Cloning and sequencing of bullfrog growth hormone complementary DNA

Fu-Ming Pan and Wen-Chang Chang

Institute of Biochemical Sciences, National Taiwan University, and Institute of Biological Chemistry, Academia Sinica, Taipei (Taiwan, China)

(Received 15 December 1987)

(Revised manuscript received 18 March 1988)

Key words: Growth hormone; cDNA sequence; Cloning; Amino acid sequence; Sequence homology; (Bullfrog)

Total mRNA was isolated from the pituitary glands of bullfrog (*Rana catesbeiana*), purified by affinity chromatography with oligo(dT)-cellulose columns. The cDNA was synthesized and cloned in *Escherichia coli* using *EcoRI* linkers and pBR322 as vector. The cDNA library was screened by hybridization with ³²P-labeled duck growth hormone (GH) cDNA. A positive clone was selected and sequenced. The full-length bullfrog GH cDNA contains 950 nucleotide pairs with an open reading frame coding for the precursor GH of 215 amino-acid residues. The partial amino-acid sequence from the protein confirms that derived from the cDNA, with Phe as the first residue in the mature bullfrog GH preceded by a 25-residue hydrophobic signal peptide. The bullfrog GH shares sequence homology with those of other vertebrate species in the following order: duck (61% protein sequence homology; 67% cDNA homology), rat (56%; 61%), human (47%; 57%) and salmon (42%; 50%).

The growth hormone (GH) of the vertebrate is synthesized and secreted by the anterior pituitary gland and required for preadult growth of the vertebrate [1,2]. It has been extensively investigated both at protein and gene levels in human and other mammals. These studies have been extended to the lower vertebrates, such as avian [3–5] and fish [6–8]. For the bullfrog, a representative of the amphibia, the biosynthesis and gene expression of GH might be more complicated and interesting because it experiences a stage of metamorphosis in its life cycle.

The bullfrog GH has been purified and characterized as follows. Fresh pituitary glands of

bullfrog (*Rana catesbeiana*) were obtained from a local supplier. They were homogenized, and separated on a Sephacryl S-200 column, followed by ion-exchange chromatography on a DEAE-cellulose column and identified as described before [9]. The purified bullfrog GH was sequenced in a pulsed-liquid protein sequencer (Applied Biosystems, Inc. 477A).

The mRNA was purified and cloned as described previously [5]. Briefly, total RNA was extracted from the fresh pituitary glands by guanidinium isothiocyanate and hot phenol. Polyadenylated RNA was selected by passing through the oligo(dT)-cellulose column twice and used for the synthesis of cDNA [10]. The double-stranded cDNA was methylated with *EcoRI* methylase before ligation with *EcoRI* linkers and then digested with *EcoRI* to generate the cohesive ends for insertion into the *EcoRI*-treated pBR322 vector.

The cDNA library was screened by hybridization with ³²P-labeled duck GH-cDNA [5]. A posi-

Abbreviation: GH, growth hormone.

Correspondence: W.C. Chang, Institute of Biological Chemistry, Academia Sinica, P.O. Box 23-106, Taipei, Taiwan, Republic of China.

A

B

Fig. 1. Orientation and enzyme map of the bullfrog GH-cDNA in plasmid pBR322 (A) and sequencing strategy (B). The solid circles indicate the starting point of dideoxy sequencing, whereas the open ones are for the end-labeling sites for chemical sequencing.

tive clone containing the bullfrog GH-cDNA was selected and sequenced by the chemical method of Maxam and Gilbert [11] and the chain-termination method of Sanger [12].

By using the duck GH-cDNA as probe, we were able to select a clone containing the full-sized bull-frog cDNA which was then sequenced by the chemical and dideoxy method as shown in Fig. 1. It is evident that this bullfrog GH-cDNA contains an open reading frame encoding the precursor GH of 215 amino-acid residues. The N-terminal amino-acid sequence of 30 residues derived from the purified bullfrog GH indicates that the mature hormone starts with a Phe in its N-terminus and completely agrees with the sequence derived from the nucleotide sequence (Fig. 2). It can be seen that the precursor GH contains a signal peptide of 25 residues followed by 190 residues in the mature

hormone. When compared with several known sequences of GHs from representative vertebrate species (Fig. 3), the bullfrog GH shows significant homology with other GHs, in accordance with the

TABLE I

SEQUENCE HOMOLOGY BETWEEN BULLFROG GROWTH HORMONE AND GROWTH HORMONES OF OTHER VERTEBRATE SPECIES

GH	Protein ^a (%)	cDNA ^b (%)
Human [14]	47	57
Rat [15]	56	61
Duck [5]	61	67
Salmon [7]	42	50

^a The precursor sequences are compared.

^b Only the coding sequences for precursor GH are compared.

50
 GGCGGAACCCCAATATACATCAGGAACACTGCTAACCTCAATCAATGGCTTCAGGGTTAG
 M A S G L G

100
 GCTCATCTCTTGACTCTTGTTGTAATATGTCTTCAAAGCCCGCAAGGGTTCAATGCCT
 S S L V L L V V I C L Q S P Q G F N A F

150
 TTCCACAGATGTCTCTTTCCAACCTCTTTACAAATGCTGTCTTCGAGCTCAGCACCTTC
 P Q M S L S N L F T N A V I R A Q H L H

200
 ACCAGATGGTTGCCGATACCTACAGAGACTATGAAAGAACATATATACCTGAAGATCAGA
 Q M V A D T Y R D Y E R T Y I P E D Q R

250 300
 GATTCAAACAAACACTTCTTATTTTCAGTGTATTGCTACTCAGAAACCATCCAGCACCAA
 F K Q T L L I S V Y C Y S E T I P A P T

350
 CAGACAAAGACAATACTCACCAGAAATCAGATATAGACCTCCTCCGCTTCTCCCTCACAC
 D K D N T H Q K S D I D L L R F S L T L

400
 TCCTGCAGTCATGGATGACCCCTATCCAAATCGTCAACAGAGTTTTTGGCAACAACCAGG
 L Q S W M T P I Q I V N R V F G N N Q V

450
 TGTTTGGAACATTGACAGAGTTTATGACAGACTGAGGGACCTCGATGAAGGACTACACA
 F G N I D R V Y D R L R D L D E G L H I

500
 TTCTCATAAGAGAGCTGGATGATGGAATGTTTCGGAACCTACGGTGTGCTGACTTTTACAT
 L I R E L D D G N V R N Y G V L T F T Y

550 600
 ATGATAAGTTTGATGTCAATCTGCGCAGTGAAGAAGGCCGGCGAAGAATTACGGTCTAC
 D K F D V N L R S E E G R A K N Y G L L

650
 TCTCCTGCTTCAAGAAAGACATGCATAAAGTAGAGACCTACCTGAAGGTGATGAAATGCC
 S C F K K D M H K V E T Y L K V M K C R

700
 GGAGATTTGTGGAGAGCAACTGCACCTTTTAATCAAGAAAAAATGTTACATCCTCCATA
 R F V E S N C T F

750
 GTCACCCCCACCTCCTGCAGGACACCCGCTGCATTTTTATATCTTCAGGAGGTAGGGCCC

800
 ACTGATAATTGTCCCTTTTAACTTCCTGACCAATTGCACATTTTTTGCAGATGTATTTT

850 900
 GCGTTGTGCTTTTGTGTCTTTGTATGTTGTTAAAAAATAAAGTATAGGAGGCATTTC

ATGA₄₆

Fig. 2. The nucleotide sequence of the bullfrog GH-cDNA with the derived amino-acid sequence

				*					**	*	*	**	
BULLFROG	MASGLGSSL--VLLVVI--CLQSPQGFNAFFQMSLSNLFNAVIRAQHLHQMVDITYR												
HUMAN	::T:SRT::--L:AFGL-L::PWL:EGS:::TIP::R::D::ML::HR:::LAF:::Q												
RAT	::ADSQTPW--L:TFSL-L::LW::EAG:L:A:P::S::A:::L:::LA:::K												
DUCK	::P:SWF:P--LFIT::TLG::W::EAAT::A:P:::A:::L:::M-E:::V:I::												
SALMON	M:QVFL:MP:::S---:FL:::AAIEN:R---:NI::S:V:::LLAQKMFN												
		*			*	*	*	*	*		*	*	*
BULLFROG	DYERTYIPEDQRF--KQTLISVYCYSETIPAPTDKDNTHQKSDI-DLLRFSLTLLQS												
HUMAN	EF:EA:::KE:KYSFL:NPQT:L:-F::S::T:SNREE:Q:::NL-E:::I::L:I::												
RAT	EF::A:::G::YS-I:NAQAAF-F:::G::EEAQ:RT:M-E:::L:I::												
DUCK	EF::S:::HTN:-NSQ-AF-:::G::DAQ::M-E:::V:I::												
SALMON	:FDG:LL:DER:QLN:IF::DF--:N:DS:VS:V:HE:::SVLK:HI:FR:IE:												
		*	*	*				*	*	*	*	*	*
BULLFROG	WMTPIQ---IVNR--VFGNNQVFGNIDR-VYDRLRDLDEGLHILIRELDDGNVR----												
HUMAN	:LE:V:---FLRS---:A:SL:Y:AS:SN:::L:K::E::IQT:MGR:E::SP:----												
RAT	:LG:V:---FLS--I:T:SLM::TS:::EK:K::E::IQA:MQ::E::SP:----												
DUCK	:L:V:---YLSK--:T:L::TS:::FEK:K::E::IQA:M::E::RSP:----												
SALMON	:EY:S:TLI:S:SLM:RNA::---:SE---K:S::KV:INL::TGSQ:::LSLDD												
		*			*	*	*	*	*	*	*	*	*
BULLFROG	NYGVLTFTYDKFDVNLNRSEEGRAKNYGLLSCFKKDMHKVETYLKVMKCRRFVESNCTF												
HUMAN	TGQIFKQ::S:::T:SHNDALL:::Y::R::D:::F:RIVQ::S::GS:G:												
RAT	IGQI:KQ:::A:M::DDALL:::L:A:::R:::A::S:A:												
DUCK	GPQL:KP:::IH:N:DALL:::L:::G:::I												
SALMON	:DSQQLPP:GNYYQ::GGDGNVRR::E:A:::T:A::KSL:A::L												

Fig. 3. Comparison of amino-acid sequence of GH precursors of bullfrog (the present study), human [14], rat [15], duck [5] and salmon [7]. The colons indicate identical residues. Gaps are introduced to maximize homology. Asterisks indicate the invariant residues for all sequences.

general concept that the GH is a well-conserved protein hormone in vertebrates. It also shows that 44 residues in the GH molecule are invariant through vertebrate evolution, suggesting that these residues may play important roles in maintaining the correct structure or conformation. Among these invariant residues, the four cysteine residues are essential for the formation of two disulfide bridges; whereas the single Trp-110 (numbering starting from the Met of bullfrog GH precursor) has been implied to form a hydrogen bond with a carboxylate ion in the hydrophobic interior of the GH molecule [13].

Comparison of amino-acid and nucleotide sequences between bullfrog GH and GHs from other species is shown in Table I. It is interesting that the bullfrog GH sequence is closer to duck GH than to fish GH. This result would shed some light on the phylogenetic and evolutionary relationship of the hormone.

Several mammalian GH genes have been reported and the 5'-untranslated sequences com-

pared [16]. These genomic sequences indicate that the 5'-untranslated region consists of around 60 base-pairs. It is obvious that this 5'-untranslated region has never been fully reverse transcribed from mRNA into cDNA. The bullfrog GH-cDNA contains 45 base-pairs in this region (Fig. 2), presumably incomplete too, but it certainly provides the full-length coding sequence for the precursor GH.

It is also of interest to note that the translation initiator site of bullfrog GH-cDNA contains a T nucleotide at the -3 position instead of a purine as predicted by the so-called Kozak rule [17]. All the other known GH-cDNAs conform to the Kozak rule for an optimal translation initiation sequence.

Acknowledgement

This work was financially supported by the National Science Council of Republic of China.

References

- 1 Miller, W.L. and Eberhardt, N.L. (1983) *Endocrine Rev.* 4, 97–130.
- 2 Chawla, P.K., Parks, J.S. and Rudman, D. (1983) *Annu. Rev. Med.* 34, 519–547.
- 3 Souza, L.M., Boone, T.C., Murdock, D., langley, K., Wypych, J., Fenton, D., Johnson, S., Lai, P.H., Everett, R., Hsu, R.Y. and Bosselman, R. (1984) *J. Exp. Zool.* 232, 465–473.
- 4 Farmer, S.W., Papkoff, H. and Hayashida, T. (1974) *Endocrinology* 95, 1560–1565.
- 5 Chen, H.T., Pan, F.M. and Chang, W.C. (1987) *Biochim. Biophys. Acta* 949, 249–253.
- 6 Farmer, S.W., Hayashida, T., Papkoff, H. and Polenov, A.L. (1981) *Endocrinology* 108, 377–381.
- 7 Sekine, S., Mizukami, T., Nishi, T., Kuwana, Y., Saito, A., sato, M., Itoh, S. and Kawauchi, H. (1985) *Proc. Natl. Acad. Sci. USA* 82, 4306–4310.
- 8 Agellon, L.B. and Chen, T.T. (1986) *DNA* 5, 463–471.
- 9 Pan, F.M. and Chang, W.C. (1986) *J. Chin. Biochem. Soc.* 15, 24–33.
- 10 Gubler, U. and Hoffmann, B.J. (1983) *Gene* 25, 263–269.
- 11 Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.* 65, 499–560.
- 12 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- 13 Bewley, T.A. and Li, C.H. (1984) *Arch. Biochem. Biophys.* 233, 219–227.
- 14 Martial, J.A., Hallewell, R.A., Baxter, J.D. and Goodman, H.M. (1979) *Science* 205, 602–607.
- 15 Seeburg, P.H., Shine, J., Martial, J.A., Baxter, J.D. and Goodman, H.M. (1977) *Nature (London)* 270, 486–494.
- 16 Vize, P.D. and Wells, J.R.E. (1987) *Gene* 55, 339–344.
- 17 Kozak, M. (1981) *Nucleic Acids Res.* 9, 5233–5252.