Volume 264, number 1, 117–120 FEBS 08409 May 1990

Nucleotide sequence determination of chicken glucagon precursor cDNA

Chicken preproglucagon does not contain glucagon-like peptide II

Shin Hasegawa, Kimio Terazono*, Koji Nata*, Tatsuyuki Takada, Hiroshi Yamamoto* and Hiroshi Okamoto*

Department of Animal Nutrition, Faculty of Agriculture, Kobe University, Kobe 657, Hyogo and **Department of Biochemistry, Tohoku University School of Medicine, Sendai 980, Miyagi, Japan*

Received 12 March 1990

cDNA clones coding for glucagon were isolated from a chicken pancreas cDNA library, and the nucleotide and amino acid sequences were determined. The amino acid sequence of chicken glucagon was HSQGTFTSDYSKYLDSRRAQDFVQWLMST, which was contained in the 151-amino acid long precursor, being preceded by a signal sequence and an amino-terminal peptide (NH_2-peptide) and followed by an intervening peptide and a glucagon-like peptide I (GLP-I). Chicken preproglucagon, however, lacked GLP-II and intervening peptide II which have been shown to be contained in mammalian glucagon precursors.

Preproglucagon; Glucagon; Glucagon-like peptide; mRNA; cDNA cloning; *Gallus gallus domesticus*

1. INTRODUCTION

Glucagon, a 29-amino acid peptide hormone, is synthesized from a larger precursor, preproglucagon [1], and acts as a regulator of hepatic and pancreatic functions [2–4]. cDNAs coding for preproglucagon have been identified in mammals (rat [5], human [6], hamster [7], guinea pig [8] and bovine [9]) and in fish (anglerfish) [10,11]. Although chicken glucagon has been isolated from the pancreas [12], the complete amino acid sequences of chicken glucagon and its precursor remain to be determined.

In the present study, we have determined the nucleotide sequence of chicken glucagon cDNA and deduced the amino acid sequence of chicken glucagon and its precursor. The precursor did not contain the glucagon-like peptide II (GLP-II) that has been found in mammalian glucagon precursors.

2. MATERIALS AND METHODS

2.1. *Construction of chicken pancreas cDNA library*

Pancreases were removed from 6-week-old male chickens (*Gallus gallus domesticus*). Total RNA was extracted from the pancreas as described [13] using cesium trifluoroacetate. Poly(A)$^+$ RNA was isolated by oligo(dT)-cellulose column chromatography [14]. A cDNA library was constructed as described previously [15] using λZAPII and *Escherichia coli* XL1-Blue (Stratagene, La Jolla, CA).

Correspondence address: H. Okamoto, Department of Biochemistry, Tohoku University School of Medicine, Sendai 980, Miyagi, Japan

The nucleotide sequence(s) presented here has (have) been submitted to the EMBL/GenBank database under the accession number no. Y07539

2.2. *Cloning of chicken preproglucagon cDNA*

A 60-base oligodeoxyribonucleotide complement of rat preproglucagon mRNA (5′-TTGAGCACGGCGGGAGTCTAGGTATTTGCTGTAGTCACTGGTGAATGTGCCCTGTGAATG-3′, nucleotide residues 217–276 in [5]) was synthesized using an Applied Biosystems Model 380B DNA synthesizer. After 5′-labeling using [γ-^{32}P]ATP (Amersham, Buckinghamshire, UK) and T4 polynucleotide kinase (Takara Shuzo, Kyoto, Japan), the oligodeoxyribonucleotide was purified [16] and used to screen plaques lifted onto nitrocellulose filters. Recombinant phage DNAs in the hybridization-positive clones were excised and recircularized in vivo [17], and the resulting Bluescript plasmids were isolated.

2.3. *DNA sequencing*

Cloned cDNAs were cleaved with various restriction endonucleases and subcloned into pBS vectors (Stratagene). The nucleotide sequence of each restriction fragment was determined as described previously [18].

2.4. *Northern blot analysis*

RNA was electrophoresed on a 1.5% agarose gel, transferred onto a nitrocellulose filter [19], and hybridized with the 5′-labeled 60-base oligodeoxyribonucleotide or with the cloned cDNA which had been labeled with [α-^{32}P]dCTP (Amersham) by the random priming method [20]. Hybridization with the oligodeoxyribonucleotide probe was carried out in 5 × SSPE (1 × SSPE; 0.18 M NaCl, 10 mM sodium phosphate buffer, pH 7.4, and 1 mM EDTA), 5 × Denhardt's solution, 40% formamide, 0.1% SDS and 100 μg/ml *E. coli* tRNA at 42°C for 19 h, followed by washing in 2 × SSPE and 0.1% SDS at room temperature for 30 min twice and in 0.1 × SSPE and 0.1% SDS at 32°C for 10 min once. Hybridization with the cDNA probe was carried out at the stringent criterion [15].

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

3. RESULTS AND DISCUSSION

A cDNA library (7.8×10^5 plaque forming units) was constructed from chicken pancreas poly(A)$^+$ RNA and screened with the 60-base oligodeoxyribonucleotide corresponding to the 20-amino acid amino terminal of rat glucagon [5]. Sixty positive clones were identified, indicating that the relative abundance of the mRNA in chicken pancreas poly(A)$^+$ RNA was approximately 0.008%. Three independent clones carrying the longest cDNA were isolated and the cDNA insert was purified. The cDNA as well as the 60-base oligodeoxyribonucleotide hybridized to a single RNA species of about 1.7 kb long in Northern blot analysis of chicken pancreas RNA (Fig. 1).

The nucleotide sequence of the cloned cDNA was determined according to the strategy as shown in Fig. 2A. The cDNA comprised 1576 nucleotides plus poly(A), and appeared to represent a nearly full-length copy of the mRNA. As shown in Fig. 2B, the cDNA had one large open reading frame of 453 nucleotides, the 5′-untranslated region of 88 nucleotides and the 1035-nucleotide long 3′-untranslated region, on the assumption that ATG at nucleotides 1–3 is the start codon and TAA at nucleotides 454–456 is the stop codon. Two polyadenylation signals, AATAAA, were present at nucleotide residues 1456–1461 and 1472–1477. The nucleotide sequence determined was identical in the cDNA inserts of the three independent clones (data not shown).

The amino acid sequence deduced from the nucleotide sequence encoded a 151-amino acid protein (Fig. 2B). The sequence of 29 amino acids from residues 55–83 was found to correspond to the chicken glucagon sequence predicted through analyses of the amino acid composition and of the partial amino acid sequence of the purified peptide from chicken pancreas [12]. Therefore, the 151-amino acid protein encoded by the cDNA was regarded as the precursor of chicken glucagon. At the amino terminus of the precursor is a 22-amino acid sequence having the characteristics of a signal peptide with a cleavage site as defined by von Heijne [21]. The signal sequence is followed by a region of 30 amino acids, Lys-Arg and the glucagon sequence. On the carboxyl terminus of glucagon are Lys-Arg, followed by a region of 24 amino acids, Lys-Arg, a 37-amino acid carboxyl-terminal peptide, and Arg-Arg-Glu. Thus, chicken glucagon is probably synthesized first as a 151-amino acid long precursor (preproglucagon): following the cotranslational removal of the 22-amino acid signal peptide from the preproglucagon, a 129-amino acid proglucagon is generated. The proglucagon is likely to be further cleaved at the pairs of basic amino acids to liberate the 30-amino acid amino-terminal peptide, glucagon, the 24-amino acid intervening peptide and the 37-amino acid carboxyl-terminal peptide.

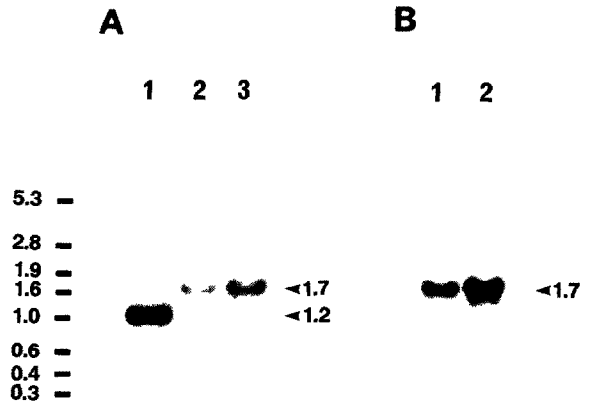

Fig. 1. Northern blot analysis of chicken pancreas RNA. RNA blots were hybridized with the 60-base oligodeoxyribonucleotide complementary to rat glucagon mRNA (A) and with the cloned chicken preproglucagon cDNA (B). A: lane 1, RNA from rat pancreatic islets of Langerhans (450 ng); 2 and 3, RNA from chicken pancreas (25 and 50 μg, respectively). The rat RNA species that hybridized to the oligonucleotide probe was 1.2 kb long, the same as reported for rat glucagon mRNA [5]. B: lanes 1 and 2, RNA from chicken pancreas (25 and 50 μg, respectively). The position to which RNA standards of the size designated (in kb) migrate is shown to the left.

As shown in Fig. 3, we compared the structure of the chicken glucagon precursor with those of rat [5] and human [6] preproglucagons and anglerfish preproglucagon I [10]. The precursors have potential processing sites at similar positions, and can be divided into several functional domains. Chicken glucagon differed by only one amino acid from rat and human glucagon and also exhibited a high homology (68%) with anglerfish glucagon.

The 37-amino acid carboxyl-terminal peptide, which was homologous to chicken glucagon itself (41%), was found to be highly homologous to rat and human GLP-I (87%) and to anglerfish GLP-I (53%), but to have only about 30% homology with rat and human GLP-II. The 37-amino acid carboxyl-terminal peptide was therefore regarded as the chicken equivalent to GLP-I. The carboxyl-terminal amino acid of chicken GLP-I was Gly, indicating that the GLP-I can be α-amidated on its carboxyl terminus as is the case with mammalian GLP-I [22]. The degree of amino acid sequence homology of other domains (signal peptide, NH_2-peptide and intervening peptide) varied from 80 to 27% between chicken on the one hand and anglerfish, rat and human on the other. The most interesting feature of chicken preproglucagon is its lacking domains corresponding to intervening peptide II and GLP-II which are present in mammalian preproglucagons, and its overall organization, which is rather similar to that of anglerfish preproglucagon. The high

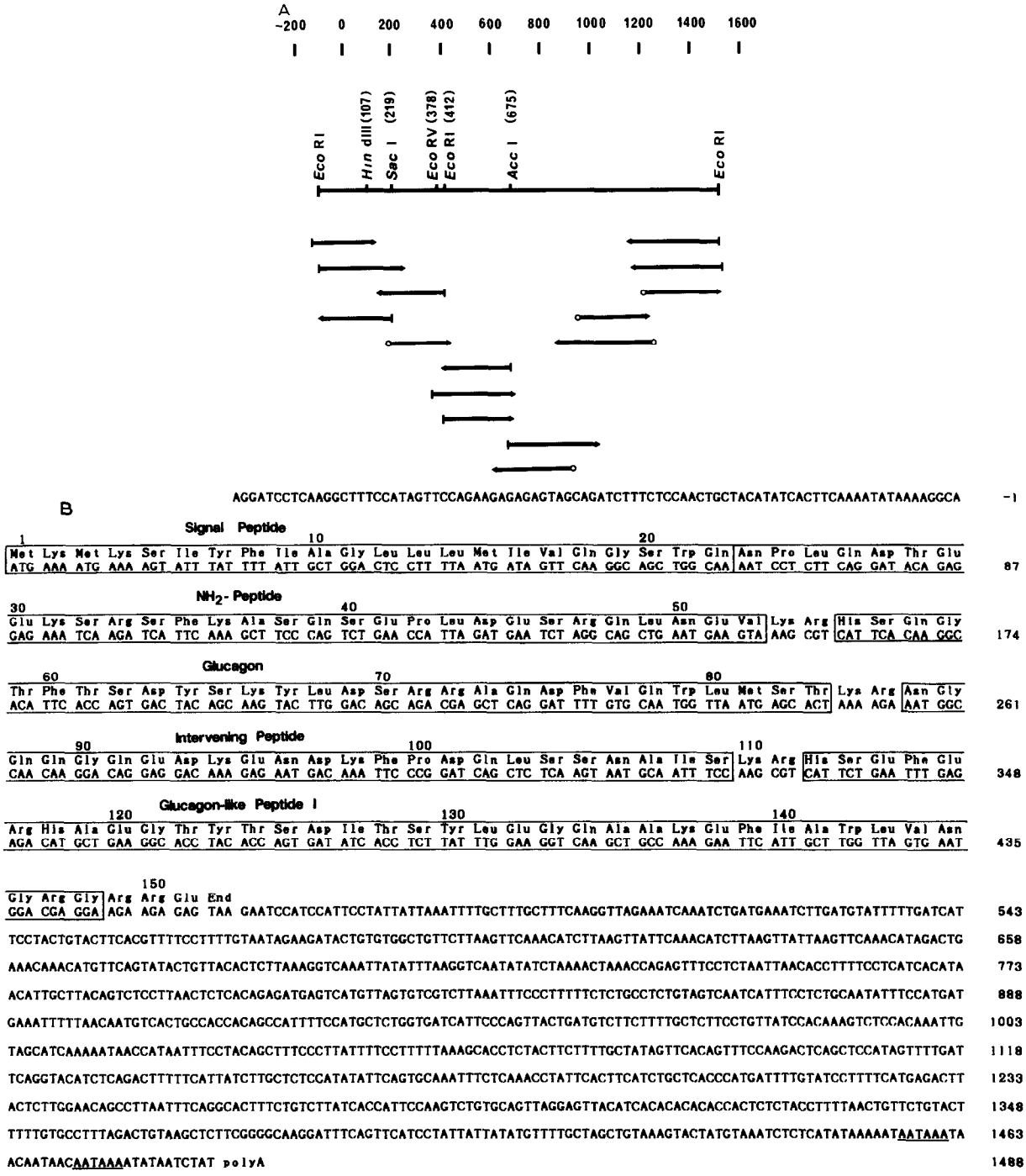

Fig. 2. (A) Restriction map and sequence strategy for cDNA encoding chicken preproglucagon. Nucleotide numbers are given at the top. Restriction sites used for subcloning and sequencing are demarcated, and numbers of the nucleotide immediately upstream of the site of cleavage are indicated in parentheses. Sequencing strategy is shown below the restriction map. Overlapping sequences were determined using commercial primers or synthetic primers (open circles). Arrows indicate the direction and extent of sequence determination. (B) Nucleotide and deduced amino acid sequences of chicken preproglucagon. Nucleotide residues are numbered in the 5′ to 3′ direction, beginning with the first residue of the ATG triplet encoding the initiator methionine, and the nucleotides on the 5′ side of residue 1 are indicated by negative numbers. Deduced amino acids are given above the nucleotide sequence and numbered beginning with the initiator methionine. Polyadenylation signals are underlined. Sequences of signal peptide, NH_2-peptide, glucagon, intervening peptide and glucagon-like peptide I are boxed.

degree of conservation of glucagon and GLP-I sequences among the avian, piscine and mammalian species suggests that the two peptides have essential biological functions. Glucagon is an important regulator of carbohydrate, protein and lipid metabolism [2,3,23]. GLP-I has been shown to be a potent insulin secretagogue [24]. The lack of GLP-II in the avian and piscine precursors suggests that this peptide has no

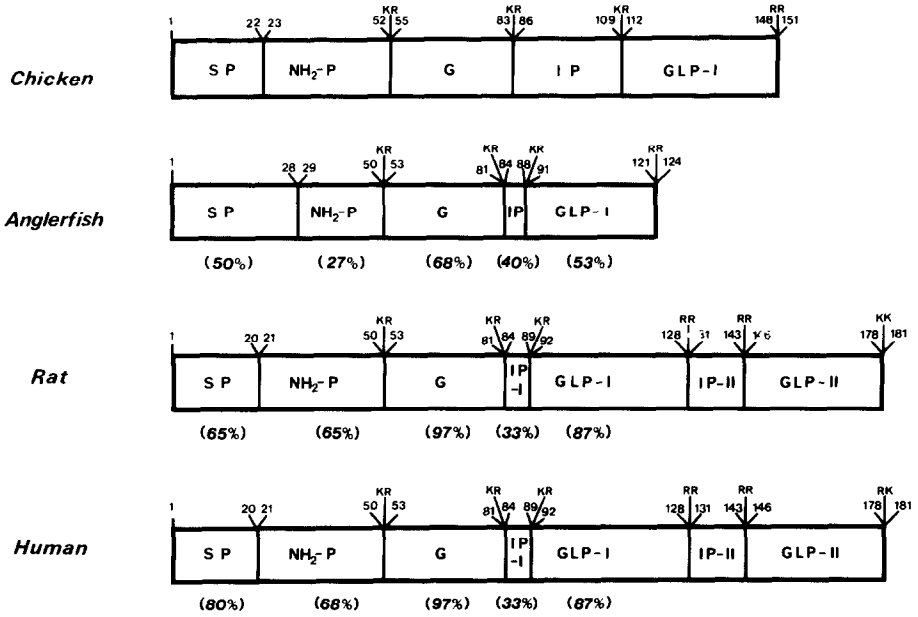

Fig. 3. Organization of functionally divided domains in chicken, anglerfish and mammalian preproglucagons. Putative proteolytic processing sites occurring at two adjacent basic residues are indicated by the single-letter code for lysine (K) and arginine (R). Numbers of the amino- and carboxyl-terminal amino acid residues of each domain are given above schematic diagrams. SP, signal peptide; NH_2-P, NH_2-peptide; G, glucagon; IP, intervening peptide; GLP-I and -II, glucagon-like peptide I and II; IP-I and -II, intervening peptide I and II. Numbers in parentheses below schematic diagrams indicate extent of homology in amino acid sequence to the chicken. Data for anglerfish (preproglucagon I), rat and human preproglucagons are from [10], [5] and [6], respectively.

essential role in mammals. In fact, there have so far been no reports concerning the biological activity of GLP-II. The physiological significance of NH_2-peptide and the intervening peptides is also unknown.

Acknowledgements: This work has been supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan.

REFERENCES

[1] Habener, J.F. (1990) in: Molecular Biology of the Islets of Langerhans (Okamoto, H. ed.) pp. 67–85, Cambridge University Press, Cambridge.
[2] Cherrington, A.D., Chiasson, J.L., Liljenquist, J.E., Jennings, A.S., Keller, U. and Lacy, W.W. (1976) J. Clin. Invest. 58, 1407–1418.
[3] Aoki, T.T., Müller, W.A., Brennan, M.F. and Cahill, G.F., Jr. (1974) Metabolism 23, 805–814.
[4] Samols, E., Marri, G. and Marks, V. (1965) Lancet 2, 415–416.
[5] Heinrich, G., Gros, P., Lund, P.K., Bentley, R.C. and Habener, J.F. (1984) Endocrinology 115, 2176–2181.
[6] Drucker, D.J. and Asa, S. (1988) J. Biol. Chem. 263, 13475–13478.
[7] Bell, G.I., Santerre, R.F. and Mullenbach, G.T. (1983) Nature 302, 716–718.
[8] Seino, S., Welsh, M., Bell, G.I., Chan, S.J. and Steiner, D.F. (1986) FEBS Lett. 203, 25–30.
[9] Lopez, L.C., Frazier, M.L., Su, C.-J., Kumar, A. and Saunders, G.F. (1983) Proc. Natl. Acad. Sci. USA 80, 5485–5489.
[10] Lund, P.K., Goodman, R.H., Dee, P.C. and Habener, J.F. (1982) Proc. Natl. Acad. Sci. USA 79, 345–349.
[11] Lund, P.K., Goodman, R.H., Montminy, M.R., Dee, P.C. and Habener, J.F. (1983) J. Biol. Chem. 258, 3280–3284.
[12] Pollock, H.G. and Kimmel, J.R. (1975) J. Biol. Chem. 250, 9377–9380.
[13] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[14] Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408–1412.
[15] Terazono, K., Yamamoto, H., Takasawa, S., Shiga, K., Yonemura, Y., Tochino, Y. and Okamoto, H. (1988) J. Biol. Chem. 263, 2111–2114.
[16] Yamamoto, H., Nata, K. and Okamoto, H. (1986) J. Biol. Chem. 261, 6156–6159.
[17] Webb, R., Reddy, K.J. and Sherman, L.A. (1989) DNA 8, 69–73.
[18] Yonekura, H., Nata, K., Watanabe, T., Kurashina, Y., Yamamoto, H. and Okamoto, H. (1988) J. Biol. Chem. 263, 2990–2997.
[19] Inoue, C., Shiga, K., Takasawa, S., Kitagawa, M., Yamamoto, H. and Okamoto, H. (1987) Proc. Natl. Acad. Sci. USA 84, 6659–6662.
[20] Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13.
[21] Von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4690.
[22] Kreymann, B., Yiangou, Y., Kanse, S., Williams, G., Ghatei, M.A. and Bloom, S.R. (1988) FEBS Lett. 242, 167–170.
[23] Goodridge, A.G. (1968) Am. J. Physiol. 214, 902–907.
[24] Schmidt, W.E., Siegel, E.G. and Creutzfeldt, W. (1985) Diabetologia 28, 704–707.