Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 10059–10063, November 1991
Biochemistry

Primary structure and functional expression of a glutaminyl cyclase

(enzyme purification/peptide hormone biosynthesis/posttranslational processing/processing enzyme/pyroglutamyl peptides)

THOMAS POHL, MICHAEL ZIMMER, KATRIN MUGELE, AND JOACHIM SPIESS

Department of Molecular Neuroendocrinology, Max Planck Institute for Experimental Medicine, Hermann-Rein-Str. 3, 3400 Göttingen, Federal Republic of Germany

Communicated by Hans Neurath, July 25, 1991

ABSTRACT **A glutaminyl cyclase (QC) that is probably involved in the biosynthesis of pyroglutamyl peptides such as gonadotropin-releasing hormone and thyrotropin-releasing hormone has been purified to homogeneity from bovine anterior pituitary. On the basis of N-terminal sequence analysis, a 2088-base-pair cDNA clone was isolated from a bovine anterior pituitary library. From the nucleotide sequence of this clone, the primary structure of a 330-residue protein and a preceding 31-residue prepropeptide sequence was deduced. By transfection of COS-7 monkey cells with a QC cDNA/pCDM8 vector construct, QC activity was expressed. Hybridization with mRNAs of various bovine tissues revealed expression of QC mainly in brain tissue.**

For the pyroglutamyl peptides thyrotropin-releasing hormone (TRH) and gonadotropin-releasing hormone (GnRH) the N-terminal pyroglutamyl modification is required for biologic activity (1, 2). The details of the biosynthesis of these hormones have not been elucidated. In the primary structures of the precursors to TRH (3, 4) and GnRH (5), sequences corresponding to glutaminyl analogs of TRH and GnRH (C-terminally deamidated and extended by glycine) are flanked by pairs of basic amino acids. On the basis of these structures it is probable that glutaminyl peptides are intermediates in the propeptide peptide conversion. Although the cyclization of N-terminal glutamine residues can occur under nonenzymatic conditions, especially under the catalytic influence of phosphate ions (6), a previous study has demonstrated (7) that tissue extract loses almost all cyclization activity after short exposure to heat. Therefore (and on the basis of other data) we have concluded that glutaminyl peptides are converted to pyroglutamyl peptides in an enzymatic reaction by glutaminyl cyclase (QC) (7). QCs have been identified in plants (8) and in mammalian tissues (7, 9) such as pituitary, hypothalamus, other parts of the brain, adrenal medulla, and B lymphocytes. Investigation of the biologic function of these enzymes has been hampered by the lack of structural data and specific antibodies, mainly because at least the mammalian enzymes are rather labile and present only in very low concentrations in the tissues tested. We report here the purification to homogeneity and DNA sequence determination* of a mammalian QC as well as confirmation of the identity of the cloned sequence by expression in COS-7 cells. On the basis of these findings it will now be possible to investigate in transfection experiments the biologic significance of this enzyme for posttranslational processing, the enzyme's maturation along the secretory pathway, its transcriptional regulation, and its mechanism of action.

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.

MATERIALS AND METHODS

Tissue Extracts. From the local slaughterhouse, 1085 pituitaries were freshly obtained and dissected. The anterior lobes (wet weight, 1979 g) were homogenized in batches of 50–100 glands with a Polytron homogenizer (Brinkmann), isotonically extracted, and centrifuged at 600 and 45,000 × *g* as described (7). The bottom part (P2b) of the precipitate of the centrifugation at 45,000 × *g* represented the crude secretory granule fraction. It was hypotonically extracted (7) to yield the supernatant Sn P2b.

QC Assay. Peptides were synthesized on a peptide synthesizer (model 430A, Applied Biosystems) by the *t*-butoxycarbonyl (t-Boc) protocol (version 1.3B). Purity of the peptides was assessed by analytical reversed-phase high-performance liquid chromatography (HPLC), plasma desorption mass spectrometry (10), and automated Edman degradation. The conversion was determined with reversed-phase HPLC on a Hewlett–Packard 1090 chromatograph and Hewlett–Packard C_{18} columns (2.1 × 100 mm; particle size, 5 μm; pore size, 120 Å). Samples were eluted at 40°C and a flow rate of 1 ml/min with a mixture of 0.1% aqueous trifluoroacetic acid (solvent A) and 0.1% trifluoroacetic acid in acetonitrile (solvent B). Usually, a linear gradient of 0–12.5% B was employed within 3 min. The activity was expressed in enzymatic units (1 unit corresponds to the formation of 10 nmol of pyroglutamyl peptide per hour). The specific activity of QC preparations was determined under saturating-substrate conditions with 400–600 nmol of H-Gln-Tyr-Ala-OH (final concentration, 0.8–1.2 mM) and an incubation time of 10–20 min; 50 μl of 9% HCl was added and an aliquot of 55 μl was applied to HPLC. Significantly higher substrate concentrations than the ones suggested on the basis of the K_m were required for activity assays of crude extracts to compensate for substrate consumption by peptide-degrading enzymes such as peptidases (11). In a typical assay of partially purified QC, 10 nmol of H-Gln-Tyr-Ala-OH (3.8 μg) was incubated at 37°C for 1 hr with either incubation buffer (0.1 M Tes buffer as sodium salt, pH 8) or hypophyseal extract (Sn P2b) containing ≈4 units of QC in a volume of 500 μl. The reaction was stopped with 50 μl of 9% HCl. The application volume was 220 μl (40% of the total reaction mixture). Protein was quantified by the method of Bradford (Bio-Rad kit).

Chromatographic Procedures. Extracted QC was subsequently purified by "fast protein liquid chromatography" (FPLC, Pharmacia). The following procedures were consecutively applied at 4°C. Anion-exchange FPLC was performed at a flow rate of 2 ml/min on a Mono Q 16/10 column (Pharmacia LKB). After equilibration with 20 mM Tris·HCl at pH 8.0 (buffer A), ≈50 mg of protein was applied per run

Abbreviations: QC, glutaminyl cyclase; IEF, isoelectric focusing; PAM, peptidylglycine α-amidating monooxygenase.

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M80626).

and eluted with a mixture of buffer A and 1 M NaCl in buffer A (buffer B).

Reversed-phase FPLC was performed on a ProRPC HR 16/10 column (Pharmacia LKB) at a flow rate of 2 ml/min. After equilibration with 60 mM ammonium formate at pH 6.5 (solvent A), ≈20 mg of protein was applied per run and eluted with a mixture of solvent A and acetonitrile (solvent B). Gel filtration FPLC was performed on Superose 12 10/30 column (Pharmacia LKB) equilibrated and eluted with 50 mM Tes·NaOH, pH 8.0/500 mM NaCl/10 mM methylamine/2 mM dithiothreitol/5% (vol/vol) glycerol at a flow rate of 0.3 ml/min. Before application, samples were desalted and concentrated with a Centriprep 10 device (Amicon). Aliquots of 300–400 μl (1 mg of protein per run) were injected.

Polyacrylamide Gel Electrophoresis. QC samples were prepared for isoelectric focusing (IEF) by three consecutive steps of FPLC. Pooled QC fractions were concentrated and desalted either in a Centricon 3 device precoated with 250 μg of cytochrome *c* or by a second reversed-phase HPLC. The final sample of 80–150 μl, containing 1–5 μg of QC, was supplemented with 4% (wt/vol) Pharmalytes 4–6.5 (Pharmacia) and 10% (vol/vol) glycerol and applied to the IEF gel. IEF tube gels (120 × 5 mm) were prepared according to O'Farrel (12) with minor modifications using Pharmalytes 3–10 and 4–6.5. After electrophoresis the gels were cut into 3-mm slices for determination of QC activity and pH. The gel slices were then applied either to sodium dodecyl sulfate (SDS)/polyacrylamide gel electrophoresis (PAGE) or to PAGE under nondenaturing conditions for electrophoresis in the second dimension. SDS/PAGE was performed in slab gels (120 × 140 × 1 mm) as described by Fling and Gregerson (13). Slices of the IEF gels were reequilibrated for SDS/PAGE in electrode buffer (10–20 min) and sealed on top of the stacking gel with 1% agarose in electrode buffer. Gels were stained with Coomassie brilliant blue R250 or with the silver staining method of Heukeshoven and Dernick (14).

The nondenaturing gels were prepared on the basis of the Pharmacia LKB Phast system for nondenaturing gels adapted for large gels (120 × 140 × 1 mm). The electrode buffer consisted of 220 mM glycine and 62.5 mM Tris, pH 8.8. The slices of the IEF gels were applied as described for SDS/PAGE. After electrophoresis under nondenaturing conditions the QC-activity-exhibiting vertical section of the slab gel was fractionated perpendicular to the direction of migration. The slices were then subjected to the QC assay and subsequently to SDS/PAGE in a third dimension. The slices were prepared for SDS/PAGE like the IEF gel slices.

Protein Sequence Analysis. Bovine anterior pituitary QC was purified by FPLC, IEF, and SDS/PAGE and electroblotted to poly(vinylidene difluoride) membrane (Millipore) (15). The protein band visualized on the blots by staining with Ponceau S was excised for Edman degradation in the presence of 3 mg of Biobrene. The analysis was performed with an automated protein sequencer (model 477A) equipped with an automated phenylthiohydantoin amino acid analyzer (model 120A, Applied Biosystems). Regular Applied Biosystems sequencer programs were used.

Preparation and Screening of the λgt11 cDNA Library. Poly(A)$^+$ RNA isolated from bovine anterior pituitaries was used to prepare an oligo(dT)-primed cDNA library in λgt11 phage vector with the Librarian XI kit (Invitrogen, San Diego). The library was screened with three different degenerate oligonucleotides designed according to Lathe's (16) criteria on the basis of the N-terminal amino acid sequence. Hybridization of duplicate filters was carried out at 65°C as described (17). Filters were subjected to three subsequent washes (10 min each) at 23°C, 55°C, and 60°C in 0.9 M NaCl/0.09 M sodium citrate, pH 7/1% SDS and exposed to Kodak x-ray film for 15 hr.

DNA Sequence Analysis. Sequence analysis of both strands was performed with a DNA sequencer (model 373A, Applied Biosystems) using the dideoxy chain-termination method (18). A series of nested deletions of these clones was generated by digestion with exonuclease III (Pharmacia) or appropriate restriction enzymes. Final gaps were bridged using synthetic oligonucleotide primers. Both strands were sequenced at least twice.

RNA Isolation and Analysis. Total RNA was purified with a modification (M.Z., unpublished results) of standard procedures, using 5 M guanidinium thiocyanate for extraction and LiCl for precipitation. Poly(A)$^+$ RNA was prepared from total RNA with the Pharmacia LKB kit and analyzed by electrophoresis (19) and hybridization on Hybond (Amersham) membranes (17). Blots were reprobed for β-actin mRNA.

Expression of QC cDNA in COS-7 Cells. The *Not* I-digested cDNA was inserted into the polycloning site of the pCDM8 expression vector (20). COS-7 cells were transfected with this construct by calcium phosphate precipitation (21). QC activity was determined in cell extracts prepared from cells 48 hr after transfection. Extracts were prepared as described for the pituitary extraction. Fraction Sn P2b and the supernatant of the 45,000 × *g* centrifugation were assayed for QC activity.

RESULTS

QC Purification and Partial Sequence Analysis. Bovine anterior pituitaries were extracted (7) to release QC from secretory granules, which had been recognized as a source of this putative processing enzyme on the basis of the subcellular QC activity distribution (7). For the development of a purification strategy it had to be taken into consideration that the enzymatic activity was irreversibly affected by low pH (<6), by most detergents, and by high concentrations of denaturing agents. The initial purification achieved with various techniques of FPLC yielded a 234-fold increase in the specific activity (Table 1). The final purification and identification of the QC protein were accomplished with two- and three-dimensional electrophoretic procedures such as IEF and PAGE under nondenaturing conditions as well as in the

Table 1. Purification of QC from bovine anterior pituitary

Fraction	Total protein, mg	Total activity, mmol/hr	Specific activity, μmol/mg per hour	Recovery, %	Purification, fold
Sn P2b	8417	8.03	0.954		
Mono Q	1286	7.6	5.9	95	6
ProRPC 10/10	131	4	30.6	50	32
Superose	42.2	1.82	43.2	23	45
ProRPC 5/2	3.6	0.81	224	10	235
IEF	0.72*	0.8	1,120	9	1,174
SDS/PAGE	0.038*	0.8†	21,280		(22,306)

*Values were estimated on the basis of stain intensity by comparison with standard protein (bovine serum albumin).
†Total activity before SDS/PAGE.

presence of SDS (Fig. 1). Since the enzymatic activity was retained after IEF and nondenaturing PAGE, QC activity could be monitored by the standard assay of the fractionated gels so that we were able to align directly the relative intensity of protein bands in gels with the corresponding level of QC activity. In a first series of two-dimensional PAGE experiments, QC-containing slices of the IEF gel were subjected to SDS/PAGE. The stain intensity of only one protein band correlated positively with the enzymatic activity of the IEF gel fractions (Fig. 1). This protein, characterized by a pI of 5.5 and an apparent molecular weight of 38,000 was electroblotted to poly(vinylidene difluoride) membranes with yields of 60–120 pmol as judged by comparison with standard proteins. Subsequent automated Edman degradation of the blotted protein revealed up to 28 N-terminal amino acids ($n = 4$) (Fig. 2). To confirm the identity of the analyzed protein band as QC, a second series of experiments was performed with a three-dimensional PAGE system. For this purpose, QC-containing slices of the IEF gel were transferred to a nondenaturing gel for PAGE. Slices of this gel with QC activity were applied to SDS/PAGE and subsequently subjected to electroblotting and Edman degradation. Only a few (up to 6) N-terminal amino acids were determined in these experiments, because of increased losses of material. However, the data were in full agreement with the results of Edman degradation after IEF and SDS/PAGE. Thus, the analyzed protein was found to exhibit QC activity and to be homogeneous under conditions (such as IEF, nondenaturing PAGE, and SDS/PAGE) utilizing different physicochemical properties of proteins. On this basis, the analyzed protein was identified as QC with a high level of confidence. Since the values for the apparent molecular weight determined with gel filtration (under nonreducing conditions, data not shown) and SDS/PAGE (under reducing conditions) did not differ significantly, it was suggested that QC is a single-chain protein.

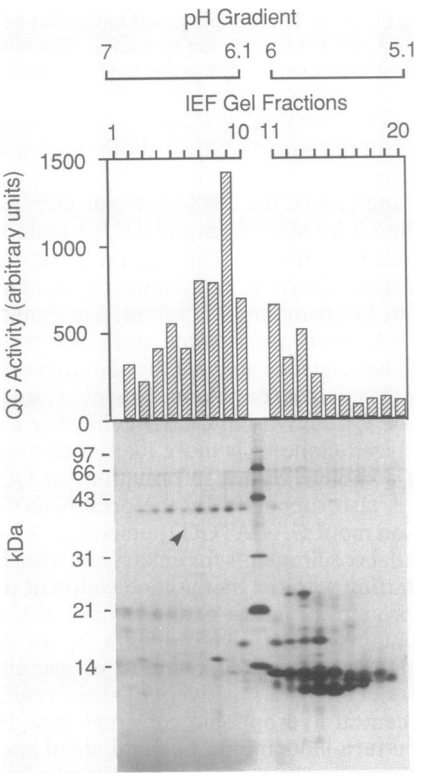

FIG. 1. Identification of QC by two-dimensional PAGE. An aliquot of a purified QC preparation was applied to IEF and subsequent SDS/PAGE. QC bands are marked with an arrowhead. Protein molecular mass markers are shown in the central lane.

Molecular Cloning of QC. We designed three sets of degenerate oligonucleotides to screen the bovine pituitary cDNA library. Only with the degenerate probe 5′-GG(C/G/T/A)GC(T/C)GT(G/C)GA(T/C)TGGAC(A/C)CA(A/G)GA(A/G)AA(A/G)AA(T/C)TA(T/C)(C/A)G(G/A)CA(A/G)CC(T/C)GC(C/T)CT(G/C)CT-3′ did we obtain 5 positive clones out of 5×10^5 clones screened. Intact inserts from these clones were isolated after digestion of the DNA with *Not* I. The longest insert was subcloned in pEMBL18 *Not* I, pUC18, and M13mp18 for further analysis and DNA sequence determination. For this purpose an additional *Not* I restriction site was inserted into the polylinker regions of pUC18 and M13mp18. The longest insert [2.1 kilobases (kb)] was subjected to sequence analysis at least twice in each orientation. The other clones (three of ≈2.1 kb, one of 0.5 kb) showed identical restriction analysis patterns with *Eco*RI and *Pst* I. Translation of the longest open reading frame within the 2088-base-pair insert resulted in a 361-residue protein including a putative prepropeptide sequence of 31 amino acids (Fig. 2). The protein sequence from the N terminus of the isolated protein matched exactly the sequence at nucleotide positions 181–264. A computer search of the May 1991 edition of the EMBL/GenBank data base using the program of Devereux *et al.* (22) did not reveal >60% sequence identity with any of the sequences listed in the data base.

Expression in COS-7 Cells. The identity of the analyzed protein as QC was demonstrated by functional expression of the corresponding cDNA in COS-7 monkey cells. The *Not* I-digested cDNA was inserted into the polycloning site of the pCDM8 expression vector. Highest rates of expression were achieved with 10 μg of pCDM8/cDNA per 10^7 COS-7 cells. The specific activity observed in extracts of transfected cells with the standard QC assay was 10,600, 11,000, and 15,000 nmol/mg per hour for 2.5, 5.0, and 10 μg of pCDM8/cDNA construct, respectively, per 10^7 transfected COS-7 cells. These values represented ≈10,000-fold higher specific activity than that of the crude pituitary extract. Significant QC activity (<0.01 nmol/mg per hour) could not be detected in the Sn P2b fraction prepared from nontransfected control cells or from COS-7 cells transfected with a construct that was identical except for the reverse orientation of the cDNA insert. The supernatant of the 45,000 × *g* centrifugation did not show QC activity.

Northern Blot Analysis. The tissue distribution of QC mRNA was examined by Northern blot analysis of poly(A)$^+$ RNA prepared from bovine pituitary, various brain regions, and peripheral tissues (Fig. 3). Highest expression was found for anterior pituitary and striatum. QC mRNA was expressed to a lesser extent in neurointermediate pituitary, cerebral cortex, hypothalamus, anterior and posterior thalamus, cerebellum, hippocampus, and retina and to a small extent in brainstem, heart atrium, heart ventricle, kidney, thymus, and skeletal muscle. No hybridization signal was observed with preparations from spleen, lung, and liver. Reprobing of all RNA blots for β-actin mRNA demonstrated that the RNA was not degraded and that equal amounts of RNA had been applied to the hybridization analysis.

DISCUSSION

We have purified QC from bovine pituitaries by a combination of FPLC and PAGE techniques using a specific assay, which was optimized in terms of speed and sensitivity.

Although the identification of QC for partial sequence analysis was accomplished in two- and three-dimensional PAGE with a high level of confidence, it was of crucial importance that the QC cDNA clone isolated on the basis of the protein chemical data produced QC activity in transfected COS-7 cells. Thus, unambiguous evidence was provided that the purified protein was the enzyme searched for. Since it has

```
   1  GGAGAAGGGGGAAGGCGAAGGACGCGCGTTCCCGGGCTCGTGACCGCCGGCGGCTCCAGGAACCGGCACCCAGACAGCCCAGGAGAGATGGCAGGCTGCAGGGACCCGCGCGTCGTGGAC  120
                                                                                                 MetAlaGlyCysArgAspProArgValValAsp   11

 121  ACCCTCCACCTGCTGCTGCTGGTGGCCGTCCTACCCCTGGCAGTCAGTGGGGTCCGTCGGGGCGCGGTAGACTGGACCCAAGAGAAGAATTATCACCAACCAGCCCTTTTGAATGTATCA  240
  12  ThrLeuHisLeuLeuLeuLeuValAlaValLeuProLeuAlaValSerGlyValArgArgGlyAlaValAspTrpThrGlnGluLysAsnTyrHisGlnProAlaLeuLeuAsnValSer   51

 241  TCTCTTCGGCAAGTTGCAGAAGGCACCAGTATTTCTGAAATGTGGCAAAATGACTTACGACCATTGCTGATAGAGCGATATCCTGGATCGCCTGGAAGCTTTGCTGCTCGCCAGCACATC  360
  52  SerLeuArgGlnValAlaGluGlyThrSerIleSerGluMetTrpGlnAsnAspLeuArgProLeuLeuIleGluArgTyrProGlySerProGlySerPheAlaAlaArgGlnHisIle   91

 361  ATGCAGAGAATTCAGAGACTTCAAGCTGACTGGGTCTTGGAAGTAGACACCTTCTTGAGTCAGACCCCCTACGGGTACCGGTCTTTCTCGAATATCATCAGCACCCTCAACCCCACTGCT  480
  92  MetGlnArgIleGlnArgLeuGlnAlaAspTrpValLeuGluValAspThrPheLeuSerGlnThrProTyrGlyTyrArgSerPheSerAsnIleIleSerThrLeuAsnProThrAla  131

 481  AAACGACACTTGGTCCTCGCCTGTCACTATGACTCCAAGTATTTTCCTCATTGGGACGACAGAGTGTTTGTGGGAGCCACTGATTCAGCCGTGCCATGTGCAATGATGCTGGAACTTGCT  600
 132  LysArgHisLeuValLeuAlaCysHisTyrAspSerLysTyrProHisTrpAspAspArgValPheValGlyAlaThrAspSerAlaValProCysAlaMetMetLeuGluLeuAla  171

 601  CGGGCCTTAGACAAGCAACTCTTTTCCTTGAAGAATATCTCAGACTCAAGGCCAGATCTTTCACTCCAGCTGATTTTCTTTGATGGAGAAGAGGCTTTTCATCTCTGGTCTCCTCAGGAT  720
 172  ArgAlaLeuAspLysGlnLeuPheSerLeuLysAsnIleSerAspSerArgProAspLeuSerLeuGlnLeuIlePhePheAspGlyGluGluAlaPheHisLeuTrpSerProGlnAsp  211

 721  TCTCTGTATGGGTCTCGCCACTTAGCTTCCAAGATGGCGTCAACACCACACCCACCTGGAGCGAGGGACACCAACCAACTGCATGGCATGGATTTATTGGTCTTACTGGATTTAATTGGA  840
 212  SerLeuTyrGlySerArgHisLeuAlaSerLysMetAlaSerThrProHisProProGlyAlaArgAspThrAsnGlnLeuHisGlyMetAspLeuLeuValLeuLeuAspLeuIleGly  251

 841  GCTCCATTCCCAACATTTCCCAACTTTTTCCCAAACACTGCCAGATGGTTTGGTAGACTTGAAGCAATTGAACATGGACTCCGTGAATTAGGTTTACTCAAGGATCACTCTTCGGAGAGG  960
 252  AlaProPheProThrPheProAsnPhePheProAsnThrAlaArgTrpPheGlyArgLeuGluAlaIleGluHisGlyLeuArgGluLeuGlyLeuLeuLysAspHisSerSerGluArg  291

 961  TGGTATTTCCGGAATTATGGTTATGGAGGTGTGATTCAGGATGACCATATTCCATTTTTAAGAAGAGGTGTTCCAGTTTTGCATCTGATATCATCTCCTTTCCCTGAAGTCTGGCACACC 1080
 292  TrpTyrPheArgAsnTyrGlyTyrGlyGlyValIleGlnAspAspHisIleProPheLeuArgArgGlyValProValLeuHisLeuIleSerSerProPheProGluValTrpHisThr  331

1081  ATGGATGACAATGAAGAAAATTTGGATAGAACAACCATTGACAATCTAAACAAAATTCTACAAGTATTTGTGTTAGAATATCTTCATTTGTAATATTCTGGTTTAGTTTACTGGTGAACT 1200
 332  MetAspAspAsnGluGluAsnLeuAspArgThrThrIleAspAsnLeuAsnLysIleLeuGlnValPheValLeuGluTyrLeuHisLeu*                              371

1201  AAATATTCTGGTGATTGCATCTAGTATCATATTCAAAAATCAAGGAATCCTTTAAAATAATCTGACTCTAGACAATGCTGTGTGGAAACTTCTATCCTATAGATTCTTTCTGTAGGTATC 1320

1321  TTTGAATATCCTAGATTTATGTCATGTCCAAATTGCTCATTAATTTTCATCTAGTGATAAAGGCAAATGTAAGAAAAACTGAAGATACATATCTCTAAAAAAACTCCTTTAGATAAGAAG 1440

1441  CTATGAAGCTAAATCAGGTGCATATAGTTAGTATCATGGTCTTCTAAAATAATATTTAATATGATGGTCAGGTGCAATGTATAGCCTTAGTTTTATGATTTTCTTATACTTGCAAATCTA 1560

1561  TTGCATATAAAACAAAACAGAGATGCATTTTGAGGACTCAATCTCTTTGAGAAATCTTTGTATGATTTATTTTATGGAAATTAAATCAAAATTAAATGCTATGTTGTAAATGTCTTTCTT 1680

1681  TATATATAAATGAAAAAAATCATATGATTTGAAGTATAGCTCAGTTTTATTGTTTACAAATTTTGCATTTTTATTGTAAAATAATATAGTCAGGACAAAAAAACGTAGAAAATAGAGAAA 1800

1801  GTAAAAAAGCCATAATTTCATTATATTAACCAAATCATCACTCATTCTAATTTTTCCTCATATATTTTCAATGAAGTTTATATGTGTTTATAATCAAACTCTATCTTCAATTTTTTGGCT 1920

1921  TTCTGTCTGTTAAATTATATAATTACCAGACATTTATTTGTTGTTTTGATCTTGGGTTACGATGAAGACAGAGATTGTTGAGTGTGTCCTCCAGAGTCAATATGAAGGTAACACAGAGAA 2040

2041  AGCCCAGCAATTATGATCATTTCTGTGTATTGTTGTATAAATTATATG  2088
```

FIG. 2. Nucleic acid and deduced amino acid sequence of the cDNA insert of clone λgt11-QC-5. The initiating methionine has been assigned position +1. The putative start (GAGATGG), stop (TAA), and polyadenylylation (AATAA, ATAAA) sites are overlined. The putative signal sequence of QC is marked by a dotted line; the putative propeptide sequence is marked by a double line, and the amino acid sequence determined by Edman degradation of purified bovine QC is underlined.

been observed (23, 24) that COS cells lack the dense secretory vesicles associated with proteolytic processing and therefore may not be able to process peptide precursors, it cannot be decided on the basis of the data presented here whether QC, a precursor to QC, or both were expressed in COS-7 cells.

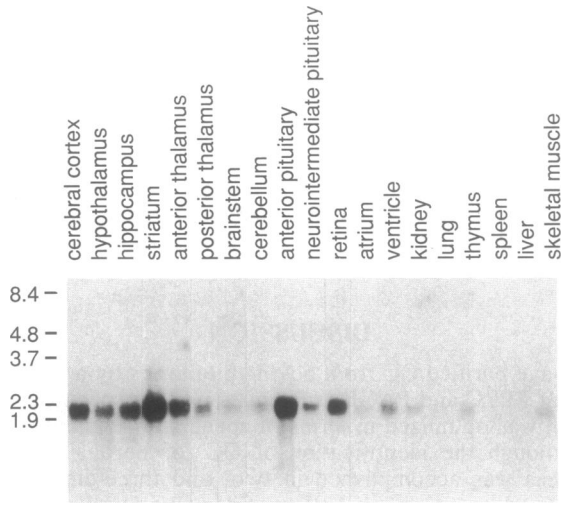

FIG. 3. Northern blot analysis of the distribution of QC mRNA in bovine pituitary, various brain regions, and peripheral tissues. Each lane contained 5 μg of poly(A)$^+$ RNA. *Bst*EII-digested phage λ DNA provided size marker shown at left in kilobases.

Sequence analysis of the 2088-base-pair cDNA insert derived from the 2.1-kb mRNA revealed a 361-amino acid open reading frame coding for QC. This open reading frame can be extended upstream from the methionine codon at nucleotides 88–90 without interruption. Therefore, we cannot exclude that the cDNA insert we isolated does not represent the complete coding sequence and that the actual translation start site might be absent from our clone. However, since we were able to express QC activity in COS-7 cells after transfection with the analyzed clone, it is more likely that the sequence contains the complete coding information for QC. This interpretation is also supported by the presence of the translation initiation motif GAGATGG (nucleotides 85–91) fulfilling the essential requirements for eukaryotic translation start sites (25). Starting with the methionine codon at nucleotides 88–90, the protein structure deduced from the nucleotide sequence contains at the N terminus a 27 (or 28)-residue sequence that possesses key features of secretory signal peptides (26), including a positively charged region at its N terminus, a central hydrophobic core region of 10 residues followed by a turn-inducing proline and small apolar amino acids (Ser-27 or Gly-28) representing typical cleavage sites for signal peptidase. Whether the signal peptide is cleaved at Ser-27 or Gly-28 has not been determined.

The finding of a signal sequence indicates that QC is targeted to the endoplasmic reticulum. QC, like peptidylgly-

cine α-amidating monooxygenase (PAM), does not exhibit a retention signal for soluble proteins of the endoplasmic reticulum [Lys-Asp-Glu-Leu (KDEL) motif] (27). Thus, QC may be present but not restricted to compartments of the endoplasmic reticulum. It had been observed (7) that the subcellular distribution of QC is parallel to the processing enzyme PAM, which has been demonstrated (28) to be located in secretory granules. This observation, together with the finding of a signal sequence and the substrate specificity of QC, supports the assumption that QC (like PAM) represents a processing enzyme.

The putative signal sequence and the protein sequence determined by Edman degradation are separated by the sequence (Gly-)Val-Arg-Arg (amino acids 28–31). Since a pair of arginine residues is recognized as a typical cleavage site for posttranslational processing, it is suggested that this sequence is a propeptide sequence removed during maturation.

As mentioned above, an apparent molecular weight of 38,000–40,000 was found in experiments using SDS/PAGE and gel filtration. On the basis of the primary structure deduced from the cDNA coding for QC, a molecular weight of 36,300 was calculated. The difference between these values may suggest the presence of nonprotein components such as carbohydrates in the QC molecule. No evidence for glycosylation was found in preliminary binding studies with carbohydrate affinity columns (data not shown). Obvious glycosylation sites were not detected in the QC sequence. However, further investigation is required to verify the absence of glycosyl residues in QC.

The pI, the apparent molecular weight, and the pH optimum found for bovine anterior pituitary QC contrast with the corresponding data of other QCs. For porcine QC a molecular weight of 55,000 and two distinct pI values of 5.7 and 7.2 were determined (9). A molecular weight of 25,000 was found for the plant enzyme (8). Similarly, the pH optima of QC from different sources were determined to be pH 8.0 (bovine pituitary), 7.2–7.5 (porcine pituitary), and 8.4 (papaya latex), respectively (7–9). Differences were also observed as to carbohydrate content and sensitivity toward inhibitors, salt, transition metals, and sulfhydryl-reactive agents. Thus, QCs of different sources exhibit significant heterogeneity.

The data presented here point also to heterogeneity of mammalian QC in the same species. On the basis of Northern blot analysis, no expression of QC was observed in bovine spleen, although the spleen is a major site of synthesis of immunoglobulins, which contain—to a significant extent—pyroglutamic residues (29). We speculate that at least two functionally distinct forms of QC may exist in mammals. Posttranslational processing of peptide hormones may be accomplished by the QC characterized here and may represent part of a regulated pathway of secretion, whereas a different form of the enzyme may be involved in the constitutive secretion of proteins such as immunoglobulins. The observed distribution of QC mRNA in the central nervous system contrasts with the corresponding data reported for the processing enzymes PAM (30) and carboxypeptidase H (31) as well as Kex2-related proteins such as PC1 and PC2 (32), which are considered as putative processing enzymes exhibiting endoproteolytic activity. On the basis of Northern blots, all proteins mentioned above are widely distributed in the brain; however, especially high expression in pituitary and hypothalamus—areas of active peptide synthesis—appears to be a common feature. The high relative expression in striatum has been observed only for QC. This high expression may not be completely explainable by the synthesis of known pyroglutamyl peptides but may point to the expression of (hormonal) pyroglutamyl peptides unknown to date.

We thank R. S. Allen for excellent technical help. We thank Dr. K. Eckart and Dr. B. Penke for the supply of synthetic peptides.

1. Vale, W. & Rivier, C. (1975) in *Handbook of Psychopharmakology*, eds. Iverson, L. L., Iverson, S. D. & Snyder, S. H. (Plenum, New York), Vol. 5, pp. 195–238.
2. Vale, W., Spiess, J., Rivier, C. & Rivier, J. (1981) *Science* **213,** 1394–1397.
3. Richter, K., Kawashima, E., Egger, R. & Kreil, G. (1984) *EMBO J.* **3,** 617–621.
4. Lechan, R. M., Wu, P., Jackson, I. M. D., Wolf, H., Cooperman, S., Mandel, G. & Goodman, R. H. (1986) *Science* **231,** 159–161.
5. Seeburg, P. H. & Adelman, J. P. (1984) *Nature (London)* **311,** 666–668.
6. Blombäck, B. (1968) *Methods Enzymol.* **11,** 398–411.
7. Fischer, W. H. & Spiess, J. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 3628–3632.
8. Messer, M. & Ottesen, M. (1964) *Biochim. Biophys. Acta* **92,** 409–411.
9. Busby, W. H., Jr., Quackenbusch, G. E., Humm, J., Youngblood, W. W. & Kizer, J. S. (1987) *J. Biol. Chem.* **262,** 8532–8536.
10. Jonsson, G. P., Hedin, A. B., Håkansson, P. L., Sundquist, B. U. R., Säve, B. G. S., Nielsen, P. F., Roepstorff, P., Johansson, K.-E., Kamensky, I. & Lindberg, M. S. L. (1986) *Anal. Chem.* **58,** 1084–1087.
11. Koger, J. B., Humm, J. & Kizer, J. S. (1989) *Methods Enzymol.* **168,** 358–365.
12. O'Farrel, P. H. (1975) *J. Biol. Chem.* **250,** 4007–4021.
13. Fling, S. P. & Gregerson, D. S. (1986) *Anal. Biochem.* **155,** 83–88.
14. Heukeshoven, J. & Dernick, R. (1985) *Electrophoresis* **6,** 103–112.
15. Matsudaira, P. (1987) *J. Biol. Chem.* **262,** 10035–10038.
16. Lathe, R. (1985) *J. Mol. Biol.* **183,** 1–12.
17. Church, G. M. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1991–1995.
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
19. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
20. Seed, B. (1987) *Nature (London)* **329,** 840–842.
21. Gorman, C. M., Moffat, L. M. & Howard, B. N. (1982) *Mol. Cell. Biol.* **2,** 1044–1051.
22. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395.
23. Lapps, W., Eng, J., Stern, A. S. & Gubler, U. (1988) *J. Biol. Chem.* **263,** 13456–13462.
24. Thomas, G., Thorne, B. A. & Hruby, D. E. (1988) *Annu. Rev. Physiol.* **50,** 323–332.
25. Kozak, M. (1986) *Cell* **44,** 283–292.
26. von Heijne, G. (1986) *Nucleic Acids Res.* **14,** 4683–4690.
27. Pelham, H. R. B. (1990) *Trends Biochem. Sci.* **15,** 483–486.
28. Eipper, B. A., Mains, R. E. & Glembotski, C. C. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 5144–5148.
29. Dayhoff, M. O. (1972) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Silver Spring, MD), Vol. 5.
30. Braas, K. M., Stoffers, D. A., Eipper, B. A. & May, V. (1989) *Mol. Endocrinol.* **3,** 1387–1398.
31. MacCumber, M. W., Snyder, S. H. & Ross, C. A. (1990) *J. Neurosci.* **10,** 2850–2860.
32. Seidah, N. G., Gaspar, L., Mion, P., Marcinkiewicz, M., Mbikay, M. & Chrétien, M. (1990) *DNA Cell Biol.* **9,** 415–424.