

Evolution of the insulin superfamily: Cloning of a hybrid insulin/insulin-like growth factor cDNA from amphioxus

(protochordate/peptide hormones/preproinsulin/somatomedin)

SHU JIN CHAN, QIU-PING CAO, AND DONALD F. STEINER

Department of Biochemistry and Molecular Biology and The Howard Hughes Medical Institute, The University of Chicago, Chicago, IL 60637

Contributed by Donald F. Steiner, August 30, 1990

ABSTRACT Although insulin and the insulin-like growth factors (IGFs) share marked similarities in amino acid sequence and biological activity, their evolutionary origins have not been resolved. To investigate this issue, we recently cloned a cDNA encoding an insulin-like peptide (ILP) from a primitive chordate species, amphioxus (*Branchiostoma californiensis*). The deduced sequence of amphioxus preproILP indicates that it is a hybrid molecule containing features characteristic of both insulin and IGF. Like proinsulin, amphioxus proILP contains a C-peptide, which is flanked by paired basic residues and is probably removed by proteolysis. However, proILP also contains an extended carboxyl-terminal peptide region that can be divided into D and E domains similar to those of proIGF. Sequence comparisons show that the amphioxus ILP A and B domains are equally homologous to those of human insulin and IGF-I and -II. Based on these results and the exon-intron organization of the amphioxus ILP gene, we propose that IGF emerged at a very early stage in vertebrate evolution from an ancestral insulin-type gene.

The insulin superfamily includes insulin (1), insulin-like growth factors (IGFs) I and II (2), and relaxin (3) in vertebrates and molluscan insulin-like protein (4), an insulin-like peptide from *Locusta migratoria* (5), and the silkworm hormone bombyxin (6) in invertebrates. Of the above hormones, insulin and the IGFs have been shown to share close similarities in primary and tertiary structure and biological activity (2, 7). The molecular cloning of cDNAs encoding insulin and IGF-I receptors has revealed that they are homologous structures with about 60% identity in amino acid sequence (8, 9). These findings have led to the proposal that insulin and IGFs may have evolved from a common ancestral form but the evolutionary date of this divergence and the nature of the primordial gene is not known. In earlier studies, we have shown (10) that a highly conserved insulin gene is present in all vertebrate species, including the agnathan hagfish.

To further investigate the evolutionary origin of insulin, we have now cloned and sequenced* the cDNA encoding an insulin-like peptide (ILP) from *Branchiostoma californiensis* (amphioxus), a primitive cephalochordate that occupies a key position in chordate development as a possible extant relative of the invertebrate progenitor from which the vertebrates emerged (11). Surprisingly, our analysis of the deduced sequence for amphioxus preproILP indicates that it has the characteristics of a hybrid insulin/IGF molecule and thus may represent a transitional form connecting insulin and IGF. In addition, the exon-intron organization revealed by the cloned amphioxus ILP gene suggests a pathway by which IGF emerged from an ancestral two-chain insulin molecule early in vertebrate evolution.

MATERIALS AND METHODS

Materials. Radioisotopes were purchased from NEN. Murine leukemia virus reverse transcriptase, *Escherichia coli* polymerase I, T4 DNA ligase, and T4 polynucleotide kinase were obtained from BRL. *EcoRI/Not I* oligonucleotide adapter was from Pharmacia. γ gt10-*EcoRI*, λ DASHII-*EcoRI*, and *in vitro* packaging extracts (Gigapack II Plus) was obtained from Stratagene. Oligonucleotides were synthesized on an Applied Biosystems model 380B DNA synthesizer.

Amphioxus RNA and DNA. Live amphioxus (30–70 mm long) were caught off the coast of southern California (Pacific Biomarine Laboratories, Venice, CA) and were shipped to Chicago for processing. RNA was extracted from whole animals by homogenization in guanidine thiocyanate as described by Chirgwin *et al.* (12). Poly(A)⁺ RNA was isolated by affinity chromatography on oligo(dT)-cellulose (Pharmacia). cDNA was synthesized using murine leukemia virus reverse transcriptase and converted into double-stranded DNA with *E. coli* DNA polymerase I according to the manufacturer's instructions. An amphioxus phage cDNA library was constructed with 1.0 μ g of double-stranded cDNA ligated with *Eco/Not I* adapters by using T4 DNA ligase, phosphorylated with T4 polynucleotide kinase, ligated to 1.0 μ g of λ gt10-*EcoRI*, and packaged *in vitro*.

High molecular weight amphioxus DNA was isolated by the method of Blin and Stafford (13). To construct a genomic library, 2.0 μ g of amphioxus DNA was digested with *EcoRI* and equal amounts of DNA were ligated separately to 1.0 μ g of λ gt10-*EcoRI* and 10 μ g of λ DASHII-*EcoRI* and packaged *in vitro*. Library screening, Northern blot analysis, and other standard recombination DNA procedures were performed essentially as described (14). DNA sequence analysis was performed using Sequenase (United States Biochemicals).

Polymerase Chain Reaction (PCR). Degenerate oligonucleotide primers corresponding to conserved amino acid sequences in the insulin A chain were sense primer, 5'-GGTATHGTNGANCARTGYTG-3' (where H is A, T, or C; N is G, A, T, or C; R is G or A; and Y is C or T), which encode Gly-Ile-Val-Glu or Asp-Gln-Cys-Cys; antisense primer contained a mixture of two degenerate oligonucleotides, 5'-TT-ARTTRCARTARTTYTSNAG-3' (where S is G or C) and 5'-TTARTTRCARTARTTYTSYAA-3', which encode the cDNA for the peptide sequence Leu-Glu or Gln-Asn-Tyr-Cys-Asn-Ter. The PCR was performed in a Perkin-Elmer/Cetus DNA Thermal Cycler for 41 cycles (each cycle was 94°C for 1 min, 50°C for 2 min, and 65°C for 3 min) and used approximately 0.1 μ g of amphioxus cDNA as template. PCR products were analyzed on an 8% polyacrylamide gel in 90

Abbreviations: IGF, insulin-like growth factor; ILP, insulin-like peptide; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends.

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M55302).

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.

mM Tris borate/1 mM EDTA buffer and visualized by staining with ethidium bromide (1 µg/ml).

Rapid amplification of cDNA ends (RACE) reactions were performed essentially as described (15, 16).

RESULTS AND DISCUSSION

In cloning the amphioxus ILP cDNA, we utilized the PCR to first amplify a 66-base-pair (bp) fragment corresponding to the A chain region. Degenerate oligonucleotide primers that correspond to highly conserved amino- and carboxyl-terminal amino acid sequences in the insulin A chain were used to amplify cDNA prepared from amphioxus poly(A)⁺ RNA (11). Analysis of the PCR products by polyacrylamide gel electrophoresis revealed a major band of the expected size (data not shown). The 66-bp band was eluted from the gel, subcloned into pGEM-4Z, and plasmid DNAs from nine independent colonies were isolated and sequenced. Of these, eight clearly did not contain insulin-related sequences since they did not contain an invariant cysteine residue corresponding to position A11, which is found in all members of the insulin superfamily. One clone pAmI5, however, did contain the codon TGC for cysteine at position A11 as well as codons for several amino acids previously found in insulins from other species (Asn-A9, Val-A10, Asp-A12, Gln-A15).

Based on the pAmI5 sequence, specific primers were synthesized and the amphioxus ILP cDNA sequence was expanded using a modified PCR protocol for RACE, devised by Frohman *et al.* (15). Plasmids containing 300–500 bp of ILP cDNA extended in the 5' and 3' direction were obtained. However, analysis of the 3' RACE cDNA revealed the unexpected finding that the deduced sequence of the amphioxus ILP A chain did not terminate at residue 21 or 22, as is the case in insulin, but contained a number of additional amino acids. To verify that this was not an artifact generated by PCR, we screened an amphioxus cDNA library prepared by conventional cloning techniques. Several clones were isolated and one λAmI6 contained a 1.9-kilobase (kb) *EcoRI* insert, which was subcloned into pGEM-4Z and sequenced.

The sequence of pAmI6, presented in Fig. 1, revealed that it contained a 915-bp open reading frame encoding 305 amino acids flanked by a 84-bp 5' untranslated region and a 636-bp 3' untranslated region. The 3' untranslated region is complete as indicated by the presence of a homopolymeric d(A) tail and the polyadenylation signal sequence AATAAA.

The deduced amino acid sequence of the amphioxus ILP precursor reveals that, for the first one-third of its sequence (residues 1–101), it is organized much like a typical preproinsulin with a signal peptide, B chain, C-peptide, and A chain. The C-peptide contains 25 amino acids flanked by a pair of basic residues (Arg-Arg) at the amino terminus and a tan-

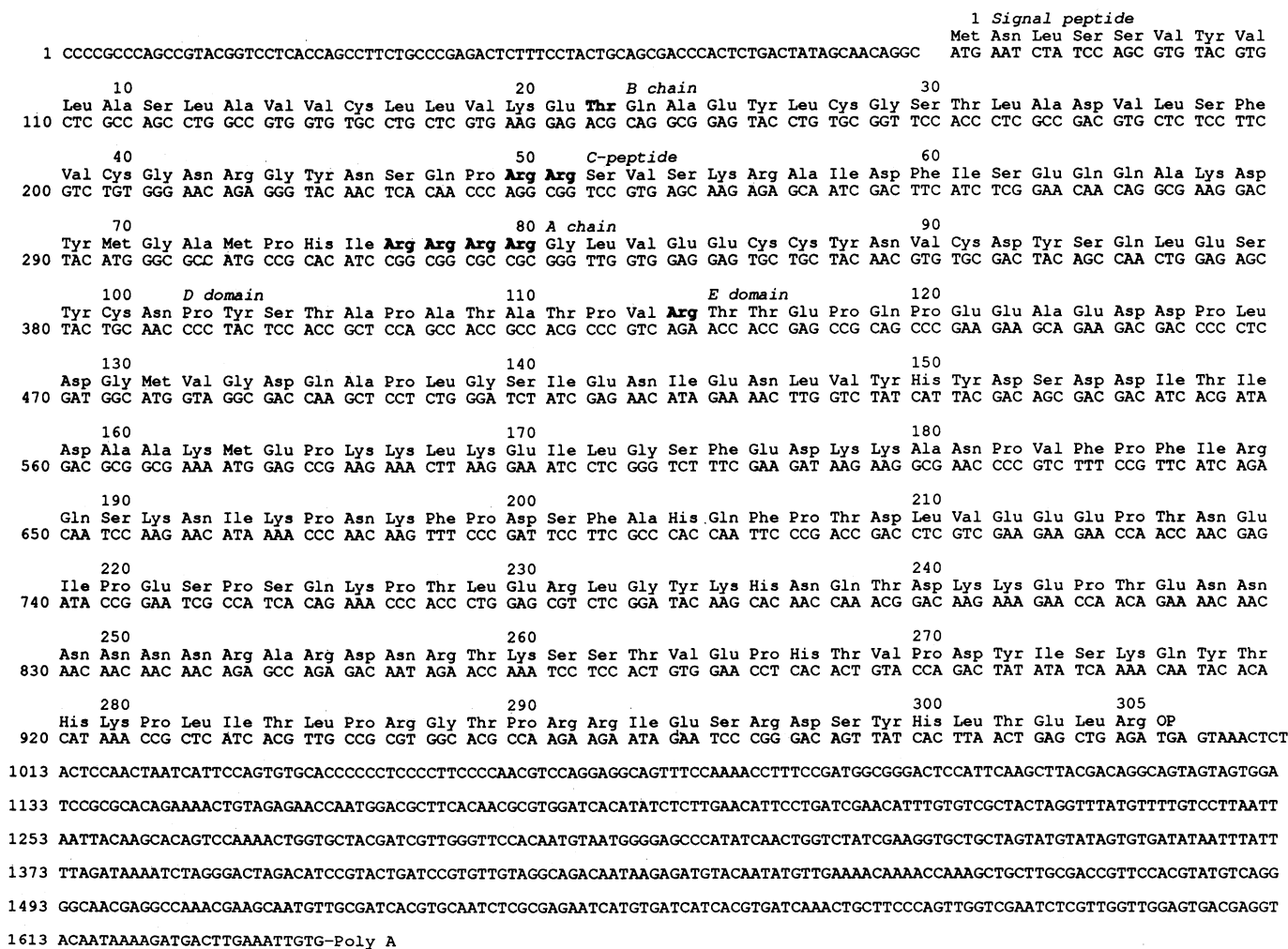


FIG. 1. Nucleotide and deduced amino acid sequence of amphioxus preproILP cDNA. A λgt10-amphioxus cDNA library (3 × 10⁵ recombinants) was probed with a 319-bp partial ILP cDNA fragment prepared from a 5' RACE reaction (12). One clone, λAmI6, contained a 1.9-kb *EcoRI* insert, which was subcloned into pGEM-4Z and sequenced. The predicted amino acid sequence for preproILP is divided into a signal peptide, B chain, C-peptide, A chain, D domain, and E domain. Residues in bold type indicate sites that may be proteolytically cleaved to produce mature ILP.

demly repeated pair of basic residues at the carboxyl end (Arg-Arg-Arg-Arg). A pair of basic residues (Lys-Arg) is also located within the C-peptide. However, a termination codon was not found at the end of the A chain and the remaining peptide sequence (residues 102–305) could be divided into putative D and E domains reminiscent of proIGF.

Northern blot hybridization of amphioxus poly(A)⁺ RNA, using as a probe a 1.1-kb *EcoRI-HindIII* fragment from pAml6, revealed a major 2.1-kb mRNA and a minor transcript of 4.1 kb (Fig. 2A). From densitometric scanning of the autoradiograph, the ratio of the 2.1-kb to 4.1-kb RNA species was approximately 6:1. We also quantitated the total amount of ILP RNA sequences present in amphioxus by dot-blot hybridization compared with known amounts of ILP-specific RNA transcribed from pAins6 with SP6 RNA polymerase. The results indicate that ILP mRNA represents about 0.1% of the total amphioxus poly(A)⁺ RNA (Fig. 2B).

Fig. 3 shows a comparison of the amphioxus ILP B and A domains with the homologous regions of human insulin and IGF-I and -II. The amphioxus ILP is identical to human insulin in 24 of 51 residues (48%); the sequence identity between amphioxus ILP and human IGF-I or -II is also 48%. Thus based on comparison of primary structures, amphioxus ILP is equally related to insulin and the IGFs.

The conserved residues found in amphioxus ILP include all residues that have been identified as being necessary to form the insulin-like tertiary conformation (17). However, the amphioxus ILP B and A domains also contain a number of other amino acid substitutions that, to our knowledge, have not been found in vertebrate insulin or IGFs. These include substitutions for Val-B12, Tyr-B16, and Ile-A2, which are all highly conserved residues in insulin. The amphioxus ILP, however, does contain asparagine at position A21, which is invariant among all the vertebrate insulins and is important for insulin biological activity (18). Another region that is important for biological activity in both insulin and the IGFs, is located at the carboxyl end of the B domain from positions B24 to B26. This region is always occupied by aromatic residues phenylalanine or tyrosine (except for certain hystricomorph insulins[†]) and has been postulated to interact with the insulin and IGF-I receptors (19). In the amphioxus ILP, a critical aromatic residue at Tyr-B24 has been conserved but the aromatic residues at B25 and B26 have been replaced with hydrophilic residues asparagine and serine.

The D domain in amphioxus ILP (12 amino acids) is similar in length to that found in human IGF-I (8 amino acids) or IGF-II (6 amino acids) but the putative E domain of (pro)ILP is much longer, 183 vs. 35 and 89 amino acids found in human proIGF-I and -II, respectively (20, 21). There is not significant sequence homology between amphioxus ILP and human IGFs in these regions. Although the D and E domains are relatively well conserved in mammals and in salmon proIGF (22), the sequence of hagfish proIGF also reveals very low sequence similarity in these domains.

Our analysis of the ILP precursor sequence, deduced from the cloned cDNA, thus suggests that a hybrid insulin/IGF-like hormone is produced in amphioxus. The sequence suggests that proILP is proteolytically cleaved like proinsulin to release a C-peptide but that it also contains extended D and E domains as found in proIGF. To further characterize the relationship among ILP, insulin, and IGF, we screened a phage library prepared from amphioxus genomic DNA and isolated two recombinants that hybridized with ILP cDNA. Restriction mapping and sequence analysis revealed that the clones contained 6.4- and 4.0-kb *EcoRI* fragments, ordered as shown in Fig. 4A, and that the ILP coding sequence is

[†]Insulins from the hystricomorph rodents, coypu and casiragua, have a deletion in this region so that the sequence is Phe-del-Tyr (where del is deletion) and cuis insulin has the sequence Phe-Phe-Ser.

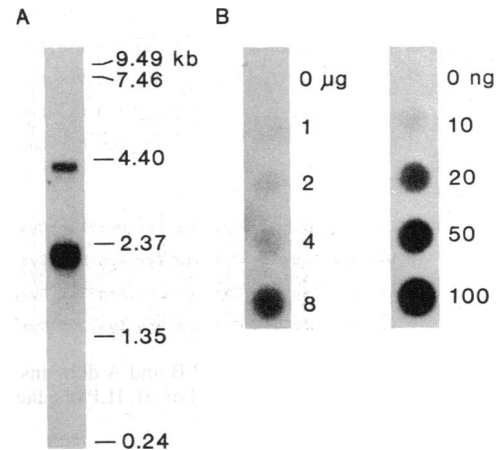


Fig. 2. Analysis of amphioxus ILP mRNA. (A) Northern blot analysis. Ten micrograms of amphioxus poly(A)⁺ RNA was electrophoresed on a 1% agarose/0.66 M formaldehyde gel, transferred onto a nitrocellulose filter, and hybridized with a labeled 1.1-kb *EcoRI-HindIII* fragment from pAml6. Molecular size markers shown are from an RNA ladder (BRL). (B) Quantitation of amphioxus ILP mRNA by dot-blot hybridization. The left column contains 0–8 μ g of amphioxus poly(A)⁺ RNA dot-blotted onto a nitrocellulose filter and hybridized with a labeled 1.1-kb *EcoRI-HindIII* fragment from pAml6; the right column contains 0–100 ng of ILP-specific RNA transcribed by SP6 polymerase from pAml6 (linearized with *Pvu* I) and dot-blotted onto the filter.

contained within two exons. One exon begins 9 bp upstream in the 5' untranslated region and encodes the signal peptide, B chain, and the first six amino acids of the C-peptide. The 3' exon-intron junction in the amphioxus ILP gene occurs at precisely the same position as found in the human (and other known vertebrate) insulin genes (23). The downstream exon encodes the remaining C-peptide, A chain, D domain, E domain, and the 3' untranslated region. At least one additional exon that has not yet been cloned must be present in the ILP gene to encode the remaining 5' untranslated region.

The above finding provides strong evidence that ILP and the vertebrate insulins share a common gene lineage, and it is possible that ILP represents an unusual insulin that, to our knowledge, has not been previously found in other vertebrates or invertebrates. However, the significant sequence similarity between ILP and the IGFs as well as the finding that proILP contains a C-terminal peptide that can be divided into putative D and E domains, suggest to us an attractive alternative hypothesis, namely, that ILP may represent an intermediate form linking the IGF genes with an ancestral insulin gene. The organization of the amphioxus ILP gene also provides a plausible pathway by which the IGF genes may have emerged from an ancestral insulin gene. As shown schematically in Fig. 4B, ILP is the first intermediate in this pathway and is formed by a nonsense-to-sense mutation in the chain-termination codon located at the end of the proto-insulin A chain. Next, we postulate that a shift in the intron donor site in the ILP gene upstream to residue Gln-48 (see Fig. 1), which is the position homologous to the donor site for intron 2 found in the IGF genes (24, 25), would delete the amino-terminal paired-basic-residue cleavage site and create a single-peptide-chain protoIGF molecule. This and additional point mutations or deletions would result in the shortened noncleaved C domain typically found in the IGFs. Subsequent events include the acquisition of an intron within the E domain and gene duplication to give rise to IGF-I and -II.

The fact that amphioxus occupies a key position on the phylogenetic tree as an early chordate, moreover, suggests to us that the appearance of the IGF genes may be a charac-

B domain																															
	1										10										20										30
Insulin	Phe	Val	Asn	Gln	His	Leu	Cys	Gly	Ser	His	Leu	Val	Glu	Ala	Leu	Tyr	Leu	Val	Cys	Gly	Glu	Arg	Gly	Phe	Phe	Tyr	Thr	Pro	Lys	Thr	
ILP	-	Gln	Ala	Glu	Tyr	Leu	Cys	Gly	Ser	Thr	Leu	Ala	Asp	Val	Leu	Ser	Phe	Val	Cys	Gly	Asn	Arg	Gly	Tyr	Asn	Ser	Gln	Pro	Arg	-	
IGF-I	-	Gly	Pro	Glu	Thr	Leu	Cys	Gly	Ala	Glu	Leu	Val	Asp	Ala	Leu	Gln	Phe	Val	Cys	Gly	Asp	Arg	Gly	Phe	Tyr	Phe	Asn	Lys	Pro	Thr	
IGF-II	Arg	Pro	Ser	Glu	Thr	Leu	Cys	Gly	Gly	Glu	Leu	Val	Asp	Thr	Leu	Gln	Phe	Val	Cys	Gly	Asp	Arg	Gly	Phe	Tyr	Phe	Ser	Arg	Pro	Ala	

A domain																					
	1										10										20
Insulin	Gly	Ile	Val	Glu	Gln	Cys	Cys	Thr	Ser	Ile	Cys	Ser	Leu	Tyr	Gln	Leu	Glu	Asn	Tyr	Cys	Asn
ILP	Gly	Leu	Val	Glu	Glu	Cys	Cys	Tyr	Asn	Val	Cys	Asp	Tyr	Ser	Gln	Leu	Glu	Ser	Tyr	Cys	Asn
IGF-I	Gly	Ile	Val	Asp	Glu	Cys	Cys	Phe	Arg	Ser	Cys	Asp	Leu	Arg	Arg	Leu	Glu	Met	Tyr	Cys	Ala
IGF-II	Gly	Ile	Val	Glu	Glu	Cys	Cys	Phe	Lys	Lys	Cys	Ser	Leu	Ala	Thr	Leu	Glu	Thr	Tyr	Cys	Ala

FIG. 3. Homology of amphioxus ILP B and A domains with human insulin and IGF-I and -II. Vertical lines indicate amino acid residues in ILP that are identical to insulin or IGF-I or -II. ILP residues in bold type indicate amino acids that, to our knowledge, have not been previously found in insulin or the IGFs.

teristic feature of the vertebrates. In support of this, the invertebrate insulin-like hormones, which have been characterized thus far, have been two-chain molecules in which the C-peptides were proteolytically removed; to date, no invertebrate single-chain IGF-like form has been described. In addition, multiple genes encoding the molluscan insulin-like peptides have been cloned by Smit (28) and sequence analyses revealed that they contain an intron that interrupts the amino-terminal coding region of the C-peptide consistent with the protoinsulin gene model.

An important corollary to the pathway proposed above is that if ILP is formed due to gene duplication and divergence,

then amphioxus should also contain a second insulin-like gene in which the A chain does not contain a C-terminal extension. Indeed, Van Noorden and Pearce (26) and Reinecke (27) have identified immunoreactive cells in the digestive tract of a European species of amphioxus (*Branchiostoma lanceolatum*) by using an antiserum to mammalian insulin. It is not known whether these cells contain ILP or a second insulin-like hormone. Southern blot analysis of amphioxus genomic DNA, however, performed under conditions of reduced stringency for hybridization, indicates that ILP is a single-copy gene (data not shown). We have also used a number of degenerate oligonucleotide primers that

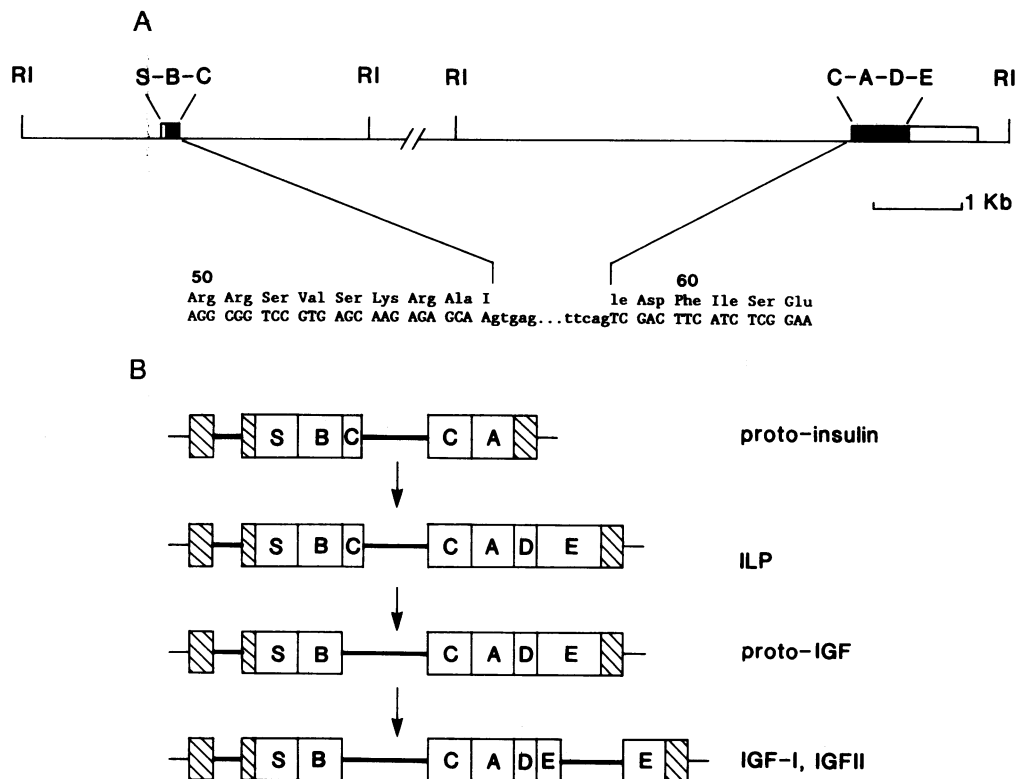


FIG. 4. (A) Organization of coding exons in the amphioxus ILP gene. Two clones, λAIG-4 and λAIG-6.4, were isolated from an amphioxus genomic phage library by hybridization with pA16. Restriction mapping and partial sequence analysis revealed that λAIG-4 contained a 4.0-kb *EcoRI* insert and a single exon spanning nucleotides 77-257 of the preproILP cDNA; AIG-6.4 contained a 6.4-kb *EcoRI* insert and a single exon that spanned nucleotides 258-1638 (see Fig. 1). The sequence of the exon-intron junction spanning the ILP C-peptide is shown. Solid boxes represent coding sequences; open boxes represent 5' and 3' untranslated regions. (B) Proposed evolutionary pathway from an ancestral insulin gene to form IGF-I and -II genes. The exon-intron organization of the protoinsulin gene is presumed to be the same as that found in the vertebrate insulin gene with three exons (boxes) and two introns (bold horizontal lines). Conversion of protoinsulin to ILP is accomplished by a nonsense to sense mutation occurring at the end of the A chain coding sequence in protoinsulin; conversion of ILP to protoIGF is accomplished by an upstream shift in the intron donor site into the B coding domain. IGF-I and -II genes evolve from protoIGF by insertion of an intron into the E coding domain and subsequent gene duplication. For simplicity of presentation, the recent finding that mammalian IGF-I and -II genes have multiple promoter sites and an alternative splicing pattern in the E domain is omitted.

correspond to the amino- and carboxyl-terminal regions of the A chain from the known vertebrate insulins, from bombesin, and the molluscan insulin-like peptide in the PCR with amphioxus cDNA as the template, but without success (S.J.C., unpublished results). It is possible that a second insulin-like gene in amphioxus contains other amino acid sequences in these regions of the A chain and thus was not amplified by the primers used in the PCR. Alternatively, we cannot exclude the possibility that during evolution a second insulin-like gene has been deleted in amphioxus.

Clearly, the definitive placement of ILP within the insulin/IGF gene family will require additional studies to more fully characterize this hormone, including studies on its biosynthesis, biological function, and interaction with its receptor. In particular, it would be of interest to determine if ILP can bind to an insulin or IGF-type receptor or to both. The identification of multiple insulin/IGF-I receptors in amphioxus would also provide additional support for the presence of a second insulin-like gene. In recent cloning experiments, we have found that amphioxus expresses at least one insulin/IGF-type receptor (S.J.C., unpublished results).

We thank Graeme I. Bell for helpful discussions and Florence Rozenfeld for her assistance in the preparation of this manuscript. This research was supported by the Howard Hughes Medical Institute and National Institutes of Health Grant DK 13914.

- Steiner, D. F., Bell, G. I. & Tager, H. S. (1989) in *Endocrinology*, ed. DeGroot, L. (Saunders, Philadelphia), pp. 1263–1289.
- Froesch, E. R. & Zapf, J. (1985) *Diabetologia* **28**, 485–493.
- Hudson, P., Haley, J., John, M., Cronk, M., Crawford, R., Haralambidis, J., Tregear, G., Shine, J. & Niall, H. (1983) *Nature (London)* **301**, 628–631.
- Smit, A. B., Vreughdenhil, E., Ebberink, R. H. M., Geraerts, W. P. M., Klootwijk, J. & Joosse, J. (1988) *Nature (London)* **331**, 535–538.
- Lagueux, M., Lwoff, L., Meister, M., Goltzene, F. & Hoffman, J. (1990) *Eur. J. Biochem.* **187**, 249–254.
- Adachi, T., Takiya, S., Suzuki, Y., Iwami, M., Kawakami, A., Takahashi, S. Y., Ishizaki, H., Nagasawa, H. & Suzuki, A. (1989) *J. Biol. Chem.* **264**, 7681–7685.
- Blundell, T. L., Bedarkar, S., Rinderknecht, D. & Humbel, R. E. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 180–184.
- Ebina, Y., Graf, L., Clauser, E., Ou, J.-H., Masiarz, F., Kan, Y. W., Goldfine, I. D., Roth, R. A. & Rutter, J. (1985) *Cell* **40**, 747–758.
- Ullrich, A., Gray, A., Tam, A. W., Yang-Feng, T., Tsubokawa, M., Collins, C., Henzel, W., Le Bon, T., Kathuria, S., Chen, E., Jacobs, S., Francke, U., Ramachandran, J. & Fujita-Yamaguchi, F. (1986) *EMBO J.* **5**, 2503–2512.
- Chan, S. J., Emdin, S. O., Kwok, S. C. M., Kramer, J. M., Falkmer, S. & Steiner, D. F. (1981) *J. Biol. Chem.* **256**, 7579–7602.
- Cade, T. J. (1979) in *Vertebrate Life*, eds. McFarland, W. N., Pough, F. H., Cade, T. J. & Heiser, J. B. (Macmillan, New York), pp. 21–51.
- Chirgwin, J. M., Przybyla, A. E., McDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
- Blin, N. & Stafford, D. W. (1976) *Nucleic Acids Res.* **3**, 2303–2306.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).
- Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
- Nishi, M., Chan, S. J., Nagamatsu, S., Bell, G. I. & Steiner, D. F. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5738–5742.
- Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M. C., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N. & Vijayan, N. M. (1988) *Philos. Trans. R. Soc. London Ser. B* **319**, 369–456.
- Gliemann, J. & Gammeltoft, S. (1974) *Diabetologia* **10**, 105–113.
- DeMeyts, P., Gu, J.-L., Shymko, R. M., Kaplan, B. E., Bell, G. I. & Whittaker, J. (1990) *Mol. Endocrinol.* **4**, 409–416.
- Jansen, M., van Schaik, F. M. A., Ricker, A. T., Bullock, B., Woods, D. E., Gabbay, K. H., Nussbaum, A. L., Sussenbach, J. S. & Van den Brande, J. L. (1983) *Nature (London)* **306**, 609–611.
- Bell, G. I., Merryweather, J. P., Sanchez-Pescador, R., Stempien, M. M., Priestley, L., Scott, J. & Rall, L. B. (1984) *Nature (London)* **310**, 775–777.
- Cao, Q.-P., Duguay, S. J., Plisetskaya, E., Steiner, D. F. & Chan, S. J. (1989) *Mol. Endocrinol.* **3**, 2005–2010.
- Steiner, D. F., Chan, S. J., Welsh, J. M. & Kwok, S. C. M. (1985) *Annu. Rev. Genet.* **19**, 463–485.
- Rotwein, P., Pollack, K. M., Didier, D. K. & Krivi, G. G. (1986) *J. Biol. Chem.* **161**, 4828–4832.
- Dull, T. J., Gray, A., Hayflick, J. S. & Ullrich, A. (1984) *Nature (London)* **310**, 777–780.
- VanNoorden, S. & Pearse, A. G. E. (1976) in *The Evolution of the Pancreatic Islets*, eds. Grillo, T. A. I., Leibson, L. & Epple, A. (Pergamon, Oxford), pp. 163–178.
- Reinecke, M. (1981) *Cell Tissue Res.* **219**, 445–456.
- Smit, A. B. (1990) Ph.D. dissertation (Free University of Amsterdam).