amino acid pairs would give rise to rCST17 (with a putative amidation site), rCST31, rCST29, rCST13, and rCST14 (Fig. 1). Alternative or partial cleavage could produce additional peptide products. rCST14 may be further processed by carboxypeptidases that would remove its C-terminal lysine.

We used the full-length rat cDNA clone to screen a mouse cerebral cortex cDNA library (generously provided by Dr. K. Hasel). Several positive clones were isolated, and their nucleotide sequences were determined. Two cDNA clones, 430 bp long, appeared to be full-length as judged by the alignment of their 5′ ends with the rat sequence (not shown). After introduction of two gaps, the mouse and rat nucleotide sequences were 86% identical (Fig. 2A). Assuming that the putative mouse translation initiation product begins at the second methionine triplet, it contains 108 amino acids compared to 112 for rat. Again, after introduction of two gaps, the putative mouse and rat proteins share 82% identity (Fig. 2B). The mouse nucleotide sequence corresponding to rCST14 and the adjacent lysine dou-
### MOUSE AND HUMAN PREPROCORTISTATIN cDNAs

#### A

<table>
<thead>
<tr>
<th>Species</th>
<th>cDNA Sequence</th>
<th>Consensus</th>
</tr>
</thead>
</table>
| Mouse Cstr | gcacagcg acctgactc gcccggtaa ggtggtc | **P**---**L** | 1
| Rat Cstr | aaagcagagt tcaagtgctt ccaagagatg gcccggtaa ggtggtc | **P**---**L** | 50
| Human Cstr | gacagccagcg tgggtaa ggtggtc | **P**---**L** | 50

#### B

<table>
<thead>
<tr>
<th>Species</th>
<th>cDNA Sequence</th>
<th>Consensus</th>
</tr>
</thead>
</table>
| Mouse Cstr | gaggagttgg ctcttgtcag gctgtaaa cctgagtc | ------**P** | 50
| Rat Cstr | gaggagttgg ctcttgtcag gctgtaaa cctgagtc | ------**P** | 50
| Human Cstr | gaggagttgg ctcttgtcag gctgtaaa cctgagtc | ------**P** | 50
much lower degree of identity to the rat sequence (71%). The human preprocortistatin deduced amino acid sequence (Fig. 2B) has 114 residues and begins with a 29-amino-acid hydrophobic probable secretory signal sequence. The sequence corresponding to the putative signal peptide of preprocortistatin contains only four iterations of CTG encoding the amino acid leucine, in contrast to six iterations of the same triplet in the rat peptide precursor or three in mouse, suggesting that this sequence is unstable and subject to expansion. Analysis of the putative processing sites in human preprocortistatin revealed that it may be cleaved at two RR sites, giving rise to hCST29 and a C-terminal 17-residue peptide that shared 13 of the last 14 residues with rat and mouse CST14, here called hCST17. The Lys-Lys pair that lies just N-terminal to cortistatin-14 in rat and mouse is not conserved in the human sequence. The other possible products that follow the signal sequence (hCST21 and hCST31) are not very conserved across species, although rCST31 and hCST31 share 13 residues clustered in their N-terminal regions that are conserved among the rat, mouse, and human prohormone sequences (Fig. 2B).

mRNA Expression

We determined the distribution of preprocortistatin mRNA by Northern blot. A band of approximately 600 nucleotides was detected in samples prepared from rat brain, cortex, and hippocampus, but not pancreas or gut (Fig. 3) or adrenal gland, liver, spleen, thymus, ovary, testes, or anterior pituitary (not shown). This pattern of expression is clearly distinct from somatostatin mRNA, which is present in several endocrine tissues. Hybridization to Northern blots of mouse tissues revealed the presence of two bands in brain but not liver, kidney, or thymus. Two bands were also observed in the human brain sample. These bands are probably due to alternate polyadenylation signals, found to be present in mouse genomic clones (L.dL., unpublished observations) and in human cDNA clones.

We previously reported that rat cortistatin is expressed in a subset of cortical and hippocampal GABAergic interneurons. To determine whether the expression of cortistatin was conserved between species, we performed situ hybridization with mouse brain tissue (Fig. 4). As in the rat, cortistatin-positive neurons were enriched in the cerebral cortex and hippocampus. In the temporal/visual cortex, cortistatin-positive neurons were especially abundant in layer VI, with very few scattered cells present in layer II–III (Figs. 4A and 4B). In the hippocampus, preprocortistatin mRNA

**FIG. 2.** (A) Alignment of the nucleotide sequences of rat, mouse, and human preprocortistatin cDNAs. The human sequence is a composite from different PCR fragments and cDNA clones, including one that showed a deletion in the coding sequence and an additional 3′ polyadenylation signal. The CTG repeat that encodes the amino acid leucine and that is of variable length between species has been underlined. The two possible polyadenylation signals are marked with an asterisk. Nucleotides conserved among all three species are shown in uppercase, otherwise, in lowercase. (B) Alignment of the deduced amino acid sequences of the rat, mouse, and human cortistatin precursors. The putative dibasic cleavage sites are indicated in bold. Consensus residues are indicated. (C) Comparison of the amino acid sequences and predicted secondary structures of rat, mouse, and human cortistatin and somatostatin from frog and mammals.
FIG. 3. Northern blots of RNA samples from rat whole brain, cortex, hippocampus, gut, and pancreas and mouse brain, liver, thymus, and kidney. The blots were hybridized with the rat cortistatin cDNA and with a cyclophilin probe (Danielson et al., 1988; not shown) as a control for loading and RNA integrity. A separate Northern blot containing an mRNA sample from whole human brain was hybridized with the human preprocortistatin cDNA sequence. In short exposures both the mouse and the human samples displayed two bands, probably generated by alternative polyadenylation signals.

could be visualized in the stratum oriens of the CA1-CA3 fields, as well as in a few neurons adjacent to the granule cell layer of the dentate gyrus. The hilar region was totally devoid of preprocortistatin-expressing cells. Strong signals could also be detected in the amygdala, especially in the medial amygdaloid nucleus (Fig. 4C). In the hypothalamus, preprocortistatin mRNA was detected in a few cells in the periventricular nucleus.

Chromosomal Mapping of Mouse Cortistatin

We mapped the cortistatin gene (gene symbol Cort) by single-strand conformation polymorphism analysis of a panel of interspecific backcross mouse DNAs. We designed primers that spanned the 3’ coding/3’ untranslated sequence of mouse cortistatin cDNA and amplified the corresponding 107-bp genomic fragment from C57BL/6J (B6) and a strain inbred from wild-derived Mus spretus (SPRET/Ei). Representative PCR fragments were sequenced to confirm their identity. A clear polymorphism that distinguished the two strains was found. To determine linkage the segregation pattern of the B6 allele was followed in a subpanel of 54 (B6 × SPRET) F1 × SPRET backcross offspring and compared to that of over 2500 genes previously mapped on the panel. The mouse cortistatin locus was found to lie on distal chromosome 4—in contrast with the somatostatin gene which maps to chromosome 16 (Lalley et al., 1987)—and was nonrecombinant with the Mtfr locus (LOD 16.3; Fig. 5). Neurological mutations that are known to reside in this region include Wallerian degeneration (Wld) and jerker (je). A quantitative trait locus for β-carboline-induced seizures has also been mapped in this region (Martin et al., 1995). This telomeric region of mouse Chr 4 show strong conserved synteny with human chromosome 1p36 (O’Brien et al., 1993), but we have not identified human neurological disorders mapping to this region for which Cort would be a compelling candidate.

DISCUSSION

We have described the cloning of the mouse and human homologues of the neuropeptide cortistatin mRNAs and mRNA distribution and gene mapping of mouse preprocortistatin. Analysis of the nucleotide and predicted amino acid sequences from rat and mouse reveals that the 14 C-terminal residues of preprocortistatin, which constitute the sequence that is most similar to somatostatin, are conserved between these species, whereas the mono- or dibasic amino acid residues whose cleavage by prohormone convertases would give rise to additional peptides are not conserved. This strongly suggests that CST14 is the only active peptide from C57BL/6J (B6) and a strain inbred from wild-derived Mus spretus (SPRET/Ei). Representative PCR fragments were sequenced to confirm their identity. A clear polymorphism that distinguished the two strains was found. To determine linkage the segregation pattern of the B6 allele was followed in a subpanel of 54 (B6 × SPRET) F1 × SPRET backcross offspring and compared to that of over 2500 genes previously mapped on the panel. The mouse cortistatin locus was found to lie on distal chromosome 4—in contrast with the somatostatin gene which maps to chromosome 16 (Lalley et al., 1987)—and was nonrecombinant with the Mtfr locus (LOD 16.3; Fig. 5). Neurological mutations that are known to reside in this region include Wallerian degeneration (Wld) and jerker (je). A quantitative trait locus for β-carboline-induced seizures has also been mapped in this region (Martin et al., 1995). This telomeric region of mouse Chr 4 show strong conserved synteny with human chromosome 1p36 (O’Brien et al., 1993), but we have not identified human neurological disorders mapping to this region for which Cort would be a compelling candidate.

DISCUSSION

We have described the cloning of the mouse and human homologues of the neuropeptide cortistatin mRNAs and mRNA distribution and gene mapping of mouse preprocortistatin. Analysis of the nucleotide and predicted amino acid sequences from rat and mouse reveals that the 14 C-terminal residues of preprocortistatin, which constitute the sequence that is most similar to somatostatin, are conserved between these species, whereas the mono- or dibasic amino acid residues whose cleavage by prohormone convertases would give rise to additional peptides are not conserved. This strongly suggests that CST14 is the only active peptide from C57BL/6J (B6) and a strain inbred from wild-derived Mus spretus (SPRET/Ei). Representative PCR fragments were sequenced to confirm their identity. A clear polymorphism that distinguished the two strains was found. To determine linkage the segregation pattern of the B6 allele was followed in a subpanel of 54 (B6 × SPRET) F1 × SPRET backcross offspring and compared to that of over 2500 genes previously mapped on the panel. The mouse cortistatin locus was found to lie on distal chromosome 4—in contrast with the somatostatin gene which maps to chromosome 16 (Lalley et al., 1987)—and was nonrecombinant with the Mtfr locus (LOD 16.3; Fig. 5). Neurological mutations that are known to reside in this region include Wallerian degeneration (Wld) and jerker (je). A quantitative trait locus for β-carboline-induced seizures has also been mapped in this region (Martin et al., 1995). This telomeric region of mouse Chr 4 show strong conserved synteny with human chromosome 1p36 (O’Brien et al., 1993), but we have not identified human neurological disorders mapping to this region for which Cort would be a compelling candidate.

DISCUSSION

We have described the cloning of the mouse and human homologues of the neuropeptide cortistatin mRNAs and mRNA distribution and gene mapping of mouse preprocortistatin. Analysis of the nucleotide and predicted amino acid sequences from rat and mouse reveals that the 14 C-terminal residues of preprocortistatin, which constitute the sequence that is most similar to somatostatin, are conserved between these species, whereas the mono- or dibasic amino acid residues whose cleavage by prohormone convertases would give rise to additional peptides are not conserved. This strongly suggests that CST14 is the only active peptide from C57BL/6J (B6) and a strain inbred from wild-derived Mus spretus (SPRET/Ei). Representative PCR fragments were sequenced to confirm their identity. A clear polymorphism that distinguished the two strains was found. To determine linkage the segregation pattern of the B6 allele was followed in a subpanel of 54 (B6 × SPRET) F1 × SPRET backcross offspring and compared to that of over 2500 genes previously mapped on the panel. The mouse cortistatin locus was found to lie on distal chromosome 4—in contrast with the somatostatin gene which maps to chromosome 16 (Lalley et al., 1987)—and was nonrecombinant with the Mtfr locus (LOD 16.3; Fig. 5). Neurological mutations that are known to reside in this region include Wallerian degeneration (Wld) and jerker (je). A quantitative trait locus for β-carboline-induced seizures has also been mapped in this region (Martin et al., 1995). This telomeric region of mouse Chr 4 show strong conserved synteny with human chromosome 1p36 (O’Brien et al., 1993), but we have not identified human neurological disorders mapping to this region for which Cort would be a compelling candidate.

DISCUSSION

We have described the cloning of the mouse and human homologues of the neuropeptide cortistatin mRNAs and mRNA distribution and gene mapping of mouse preprocortistatin. Analysis of the nucleotide and predicted amino acid sequences from rat and mouse reveals that the 14 C-terminal residues of preprocortistatin, which constitute the sequence that is most similar to somatostatin, are conserved between these species, whereas the mono- or dibasic amino acid residues whose cleavage by prohormone convertases would give rise to additional peptides are not conserved. This strongly suggests that CST14 is the only active peptide from C57BL/6J (B6) and a strain inbred from wild-derived Mus spretus (SPRET/Ei). Representative PCR fragments were sequenced to confirm their identity. A clear polymorphism that distinguished the two strains was found. To determine linkage the segregation pattern of the B6 allele was followed in a subpanel of 54 (B6 × SPRET) F1 × SPRET backcross offspring and compared to that of over 2500 genes previously mapped on the panel. The mouse cortistatin locus was found to lie on distal chromosome 4—in contrast with the somatostatin gene which maps to chromosome 16 (Lalley et al., 1987)—and was nonrecombinant with the Mtfr locus (LOD 16.3; Fig. 5). Neurological mutations that are known to reside in this region include Wallerian degeneration (Wld) and jerker (je). A quantitative trait locus for β-carboline-induced seizures has also been mapped in this region (Martin et al., 1995). This telomeric region of mouse Chr 4 show strong conserved synteny with human chromosome 1p36 (O’Brien et al., 1993), but we have not identified human neurological disorders mapping to this region for which Cort would be a compelling candidate.
FIG. 4. In situ hybridization in mouse brain. (A) Dark-field micrograph of a section through the mouse cortex. Note the presence of scattered cells in the deep layers of the neocortex and hippocampal CA1 field (arrows). (B) High-magnification of a cortistatin-positive cell in layer VI. (C) Dark-field image of the mouse amygdala hybridized with a cortistatin riboprobe. The amygdala and several regions of the hypothalamus (not shown) showed stronger signals in the mouse compared to the rat.
MOUSE AND HUMAN PREPROCORTISTATIN cDNAs

Cortistatin-14 sequence with one conservative substitution. This suggests that the critical amino acids for cortistatin function reside in the loop formed by the two cysteines and, possibly, in the N-terminal proline and C-terminal lysine, although the latter may be processed by carboxypeptidases in the secretory pathway (Fricker, 1988). We cannot rule out the possibility that the human CST species is further processed at a single R site, to generate CST14 with an additional N-terminal methionine substitution.

Recently, a second vertebrate somatostatin gene has been reported in the frog Rana ridibunda (Tostivint et al., 1996). Frog somatostatin II has two amino acid substitutions relative to somatostatin I: a Pro at position 2 and a Met at position 13 (Fig. 2C). Thus, the N-terminal proline residue may be critical for the specific actions of somatostatin II in frog and cortistatin in rat, mouse, and human. However, somatostatin II is an unlikely predecessor of cortistatin, as the nucleotide and amino acid sequences of the precursors are quite divergent. As during the evolution of tetrapods several gene duplications may have occurred, the existence of more members of the somatostatin/cortistatin family in mammals cannot be ruled out.

Analysis of mouse preprocortistatin mRNA expression showed an overall coincidence with the pattern described in rat. However, mouse preprocortistatin mRNA seems less abundant than its rat counterpart, as judged by Northern blot and in situ hybridization. In the mouse visual cortex, cortistatin-positive cells were abundant only in the deep layers, whereas in rat, cortistatin signals covered the entire thickness of the cortex. Also, we could detect some cortistatin-positive cells in the mouse periventricular hypothalamic area and in the amygdala, regions that were negative in the rat. Small differences in the expression of neuropeptides between species have been reported for galanin (Benzing et al., 1993) and other neuropeptides (Blahser, 1992), although the functional implications are unknown.

The rat DNA sequence for preprocortistatin contains six repetitions of the trinucleotide CTG in the region corresponding to its putative signal peptide, whereas the mouse sequence contains three and the human displays four iterations of the same triplet. The instability of CTG repeats has been shown to be responsible for several neurological diseases in humans as well as in mouse models. Expansion of the CTG repeat of cortistatin would likely impair its processing into a mature, active peptide. Alternatively, an expanded polyleucine stretch could produce gain-of-function mutations.

ACKNOWLEDGMENTS

We thank The Jackson Laboratory mapping service for providing backcross DNAs. This work was supported by grants from NS33396, GM32355, and Digital Gene Technologies. W.N.F. holds a Klingenstein Fellowship in the Neurosciences.
REFERENCES


