Volume 234, number 1, 160–164 FEB 06032 July 1988

# Cloning and sequence analysis of human pituitary cDNA encoding the novel polypeptide 7B2

Gerard J.M. Martens

*Departments of Animal Physiology and Molecular Biology, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands*

Received 21 April 1988

Application of a differential hybridization technique led to the isolation of a human pituitary cDNA clone encoding the complete structure of the polypeptide 7B2. This protein of unknown function, which is sorted to secretory granules, appears to be present selectively in neurons and endocrine cells. The polypeptide chain of human 7B2, preceded by a cleaved signal peptide, comprises 185 amino acids (a calculated $M_r$ of 20793). Interesting features of the highly-conserved 7B2 structure include (i) a serine phosphorylation consensus sequence, (ii) the occurrence of three pairs of dibasic amino acids representing potential proteolytic cleavage sites and, in particular, (iii) the presence of three regions homologous to GTP-binding domains giving 7B2 structural characteristics of a GTP-binding protein.

7B2 protein; cDNA sequence; Differential hybridization; Peptide hormone; Protein secretion; (Human pituitary)

## 1. INTRODUCTION

The novel protein 7B2 appears to be selectively present in small amounts in neurons and endocrine cells [1–3]. Ultrastructural studies have shown that the 7B2 protein is compartmentalized within peptide hormone- and neuropeptide-containing secretory granules [4,5]. This 21 kDa protein has been partially sequenced following its isolation from porcine and human pituitary glands, the major source of the protein [6,7]. Both the involvement of 7B2 in the secretory process (protein traffic or prohormone maturation) as well as the possibility that 7B2 itself has some biological activity on target cells have been suggested [3,5]. Recent evidence indicated that immunoreactive 7B2 may be a plasma marker for endocrine tumors [8]. Clearly, the 7B2 protein is an interesting polypeptide which deserves further attention. As a step towards elucidation of the function of 7B2, we present in this report the nucleotide sequence of cloned human pituitary cDNA which codes for the complete primary structure of the 7B2 protein.

*Correspondence address:* G.J.M. Martens, Department of Animal Physiology, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number Y00757

## 2. MATERIALS AND METHODS

We applied a differential hybridization technique in order to identify a gene whose expression is associated with the secretory function of a peptide-secreting cell. A consideration of the physiological function of the secretory cells (melanotrophs) of the amphibian pituitary gland, namely the synthesis and secretion of melanophore-stimulating hormone (MSH) in animals on a black background [9,10], has allowed us to apply the differential hybridization technique. A pituitary cDNA library of black-adapted South-African clawed toads, *Xenopus laevis*, was constructed [11]. Single-stranded cDNA probes were prepared by oligo(dT) priming of 2.5 $\mu$g RNA isolated from intermediate pituitary glands of black-adapted toads (biosynthetically active cells) and white-adapted animals (biosynthetically inactive cells). With these probes 20000 colonies of the *Xenopus* pituitary cDNA library were screened. One of the differentially hybridizing *Xenopus* pituitary cDNA clones (clone pX9) was then used to screen a human pituitary cDNA library in the vector λgt10 (generously provided by Dr Peter van Wezenbeek, Organon BV, The Netherlands). The screening of $8 \times 10^5$ clones (standard hybridization solution, 58°C) with pX9 nick-translated to $5 \times 10^8$ cpm per $\mu$g DNA as probe resulted in the



*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/88/$3.50 © 1988 Federation of European Biochemical Societies

isolation of fourteen hybridization-positive clones. The nucleotide sequence reported here was assembled from the two overlapping cDNA inserts of clones λH6 (insert length 1 kb) and λH7 (1.2 kb). Sequencing on both strands and with overlapping M13 subclones was performed with the dideoxy chain termination method [12].

In Northern blot analysis, glyoxylated RNA samples were fractionated on a 1.4% agarose gel. After transfer [13], the blots were hybridized in standard hybridization solution in the presence of 50% formamide with nick-translated insert DNA of human pituitary 7B2 cDNA clone λH6 ($6 \times 10^8$ cpm per $\mu$g DNA).

## 3. RESULTS

### 3.1. *Cloning and nucleotide sequence of human pituitary 7B2 cDNA*

In the differential screening of the *Xenopus* pituitary cDNA library, one cDNA clone was isolated and then used to screen a human pituitary cDNA libray under low-stringency hybridization conditions. The nucleotide sequence of the two longest overlapping hybridization-positive human pituitary cDNA clones comprises 1152 bp, not including the poly(A) tail (fig.1). It contains 633 bp coding region and 491 bp 3′-untranslated region which includes two polyadenylation signals (AATAAA). The open reading frame codes for a protein of 211 amino acids. A computer homology analysis (EMBL nucleotide and NBRF protein data base) revealed that the N-terminal region of this protein (residues 1–77; fig.1) is identical to the N-terminal 77 amino acids that have been sequenced for the human pituitary protein 7B2 [7]. No significant homology with any other protein or DNA sequence was found.

```
                                                                  5'----CGCTCCTCGGGCTGCCCCTCGGTTGACA    28

-26                    -20                                 -10                                  -1
Met Val Ser Arg Met Val Ser Thr Met Leu Ser Gly Leu Leu Phe Trp Leu Ala Ser Gly Trp Thr Pro Ala Phe Ala
ATG GTC TCC AGG ATG GTC TCT ACC ATG CTA TCT GGC CTA CTG TTT TGG CTG GCA TCT GGA TGG ACT CCA GCA TTT GCT   106

  1                                10                                  20
Tyr Ser Pro Arg Thr Pro Asp Arg Val Ser Glu Ala Asp Ile Gln Arg Leu Leu His Gly Val Met Glu Gln Leu Gly
TAC AGC CCC CGG ACC CCT GAC CGG GTC TCA GAA GCA GAT ATC CAG AGG CTG CTT CAT GGT GTT ATG GAG CAA TTG GGC   184

            30                                  40                                  50
Ile Ala Arg Pro Arg Val Glu Tyr Pro Ala His Gln Ala Met Asn Leu Val Gly Pro Gln Ser Ile Glu Gly Gly Ala
ATT GCC AGG CCC CGA GTG GAA TAT CCA GCT CAC CAG GCC ATG AAT CTT GTG GGC CCC CAG AGC ATT GAA GGT GGA GCT   262

                    60                                  70
His Glu Gly Leu Gln His Leu Gly Pro Phe Gly Asn Ile Pro Asn Ile Val Ala Glu Leu Thr Gly Asp Asn Ile Pro
CAT GAA GGA CTT CAG CAT TTG GGT CCT TTT GGC AAC ATC CCC AAC ATC GTG GCA GAG TTG ACT GGA GAC AAC ATT CCT   340

     80                                  90                                  100
Lys Asp Phe Ser Glu Asp Gln Gly Tyr Pro Asp Pro Pro Asn Pro Cys Pro Val Gly Lys Thr Asp Asp Gly Cys Leu
AAG GAC TTT AGT GAG GAT CAG GGG TAC CCA GAC CCT CCA AAT CCC TGT CCT GTT GGA AAA ACA GAT GAT GGA TGT CTA   418

                    110                                 120                                 130
Glu Asn Thr Pro Asp Thr Ala Glu Phe Ser Arg Glu Phe Gln Leu His Gln His Leu Phe Asp Pro Glu His Asp Tyr
GAA AAC ACC CCT GAC ACT GCA GAG TTC AGT CGA GAG TTC CAG TTG CAC CAG CAT CTC TTT GAT CCG GAA CAT GAC TAT   496

                                140                                 150                    •
Pro Gly Leu Gly Lys Trp Asn Lys Lys Leu Leu Tyr Glu Lys Met Lys Gly Gly Glu Arg Arg Lys Arg Arg Ser Val
CCA GGC TTG GGC AAG TGG AAC AAG AAA CTC CTT TAC GAG AAG ATG AAG GGA GGA GAG AGA CGA AAG CGG AGG AGT GTC   574

            160                                 170                                 180
Asn Pro Tyr Leu Gln Gly Gln Arg Leu Asp Asn Val Val Ala Lys Lys Ser Val Pro His Phe Ser Asp Glu Asp Lys
AAT CCA TAT CTA CAA GGA CAG AGA CTG GAT AAT GTT GTT GCA AAG AAG TCT GTC CCC CAT TTT TCA GAT GAG GAT AAG   652

Asp Pro Glu ***
GAT CCA GAG TAA AGAGAAGATGCTAGACGAAAACCCACATTACCTGTTAGGCCTCAGCATGGCTTATGTGCACGTGTAAATGGAGTCCCTGTGAATGAC   751

AGCATGTTTCTTACATAGATAATTATGGATACAAAGCAGCTGTATGTAGATAGTGTATTGTCTTCACACCGATGATTCTGCTTTTTGCTAAATTAGAATAAGA   854

GCTTTTTTGTTTCTTGGGTTTTTAAAATGTGAATCTGCAATGATCATAAAAATTAAAATGTGAATGTCAACAATAAAAAGCAAGACTATGAAAGGCTCAGATT   957

TCTTGCAGTTTAAAATGGTGTCTGAGGTTGTACTATTTTGGCCAAGTCTGTAGAAAGCTGTCATTTGATTTTGATTATGTAGTTCATCCAGCCCTTGGGCATT  1060

GTTATACACCAGTAAAGAAGGCTGTACTCAAGAGGAGGAGCTGACACATTTCACTTGGCTGCGTCTTAATAAACATGAATGCAAGCATTGGC(A)n----3'  1152
```

Fig.1. Nucleotide sequence and deduced amino acid sequence of human pituitary cDNA encoding the 7B2 preprotein. Amino acid sequence numbering starts at the N-terminal residue of the mature 7B2 protein with the presumptive signal peptide sequence being indicated by negative numbering. The dot above Ser-155 indicates a potential cAMP-dependent protein kinase phosphorylation site [16]. The two signals for polyadenylation (AATAAA) are overlined.

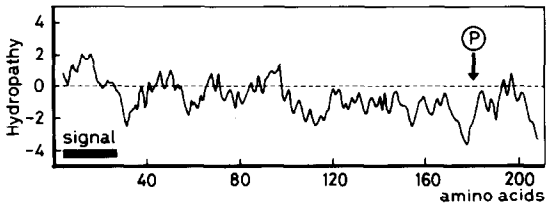


Fig.2. Hydropathy profile of the human 7B2 preprotein. The profile was plotted using the algorithms of Kyte and Doolittle [15] with a window size of seven residues. Positive points in the ordinate denote hydrophobic regions within the protein. Arrow indicates position of a potential cAMP-dependent protein kinase phosphorylation site (corresponding to Ser-155 in fig.1).

The cloned cDNA thus codes for the human 7B2 preprotein with an 18–26 amino acid signal peptide sequence for transport across endoplasmic reticulum membranes and a 185 amino acid mature 7B2 protein. It is not yet clear which one of the three methionine residues in the signal peptide is the initiator methionine. The prepeptide of 7B2 ends in Ala-Phe-Ala (fig.1); Ala-X-Ala is the most frequently occurring tripeptide preceding a signal peptide cleavage site [14]. The molecular mass of the mature protein deduced from the cDNA is 20793 Da, in agreement with that previously determined for 7B2 by SDS-polyacrylamide gel electrophoresis (21 kDa) [6].

### 3.2. *Analysis of the human 7B2 sequence*

The hydropathy plot according to Kyte and Doolittle [15] shows that the N-terminal portion of human 7B2 is more hydrophobic than the C-terminal region (fig.2). The 7B2 protein does not contain a long hydrophobic domain indicative of a membrane-spanning protein.

Human 7B2 contains two interesting potential targets for post-translational modifications. The first site concerns the presence of a consensus sequence [16] for serine phosphorylation by cAMP-dependent protein kinase (Arg-Arg-Lys-Arg-Arg-Ser$^{155}$; fig.1). Second, the 7B2 protein contains three sites with two or more adjacent basic amino acids (residues 138–139; 150–154; 171–172; fig.1). Such sites are potential recognition sites for proteolytic enzymes during the processing of prohormones to smaller bioactive peptides [17]. No potential site for N-linked glycosylation (Asn-X-Ser/Thr) [18] is present in 7B2.

An extensive search for amino acid sequence homology between the human 7B2 structure and functional domains of regulatory proteins was conducted. This search revealed within the 7B2 protein structure the presence of three regions which share significant homology with GTP-binding proteins in domains presumed to participate in GTP binding and hydrolysis (fig.3). These regions, indicated as regions A, B and C in fig.3, have been assigned on the basis of their conservation within GTP-binding proteins, X-ray crystallographic studies and site-directed mutagenesis experiments [19–21]. From the extensive homology shown in fig.3 it appears that the 7B2 protein is structurally related to GTP-binding proteins.

The 7B2 protein is highly conserved during 350 million years of vertebrate evolution since the human and *Xenopus* proteins exhibit an overall amino acid sequence homology of 83% and, in an N-terminal region comprising 80 amino acids, even 95% homology (conservative amino acid substitutions not included).

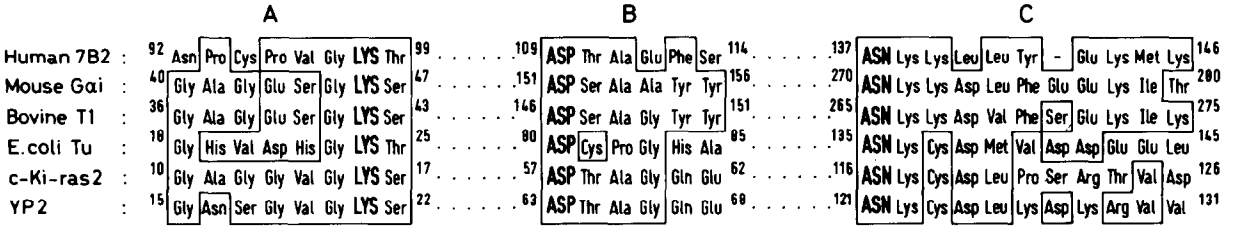


Fig.3. Comparison between human 7B2 protein regions and regions within GTP-binding proteins. The Lys, Asp and Asn residues shown in capital letters in regions A, B and C, respectively, have been implicated to be directly involved in GTP binding and hydrolysis [19–21]. Sets of identical or conservative [25] amino acid residues are boxed. Sequences of mouse $G_{\alpha i}$ [20], bovine transducin T1 [26], *E. coli* elongation factor Tu [27], *c-Ki-ras* oncogene product [28] and yeast GTP-binding protein YP2 [29].

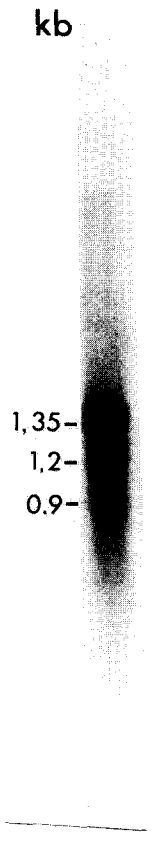


Fig.4. Northern blot analysis of 7B2 mRNA. Total cellular RNA (25 $\mu$g) from human pituitary was hybridized with human pituitary 7B2 cDNA clone λH6. The autoradiogram was exposed for 20 h at −70°C with two intensifying screens.

### 3.3. *The 7B2 mRNA*

We used the cloned human 7B2 cDNA in Northern blot analysis and detected a ~1.35 kb transcript in human pituitary (fig.4), indicating that the sequence presented in fig.1 represents the nearly complete human 7B2 mRNA sequence (allowing for a poly(A) tail of 100 nucleotides). The presence of two additional minor transcripts of ~1.2 kb and ~0.9 kb (fig.4) suggests the occurrence of 7B2-related mRNA species in the human pituitary gland.

## 4. DISCUSSION

The biological role of the widespread and evolutionarily conserved 7B2 protein is not known. The primary structure of human 7B2 reported here may lead to a better understanding of the function of this protein. The polypeptide chain of 7B2 is preceded by a cleaved signal peptide, indicating that during biosynthesis this protein is targeted to the membrane of the rough endoplasmic reticulum. The occurrence of three pairs of basic residues in the 7B2 structure raises the possibility that 7B2 is a precursor protein for smaller peptides.

One feature of the 7B2 structure deserves particular attention, namely the presence of putative GTP-binding domains. This might indicate that 7B2 represents a member of a novel family of low-molecular mass, secretory granule-associated GTP-binding proteins. Interestingly, recent evidence suggests that GTP-binding proteins participate in translocation of proteins across the endoplasmic reticulum [22], intercisternal protein transport within the Golgi stack [23] and fusion of secretory granule and plasma membranes during triggered exocytotic secretion [24]. Our cloning of the 7B2 cDNA provides tools to study whether the 7B2 protein is a GTP-binding protein involved in the pathway of protein secretion.

*Acknowledgements:* The author thanks Myriam van Miert for technical assistance, Dr Peter van Wezenbeek for providing the human pituitary cDNA library, Marleen Verbeeck and Peter Burbach for Northern blots, Dr Jack Leunissen for computer analyses and Dr Bruce Jenks for critical reading of the manuscript. This work was supported by the Netherlands Organization for Scientific Research (NWO).

## REFERENCES

[1] Iguchi, H., Chan, J.S.D., Seidah, N.G. and Chrétien, M. (1984) Neuroendocrinology 39, 453–458.
[2] Iguchi, H., Chan, J.S.D., Dennis, M., Seidah, N.G. and Chrétien, M. (1985) Brain Res. 338, 91–96.
[3] Iguchi, H., Chan, J.S.D., Seidah, N.G. and Chrétien, M. (1987) Peptides 8, 593–598.
[4] Marcinkiewicz, M., Benjannet, S., Cantin, M., Seidah, N.G. and Chrétien, M. (1986) Brain Res. 380, 349–356.
[5] Marcinkiewicz, M., Benjannet, S., Seidah, N.G., Cantin, M. and Chrétien, M. (1987) Cell Tissue Res. 250, 205–214.
[6] Hsi, K.L., Seidah, N.G., De Serres, G. and Chrétien, M. (1982) FEBS Lett. 147, 261–266.
[7] Seidah, N.G., Hsi, K.L., De Serres, G., Rochemont, J., Hamelin, J., Antakly, T., Cantin, M. and Chrétien, M. (1983) Arch. Biochem. Biophys. 225, 525–534.

[8] Suzuki, H., Ghatei, M.A., Williams, S.J., Uttenthal, L.O., Facer, P., Bishop, A.E., Polak, J.M. and Bloom, S.R. (1986) J. Clin. Endocrinol. Metab. 63, 758–765.
[9] Jenks, B.G., Van Overbeeke, A.P. and McStay, B.F. (1977) Can. J. Zool. 55, 922–927.
[10] Martens, G.J.M., Weterings, K.A.P., Van Zoest, I.D. and Jenks, B.G. (1987) Biochem. Biophys. Res. Commun. 143, 678–684.
[11] Martens, G.J.M., Civelli, O. and Herbert, E. (1985) J. Biol. Chem. 260, 13685–13689.
[12] Sanger, F., Nicklen, S. and Coulson, A.P. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[13] Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201–5205.
[14] Perlman, D. and Halvorson, H.O. (1983) J. Mol. Biol. 167, 391–409.
[15] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[16] Feramisco, J.R., Glass, D.B. and Krebs, E.G. (1980) J. Biol. Chem. 255, 4240–4245.
[17] Douglass, J., Civelli, O. and Herbert, E. (1984) Annu. Rev. Biochem. 53, 665–715.
[18] Pless, D.D. and Lennarz, W.J. (1977) Proc. Natl. Acad. Sci. USA 74, 134–138.
[19] Jurnak, F. (1985) Science 230, 32–36.
[20] Sullivan, K.A., Liao, Y.-C., Alborzi, A., Beiderman, B., Chang, F.-H., Masters, S.B., Levinson, A.D. and Bourne, H.R. (1986) Proc. Natl. Acad. Sci. USA 83, 6687–6691.
[21] Wagner, P., Molenaar, C.M.T., Rauh, A.J.G., Brokel, R., Schmitt, H.D. and Gallwitz, D. (1987) EMBO J. 6, 2373–2379.
[22] Connolly, T. and Gilmore, R. (1986) J. Cell Biol. 103, 2253–2261.
[23] Melançon, P., Glick, B.S., Malhotra, V., Weidman, P.J., Serafini, T., Gleason, M.L., Orci, L. and Rothman, J.E. (1987) Cell 51, 1053–1062.
[24] Barrowman, M.M., Cockcroft, S. and Gomperts, B.D. (1986) Nature 319, 504–507.
[25] Dayhoff, M.O. (1978) in: Atlas of Protein Sequence and Structure (Dayhoff, M.O. ed.) pp.345–352, National Biomedical Research Foundation, Washington, DC.
[26] Tanabe, T., Nukada, T., Nishikawa, Y., Sugimoto, K., Suzuki, H., Takahashi, H., Noda, M., Haga, T., Ichiyama, A., Kangawa, K., Minamino, N., Matsuo, H. and Numa, S. (1985) Nature 315, 242–245.
[27] McGrath, J.P., Capon, J.D., Smith, D.H., Chen, E.Y., Seeburg, P.H., Goeddel, D.V. and Levinson, A.D. (1983) Nature 310, 644–649.
[28] Laursen, R.A., L'Italien, J.J., Nagarkatti, S. and Miller, D.L. (1981) J. Biol. Chem. 256, 8102–8109.
[29] Gallwitz, D., Donath, C. and Sander, C. (1983) Nature 306, 704–707.