Postprint version. Original publication in:
Molecular Biology (USSR) (1988) 22, 1, 122 - 127
Molekulyarnaya Biologiya (1988) 22, 1, 145 - 150

GENETIC ENGINEERING OF PEPTIDE HORMONES. III. CLONING OF cDNA OF PORCINE GROWTH HORMONE AND CONSTRUCTION OF GENE FOR EXPRESSION OF HORMONE IN BACTERIA

G. S. Zhvirblis, V. G. Gorbulev, P. M. Rubtasov, B. K. Chernov, Yu. B. Golova, G. E. Posmogova, K. G. Skryabin, and A. A. Baev

UDC 577.219.175.3

Results are presented of cloning cDNA of procine growth hormone, analysis of its primary structure, and creation of a construction capable of expression of this cDNA in *Escherichia coli* cells. It is shown that in the population of mRNA coding porcine growth hormone, heterogeneity is noted which is manifested not only at the level of the nucleotide sequence, but also is reflected in the amino acid sequence of the mature hormone.

Growth hormone, or somatotropin, belongs to a numerous group of peptide hormones synthesized in the adenohypophysis of vertebrates. It is initially synthesized in the form of a prehormone containing a leader (or signal) sequence which is detached in the process of secretion of the polypeptide into blood.

Growth hormone displays species specificity, although in a number of cases, somatotropin of one species manifests biological activity in relation to several species standing at a lower level in the evolutionary order.

At the cellular level the mechanism of effect of growth hormone has been studied in inadequate detail. At the level of the whole organism, somatotropin manifests a pleiotropic effect expressed in accelerated growth of the skeleton, increase in weight of animals computed on the same consumption of food, and change in metabolism of glucose and lipids [1]. Because of this physiological effect, growth hormone is important for intensification of agriculture.

Advances in genetic engineering now permit proteins previously not easily accessible to be obtained in large quantities using microorganisms. The first step along the path toward biotechnological production of porcine growth hormone is construction of the gene coding this protein and suitable for expression in cells of microorganisms. Results are presented in this article of the cloning and determination of primary structure of cDNA of porcine growth hormone and the construction of such a gene on its basis.

EXPERIMENTAL

Matrix RNA was isolated from swine hypophyses frozen in liquid nitrogen according to Chirgwin et al. [2], using guanidine thiocyanate and subsequent centrifugation through a CsCl carrier. Residues of RNA were suspended in guanidine hydrochloride, washed with ethanol, and then reprecipitated twice with ethanol from 0.3 M sodium acetate. Yield of total RNA constitute 10-30 mg from 5-10 g of tissue.

The RNA containing poly(A) was separated from other RNA by chromatography on oligo(dT)-cellulose (Sigma).

Synthesis of cDNA and cloning was conducted according to a method described previously [3]. Revertase was obtained from V. Kavsan (Institute of Molecular Biology and Genetics, Academy of Sciences of the Ukrainian SSR, Kiev). An inhibitor of ribonuclease, RNAazin (PL Pharmacia), was used in synthesis of the first chain.

Institute of Molecular Biology, Academy of Sciences of the USSR, Moscow. Translated from Molekulyarnaya Biologiya, Vol. 22, No. 1, pp. 145-150, January-February, 1988.

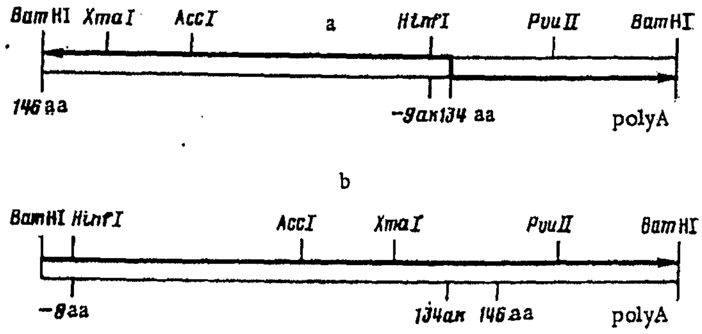

Fig. 1. Restrictase maps of insertion from plasmid pPGH2 (a) and cDNA of porcine growth hormone according to Seeburg, et al. [10] (b). Arrows designate direction of coding chain of cDNA from 5'- to 3'-end. Localization of codons corresponding to certain amino acids (aa) are given in numbers designating the position of the given amino acid in porcine growth hormone. Numbering proceeds from the first amino acid of mature protein; amino acids of the signal peptide are noted with negative numbers.

Strain HB101 of *Escherichia coli* was used for transformation. Colonies were hybridized with a ^{32}P-labeled probe according to a known method [4].

The *E. coli* DNA-polymerase I, large fragment of DNA-polymerase I, and DNA-ligase of phage T4 were obtained from Amercham; HindIII, EcoRI, BamHI, Sau3AI, XmaI, and Eco471 restrictases were obtained from Ferment non-governmental organization (Vil'nyus), ApaI restrictase from Biolaboratories, AccI restrictase from Amersham, and S1 nuclease from Sigma.

Plasmic DNA was cleaved in buffer containing 10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.5, 1 mM dithiothreitol, and NaCl, the concentration of which varied from 0 to 150 mM depending on restrictase used.

Nucleotide sequences of DNA were determined according to Maxam and Gilbert [5] and Sanger et al. [6], as modified by Hattori and Sakaki [7]. Oligonucleotides for the 5'-terminal part of porcine growth hormone were synthesized chemically [8].

Treatment of oligonucleotides with kinase and ligase was conducted in the following way: 5 μg of each preparation were incubated with kinase (4 units) 15 min at 37°C in buffer containing 50 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 5 mM dithiothreitol[γ-^{32}P]ATP with a radioactivity of 5 μCi. Then, "cold" ATP was added to a final concentration of 0.5 mM and incubated another 30 min at 37°C. Ligation of oligonucleotides was conducted for 12 h at 4°C in the same buffer with DNA-ligase of phage T4 (50 units). Ligase mixture was treated with EcoRI and ApaI restrictases and the synthetic fragment was separated from 8% polyacrylamide gel.

RESULTS AND DISCUSSION

The matrix for synthesis of cDNA used was RNA from swine adenohypophysis purified on a column with oligo(dT)-cellulose and containing poly(A). Double-stranded cDNA, after treatment with S1 nuclease and attachment of HindIII or BamHI linkers, was fractionated by electrophoresis in 5% polyacrylamide gel. Fragments 500-900 basepairs (b.p.) in size were eluted and inserted into plasmid pBR322 cleaved at the appropriate section. After transformation in strain HB101, recombinant clones were selected according to sensitivity to tetracycline.

Clones containing cDNA of porcine growth hormone were chosen by hybridization of colonies, using cloned cDNA of bovine growth hormone as the probe [9].

As a result, two clones were chosen: pPGH1, carrying an insertion 500 b.p. in size, and pPGH2, with an insertion of 760 b.p.

Analysis of the nucleotide sequence of cDNA from clone pPGH1 and comparison of it with the structure of porcine growth hormone cDNA published previously [10] showed that the inser-

```
      -11                               -6
——G C T C T G C C T G C C C T G G A C T ——
  :   : : : : : : : : : : :   : :   : :
——C T A G A C G G A C G G G C C C C G A ——
  136       A                 131
```

Fig. 2. Self-complementary section of cDNA of porcine growth hormone. Numbering of codons is the same as in Fig. 1.

```
          10                20              30              40             50
ACT CAG GAG GTG GGA GCC TTC CCA GCC ATG CCC TTG TCC AGC CTA TTT GCC AAC
thr gln glu val gly ala phe pro ala met pro leu ser ser leu phe ala asn
     -5                 -1   1               5                   10
     60              70             80             90             100
GCC GTG CTC CGG GCC CAG CAC CTG CAC CAA CTG GCT GCC GAC ACC TAC AAG GAG
ala val leu arg ala gln his leu his gln leu ala ala asp thr tyr lys glu
         15                  20                      25                  30
110          120              130            140            150         160
TTT GAG CGC GCC TAC ATC CCG GAG GGA CAG AGG TAC TCC ATC CAG AAC GCC CAG
phe glu arg ala tyr ile pro glu gly gln arg tyr ser ile gln asn ala gln
                     35                  40                  45
        170             180             190            200            210
GCT GCC TTC TGC TTC TCG GAG ACC ATC CCG GCC CCC ACG GGC AAG GAC GAG GCC
ala ala phe cys phe ser glu thr ile pro ala pro thr gly lys asp glu ala
    50                  55                  60                  65
    220          230            240            250            260         270
CAG CAG AGA TCG GAC GTG GAG CTG CTG CGC TTC TCG CTG CTG CTC ATC CAG TCG
gln glu arg ser asp val glu leu leu arg phe ser leu leu leu ile gln ser
            70                  75                  80
           280          290            300            310          320
TGG CTC GGG CCC GTG CAG TTC CTC AGC AGG GTC TTC ACC AAC AGC CTG GTG TTT
trp leu gly pro val gln phe leu ser arg val phe thr asn ser leu val phe
85                      90                  95                  100
     330           340            350            360           370
GGC ACC TCA GAC CGC GTC TAC GAG AAG CTG AAG GAC CTG GAG GAG GGC ATC CAG
gly thr ser asp arg val tyr glu lys leu lys asp leu glu glu gly ile glu
        105                 110                 115                 120
380          390            400            410            420            430
GCC CTG ATG CGG GAG CTG GAG GAT GGC AGC CCC CGG GCA GGA CAG ACC CTC AAG
ala leu met arg glu leu glu asp gly ser pro arg ala gly gln thr leu lys
                125                 130                 135
        440          450            460           470            480
CAA ACC TAC GAC AAA TTT GAC ACA AAC TTG CGC AGT GAT GAC GCG CTG CTT AAG
gln thr tyr asp lys phe asp thr asp leu arg ser asp asp ala leu leu lys
    140                 145                 150                 155
    490          500            510            520           530          540
AAC TAC GGG CTG CTC TCC TGC TTC AAG AAG GAC CTG CAC AAG GCT GAG ACA TAC
asn tyr gly leu leu ser cys phe lys lys asp leu his lys ala glu thr tyr
            160                 165                 170
            550          560            570            580          590
CTG CGG GTC ATG AAG TGT CGC CGC TTC GTG GAG AGC AGC TGT GCC TTC TAGTTGC
leu arg val met lys cys arg arg phe val glu ser ser cys ala phe  *
175                 180                 185                 190
    600       610       620       630       640       650       660
TGGGCATCTCTGTTGCCCCTCCCAGTACCTCCCCTGACCCTGGAAAGTGCCACCCCAATGCCTGCTGTCC
   670       680       690       700       710
TTTCCTAATAAAACCAGGTTGCATTGTAAAAAAAAAAAAAAAAAA
```

Fig. 3. Reconstructed nucleotide and amino acid sequence of cDNA of porcine growth hormone.

tion is an incomplete copy of the mRNA of porcine somatotropin and begins from the codon corresponding to the 66th amino acid of the mature protein. The insertion contains two-thirds of the coding region of growth hormone mRNA, the entire 3'-untranslated region (105 b.p.), and terminates with a sequence of poly(A) (18 b.p.).

When sequencing the insertion from clone pPGH2, it was found that it contains the entire cDNA sequence of mature growth hormone and, in addition, nine codons of the signal peptide. Further, it turned out that serious reorganizations occurred in cDNA from clone pPGH2 in the process of cloning which disrupted the structure of the cDNA coding section (Fig. 1).

Analysis of the nucleotide sequence in the region located next to the section in which reorganization occurred revealed a very extended zone of complementation between the 5'-terminal sequence of mRNA (codons from the −11th to the −6th amino acid) and middle section of mRNA (codons of amino acids 130–136) of porcine growth hormone (Fig. 2). Such strong

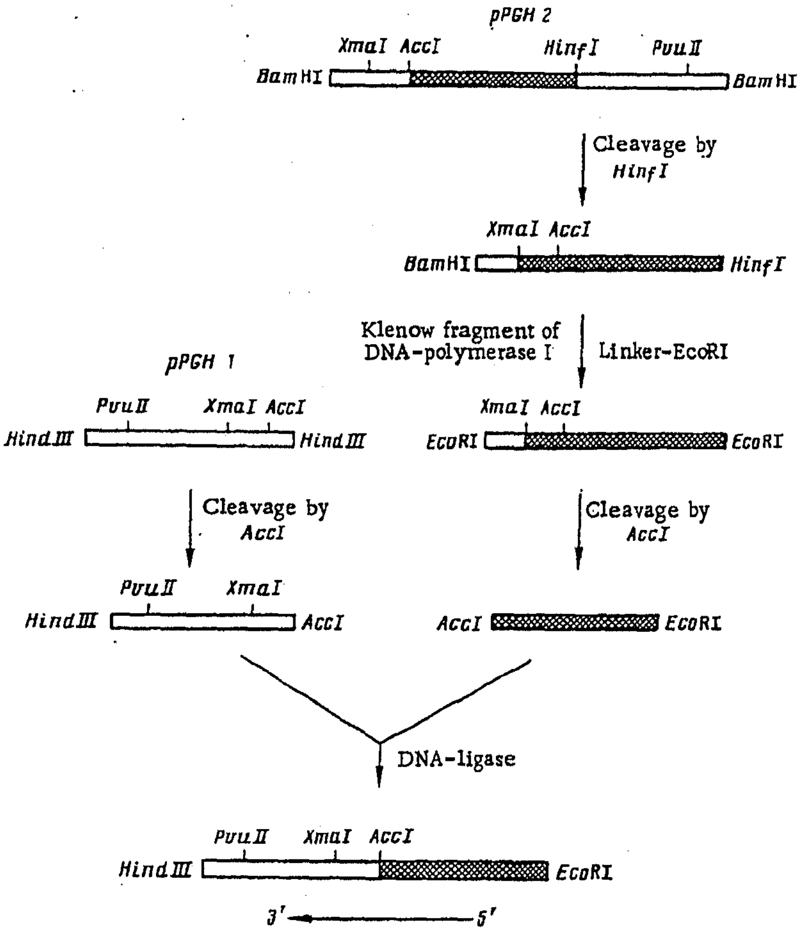

Fig. 4. Plan for reconstruction of complete and accurate sequence of cDNA coding mature porcine somatotropin with six codons corresponding to section of signal peptide. Lower horizontal arrow indicates direction of coding cDNA chain.

TABLE 1. Comparison of cDNA of Porcine Growth Prehormone Obtained by Different Authors

Site of difference	Variants of cDNA of porcine growth prehormone			
	Seeburg et al. [10]	Movva and Schulz [11]	pPGH1	pPGH2
—2a. c.	GGA	GGC	—	GGA
25 »	GCC	GCT	—	GCC
48 »	CAG	CAA	—	CAG
56 »	ACC	ACG	—	ACC
100 »	CTG	TTG	CTG	CTG
127 »	GAA	GAG	GAG	GAA
134 »	GGA	GGC	GGA	GGA
136 »	ATC	ATC	ACC	ATC
190 »	TTC	TCC	TTC	TTC
44—51 n.	GGTGCCCT	GACCT	GACCCT	GACCCT
83—87 »	TCCCT	TCCT	TCCT	TCCT
94—97 »	ACCA	ACA	ACCA	ACCA
105—107 »	TCG	TCG	TTG	TTG

*Sites of differences in coding part of prehormone cDNA are designated by number of codon (c.) of corresponding amino acid. The 3'-terminal untranslated section is numbered from the first nucleotide (n.) of the terminal codon according to the variant of Seeburg et al. [10].

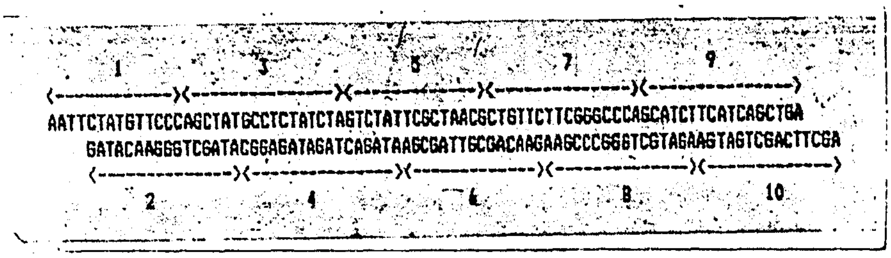

Fig. 5. ... and oligonucleotides from which the synthetic fragment was collected during ligation are indicated by numbers [sic; part of caption missing in Russian original].

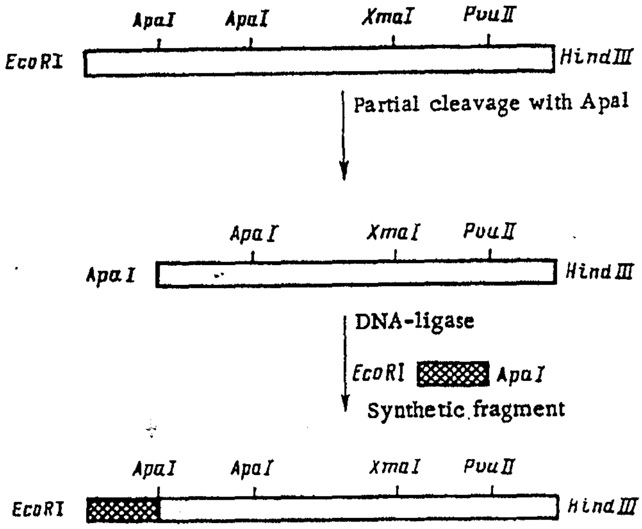

Fig. 6. Plan for construction of cDNA of porcine growth hormone with replacement of natural 5'-terminal region with synthetic.

complementation (16 b.p. out of 20) apparently led to formation of hairpins during synthesis of cDNA and ultimately caused reorganization of cloned cDNA.

Comparison of nucleotide sequences of insertions of cDNA of clones pPGH1 and pPGH2 with each other and with results published previously [10, 11] showed that the cDNA of pPGH1 and pPGH2 differ not only from the published cDNA, but also from each other (Table 1). Substitutions affect both the untranslated region and the coding part of cDNA. One such substitution, identified in pPGH1 in the codon corresponding to the 136th amino acid of mature somatotropin, leads to disappearance of one Sau3AI section (absence of Sau3AI section was demonstrated by restrictase analysis and during analysis of the sequence of this section). This substitution moreover changes the meaning of the codon: instead of isoleucine, threonine appears in the protein. This substitution is absent in cDNA from clone pPGH2. It follows from this that in the case of porcine growth hormone, polymorphism exists not only at the level of mRNA, but also at the level of the protein molecule.

Using the presence of the unique AccI section in cDNA from clones pPGH1 and pPGH2, we reconstructed the complete and accurate sequence of cDNA coding mature porcine somatotropin (Fig. 3). The plan for reconstruction is presented in Fig. 4.

It is known from data presented in the literature [10] that cDNA of porcine growth hormone, which contains a natural 5'-terminal part of the coding region of mRNA is expressed in *E. coli* with low effectiveness. Therefore, we replaced the region of cDNA coding the first 15 amino acids of mature protein with a synthetic region (Fig. 5), not changing the amino acid sequence of the hormone and taking into account optimal utilization of codons in *E. coli* and supplied by restrictase cleavage sites required for further work in inserting this cDNA into plasmid with a strong promoter. The plan for construction of cDNA carrying the synthetic 5'-terminal region and natural 3'-terminal region is shown in Fig. 6.

Hybrid cDNA obtained in this way was cloned in pUC19, and the correctness of its nucleotide structure was tested by determining the nucleotide sequence using a modification of the method of Sanger and a double-stranded matrix [6].

The authors are grateful to A. A. Shul'ge (Institute of Molecular Biology, Academy of Sciences of the USSR) for providing the methods for determining sequence of double-stranded DNA according to Singer using labeled primers.

LITERATURE CITED

1. C. S. Chung, T. D. Etherton, and J. P. Wiggins, J. Anim. Sci., 60, 118-130 (1985).
2. J.M. Chirgwin, A. E. Przybyla, R. J. Macdonald, and W. J. Rutter, Biochemistry, 18, 5294-5299 (1979).
3. T. Maniatis, E. F. Fritsch, and J. Sambrook, Molecular Cloning: Methods of Genetic Engineering [Russian translation], Mir, Moscow (1984), pp. 205-240.
4. M. Grunstein and D. S. Hogness, Proc. Natl. Acad. Sci. USA, 72, 3961-3965 (1975).
5. A. M. Maxam and W. Gilbert, Proc. Natl. Acad. Sci. USA, 74, 560-564 (1977).
6. F. Sanger, S. Bicklen, and A. R. Coulson, ibid, 5463-5467 (1977).
7. M. Hattori and Y. Sakaki, Anal. Biochem., 152, 232-238 (1986).
8. B. K. Chernov, et al., Dokl. Akad. Nauk SSSR, 291, 1131-1134 (1986).
9. V. G. Gorbulev, P. M. Rubtsov, and K. G. Skryabin, Proceedings of the Fifth International Symposium on Metabolism and Enzymology of Nucleic Acids Including Gene Manipulations, Bratislava (1984), p. 233.
10. P. H. Seeburg, S. Sias, J. Adelman, H. A. de Boer, J. Hayflick, P. Jhurani, V. D. Goeddel, and H. L. Heyneker, DNA, 2, 37-45 (1983).
11. N. Movva and M.-F. Schulz, Patent EP 0 104 920 A1 (1984).