Biochimica et Biophysica Acta, 1088 (1991) 47–56

© 1991 Elsevier Science Publishers B.V. (Biomedical Division) 0167-4781/91/$03.50
ADONIS 016747819100057J

BBAEXP 92190

Molecular cloning of porcine α_1-microglobulin/HI-30 reveals developmental and tissue-specific expression of two variant messenger ribonucleic acids

Amir Tavakkol

Department of Dermatology, The University of Michigan, Ann Arbor, MI (U.S.A.)

(Received 17 May 1990)
(Revised manuscript received 14 August 1990)

Key words: α_1-Microglobulin; Tissue-specific gene expression; mRNA; Gene regulation; cDNA cloning; (Pig)

A 1008 basepair (bp) cDNA clone encoding 335 amino acids followed by an inframe TGA translation termination codon and a 295-nucleotide 3′ untranslated (UT) region has been isolated from a pig liver cDNA library. Based on the deduced amino acid and nucleotide sequence homology to a human cDNA (Kaumeyer, J.F., Polazzi, J.O. and Kotick, M.P. (1986) Nucleic Acids Res. 14, 7839–7850), the 5′ amino terminus was found to code for α_1-microglobulin (α_1-M), a 183 amino acid protein belonging to the lipocalin protein superfamily (Pervaiz, S. and Brew, K. (1985) Science 228, 335–337). The 3′ half encoded HI-30 which constitutes the Kunitz-type proteinase inhibitory (L-chain) domain of porcine inter-alpha-trypsin inhibitor (IαTI). In Northern blot hybridization, this cDNA identified two equally abundant mRNA species of approx. 1.3 kb and 1.6 kb in length. However, a 125 bp cDNA probe derived from the 3′ UT region of the cDNA hybridized only to the 1.6 kb mRNA. The differences observed in the 3′ UT region of these mRNAs suggest the utilization of alternative polyadenylation signals or presence of unprocessed nuclear RNA. Densitometric scanning of Northern blots indicated that α_1-M/HI-30 mRNA levels were higher (5–8-fold) in fetal and neonatal liver compared to that of primiparous pigs. In contrast, the RNA levels did not change significantly during pregnancy. Dot blot analysis of RNA indicated liver to be the major site of α_1-M/HI-30 mRNA expression with lower levels observed in the stomach. The results suggest that modulation of α_1-M/HI-30 gene expression could play a role during porcine growth. Increased IαTI L-chain mRNA levels may be particularly important in fetal and neonatal development when regulation of the inflammatory response and protection of macromolecules from proteolytic degradation is vital to survival and sustained growth.

Introduction

Human inter-alpha-trypsin inhibitor (IαTI) is a complex plasma protein of M_r 250000 originally identified as protein π [3]. Although thought for sometime to be a single chain polypeptide [4,5], IαTI has now been shown to be composed of two non-inhibitory heavy chains linked to a light (L) chain proteinase inhibitory domain known as HI-30 [6]. Polypeptides for two heavy and the single light chains appear to be synthesized by three different mRNA species [7]. A newly identified pre-alpha-trypsin inhibitor (PαTI; M_r 125000) utilizes the common L-chain Kunitz inhibitor but contains only a single heavy chain [8]. A glycosaminoglycan chain linked to serine-10 residue of L-chain (HI-30) appears to be responsible for physical association of the heavy and light chains in both IαTI and PαTI [8]. In human, the intracellular precursor for the L-chain peptide is composed of two functionally unrelated proteins: (1) α_1-microglobulin (HC), an acute phase reactant belonging to the lipocalin superfamily of plasma proteins that circulate in plasma both free and as a complex with IgA [9,10] and impedes antigen stimulation of lymphocytes [11] as well as inhibits neutrophil chemotaxis [12], and (2) a serine proteinase inhibitor (HI-30) which contains two tandem Kunitz-type domains and is responsible for the proteinase inhibitory activity of IαTI and PαTI [6,8].

Although the nucleotide sequence for the single common L-chain of human IαTI and PαTI have been determined by several laboratories [1,13–14], no cDNA clones have yet been identified from any other species. In an attempt to characterize porcine insulin-like growth

The sequence data in this paper have been submitted to the EMBL/Genbank Data Libraries under the accession number X52087.

Correspondence: A. Tavakkol, Department of Dermatology, R6558, Kresge I, The University of Michigan Medical School, Ann Arbor, MI 48109-0528, U.S.A.

factor I (IGF-I) [15], we originally employed immunoscreening of a pig liver cDNA library constructed in a lambda gt11 expression vector. A rabbit antiserum raised to human IGF-I [16] identified a cross-reacting lambda bacteriophage clone (CLT8c) that remained antibody positive after several rounds of screening. Further analysis of the CLT8c phage revealed three *Eco*RI restriction fragments that did not hybridize to a ^{32}P-labeled porcine IGF-I cDNA probe. We have determined the identity of these *Eco*RI fragments as encoding portions of HI-30 and serum albumin. The original cDNA isolate for IαTI L-chain (HI-30) has now been utilized to isolate a nearly full-length cDNA clone, and to examine tissue specificity and temporal expression of α_1-M/HI-30 mRNA during porcine development. Also, we provide evidence that porcine α_1 M and HI-30, similar to that of human, are co-translated from the same mRNA species [1]. However, in contrast to human, the intracellular precursor for α_1-M/HI-30 is encoded on two distinct mRNAs in the pig.

Materials and Methods

The *Escherichia coli* protein blocking extract, *EXO* III nuclease digestion kit, plasmid sequencing vector (pGEM-4Z), SP6 and T7 primers were all from Promega Biotech (Madison, WI). All other reagents were of molecular biology grade and obtained from various commercial sources.

Animals

The tissues were collected from pregnant sows at days 6, 12, 29, 40, 45, 60, 90, 105 and 110 of gestation (term = day 115), frozen immediately in liquid nitrogen and stored at −70 °C until use. The liver was removed from day 40, 90, 110 porcine fetuses and day 0 and 21 neonates. Fetal intestine and stomach at 110 days were also removed. Animal use protocols were conducted under guidelines established by the American Veterinary Medical Association and the National Institutes of Health and were previously approved by the Institutional Laboratory Animal Care and Use Committee of The Ohio State University.

Screening of pig liver cDNA library

(a) Immunoscreening. An adult pig liver cDNA library (Clontech; Palo Alto, CA) constructed in lambda gt11 expression vector [17] was screened [18] with a rabbit polyclonal anti-human IGF-I antiserum (UB286), provided by Drs L.E. Underwood and J.J. Van Wyk, Division of Pediatric Endocrinology, University of North Carolina, Chapel Hill, and the National Hormone and Pituitary Program. The dilution of antiserum used in the screening was emperically determined (1/5000 dilutions of the original concentration).

(b) Hybridization screening. This was performed by a modification of the established procedures [18] and as previously described [15]. The plasmid DNA was digested with *Eco*RI and the cDNA insert was gel purified (see below). All cDNA probes were labeled with deoxyribonucleotide [^{32}P]triphosphates by nick translation or random-primer labeling with specific activities of (3–5) · 10^8 cpm/μg DNA usually obtained. The probe was used at a concentration of (1–2) · 10^6 cpm/ml. Pre- and post-hybridization conditions have been described elsewhere [15].

Preparation of bacteriophage DNA

Large-scale DNA preparation from the recombinant bacteriophages were carried out using the plate lysate method [18]. Small scale phage DNA preparation was performed by DEAE cellulose (DE52) procedure as outlined below. Briefly, 1.5 ml of a phage lysate containing (5–10) · 10^9 pfu/ml was mixed with an equal volume of DEAE cellulose (pH 6.7) and the mixture was inverted 20–30-times. Bacterial DNA and RNA were pelleted by centrifugation for 10 min at 4 °C, the supernatant containing the phage removed, DNA extracted with phenol/chloroform after proteinase K treatment and precipitated in ethanol at −20 °C for 16–18 h. 1–3 μg of phage DNA was digested with *Eco*RI and electrophoresed on a 0.8% agarose slab gel, for ethidium bromide staining or blot hybridization.

Subcloning and probe purification

Standard procedures were employed for subcloning cDNA clones [18]. cDNA probe purification and elution from acrylamide gel have been previously described [15]. The 3′ UT region specific probe (AT 125) was isolated from the 1.3 kb cDNA insert of sti.1.10 plasmid (see Fig. 1) by double digestion with *Ava*I and *Eco*RI, followed by elution from the gel.

RNA isolation, Northern and dot blot analysis

Tissue RNA was prepared using the guanidine isothiocyanate procedure [18]. The RNA was denatured and fractionated on 1.2% formaldehyde-agarose gels and transferred to a Biotrans membrane (Schleicher and Schuell, Keene, NH). Blot hybridization was carried out at 42 °C [15]. RNA dot blotting was performed in the following manner. Samples containing known amounts of total cellular RNA were denatured in 20 mM Tris-HCl (pH 7.0), 50% (v/v) deionized formamide and 6% (v/v) formaldehyde. They were heated to 65 °C for 15 min and an equal volume of 20 × SSC was added. The RNA was bound to a Biotrans membrane using a microsample filtration manifold (Schleicher and Schuell). Filters were washed with 20 × SSC, air dried and baked for 2 h at 80 °C. The dot blot hybridization was performed in a rapid hybridization buffer (Amersham, Arlington Heights, IL) following the

manufacturer's recommendation. RNA and dot blots were quantified by laser densitometer (LKB) scanning of autoradiograms. Values for α_1-M/HI-30 signals from each time point were corrected by normalizing against serum albumin mRNA values in order to account for individual variation and RNA loading differences.

DNA sequencing

DNA restriction fragments were subcloned into pGEM-4Z sequencing vector. The cDNA inserts were sequenced in both orientations and on both strands using the Sanger dideoxy chain termination method [19], SP6 and T7 primers and Sequenase (United States Biochemicals). To sequence the largest cDNA, it was necessary to generate several *Pst*I and *Pst*I-minus subclones in pGEM-4Z. We took advantage of *Pst*I sites present in the cDNA insert (Fig. 1A) and the linker region of pGEM-4Z. Plasmid DNA of the largest clone (sti.1.10) was digested with *Pst*I followed by fractionation of the product in a 5% PAGE gel (see above). The *Pst*I restricted fragments were then recovered and subcloned into pGEM-4Z for sequencing. In order to construct a *Pst*I-minus subclone, the gel slice containing the vector DNA and the remaining portion of the cDNA insert was recovered. The DNA fragment was similarly eluted from the gel slice, religated and used for transformation of *E. coli* competent cells. Portions of the sequence data were also compiled after the construction of unidirectional deletions with *Exonuclease* III strategy [20]. The sequence data were analyzed on an IBM XT computer with an aid of Microgenie program (Beckman Instruments, Palo Alto, CA).

Results

Isolation and characterization of porcine α_1-microglobulin / HI-30 and serum albumin cDNA clones

Our initial goal was to isolate cDNA clones corresponding to porcine insulin-like growth factor I (IGF-I). Since porcine IGF-I immunoreacts with the rabbit antiserum raised to human IGF-I [21], we chose to employ this antiserum for isolation of IGF-I from a pig liver cDNA library. The antiserum (UB286) is specific for human IGF-I and exhibits less than 0.5% cross-reactivity with insulin-like growth factor II and essentially no reaction with insulin. The liver was used as the starting tissue since it was known to be the major site of IGF-I synthesis in human [22]. Of 15 immunoreactive recombinant bacteriophages initially identified by screening a total of $5 \cdot 10^5$ plaques, only one (CLT8c) remained antibody positive through several rounds of immunoscreening. Digestion of CLT8c phage with *Eco*RI yielded three cDNA inserts of 719 (sti.3), 636 (sti.1) and 456 (sti.2) basepairs (bp). Surprisingly, these cDNA inserts did not hybridize to a ^{32}P-labeled porcine IGF-I cDNA probe and were negative in an epitope selection experiment (data not shown).

To determine the identity of the cDNA inserts, they were subcloned into pGEM-4Z plasmid vector and sequenced from both orientations using the dideoxy chain termination method [19]. A search of the gene data bank revealed extensive sequence homology between one of the inserts and the L-chain inhibitor of human IαTI [1]. Subsequently, this cDNA clone (sti.1) was used to rescreen the liver library to obtain larger cDNA fragments. A number of hybridization positive bacteriophages were obtained representing an α_1-microglobulin/HI-30 (α_1-M/HI-30) mRNA abundance of 0.5% in pig liver cDNA libary. *Eco*RI-digested gel-fractionated phage DNA was then hybridized to the original isolate resulting in identification of two additional cDNA clones of approximately 1.0 kb (sti.1.4) and 1.3 kb (sti.1.10) which differed by the presence of 196-nucleotide long 3′ UT segment unique to sti.1.10. The two other cDNA clones (sti.2 and sti.3) were identified after DNA sequence analysis and RNA blotting and found to correspond to porcine serum albumin (data not shown). Fig. 1C depicts the strategy used in sequencing these clones and their alignment with a cDNA recently reported for pig serum albumin [23]. The partial restriction map and sequencing strategy used in determining the complete nucleotide sequence of L-chain clones are depicted in Fig. 1A and B. The largest cDNA fragment (sti.1.10) consisted of a 1008 bp open reading frame encoding 335 amino acids followed by an inframe TGA translation termination codon and a 295-nucleotide 3′ UT region. By analogy to a human cDNA [1], the pig clone contained nucleotide sequences for 5′ amino terminal signal peptide and complete sequence for α_1-M/HI-30 mRNA and a large segment of 3′ UT region. The sequence of the original clone (sti.1) was found to be identical to position 268–861 and 514–1107 of the two others (sti.1.4 and sti.1.10), respectively. Sti.1.4 cDNA clone also contained identical sequence to nucleotide 247–1107 of sti.1.10 but both sti.1 and sti.1.4 differed from sti.1.10 in three nucleotides, one of which (A → G) was located at position 555 in the coding sequence of sti.1.10. This nucleotide did not change the amino acid usage but generated an *Ava*I restriction site in sti.1.10 (Fig. 2) which was absent from the two smaller cDNAs (Fig. 1A). The two other nucleotide differences (C → T and T → C) were located in the 3′ UT region of sti.1.10 at positions 1016 and 1061, respectively. The putative polyadenylation signal AATAAA was located 70 nucleotides beyond the TGA and mapped to the same position in all three cDNA clones (Fig. 1A). In both sti.1 and sti.1.4 but not sti.1.10, this signal was followed by a poly(A) tract of varying length beginning at identical positions 594 and 861, respectively.

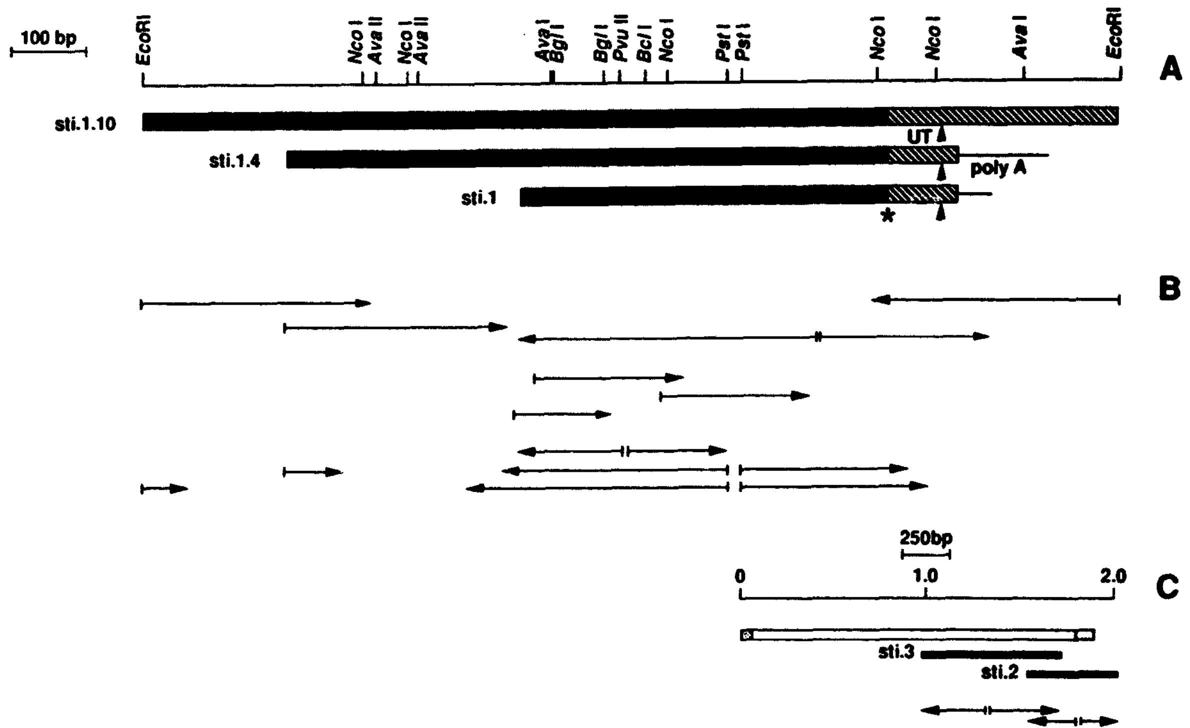

Fig. 1. Restriction map and sequencing strategy for determining α_1-M/HI-30 and serum albumin nucleotide sequences. (A) Restriction sites and the physical map of three independently isolated α_1-M/HI-30 cDNA. The smallest cDNA (sti.1) was used to rescreen the library for isolation of sti.1.4 and sti.1.10 clones. Coding sequences are shown in solid and non-coding 3′ untranslated (UT) regions are cross-hatched. The thin lines represent poly(A) tracts in sti.1 and sti.1.4. Short thick vertical arrows below the 3′ UTs indicate identical locations of polyadenylation signals. The translation termination codon TGA is marked by an asterisk and maps to the same position in all isolates. (B) Long arrows indicate the bilateral order and direction of various subclones used in DNA sequence analysis of α_1-M/HI-30 mRNA. (C) The previously reported cDNA clone for porcine serum albumin is shown as an open box, with the 5′ and 3′ UT regions dotted. The thick lines below the published pig serum albumin cDNA [23] indicate the two *Eco*RI fragments isolated in this work. Sti.2 extends the previously reported cDNA by 104 nucleotide in the 3′ direction.

Evolutionary considerations

The 5′ amino terminus of sti.1.10 contained coding sequences for a 183 amino acid polypeptide known as α_1-microglobulin [24], or protein HC [6,9] in human, while the carboxy-terminus encoded two Kunitz proteinase inhibitory domains (HI-30) that constitute the L-chain of IαTI and PαTI [6,8]. The nucleotide and deduced amino acid sequences of this cDNA (sti.1.10) have been presented in Fig. 2. The putative proteolytic cleavage site for release of α_1-microglobulin and HI-30; Arg Arg [1] occurs at amino acid residues 188 and 189 of sti.1.10. Derived amino acid sequences of human and pig proteins encoded by α_1-M/HI-30 mRNA exhibited 80% homology. Amino acid sequence of α_1-M is depicted in Fig. 3 along with those of human, rat, rabbit and the guinea-pig [25,26] to demonstrate sequence conservation. Except for human and a partial sequence for rat α_1-M, no other cDNAs were available from other species thus partial amino acids for these peptides have been compared.

HI-30 was isolated from human urine (syn UTI) [27] and shown to be identical to that found in serum [28]. Partial amino acid sequence for porcine homolog of HI-30 (PI-30) has been characterized from plasma [29,30]. Our cDNA-derived amino acid sequences differed in nine positions (residues: 258, 269, 277, 282,

Fig. 2. Nucleotide and deduced amino acid sequences of porcine α_1-M/HI-30 mRNA. The nucleotide sequences are numbered above the lines and amino acids on the side. The reactive site of the two protease inhibitory domains I and II (leucine and arginine, respectively) and the preceding three amino acids are boxed. The two arginines at the site of proteolytic cleavage are indicated by asterisks. The 19 amino acid long repeats flanking the reactive site in domain II are shown in bold type. Only 4 out of 19 amino acids are different in the repeats. The nucleotide variations detected between the three overlapping cDNAs are indicated directly above the variant nucleotides by the use of lower case letters. The new *Ava*I site generated by A → G substitution has been indicated. Putative N-linked glycosylation sites are overlined whereas the polyadenylation signal and *Eco*RI sites are underlined. The two potential poly(A) addition signals are shown by arrows.

```
                                                                         GAATTCCG

                                        30                                  60
GTG AGC GCC AGC CCT GTG CTG ACA TTG CCC AAT GAC ATC CAG GTG CAG GAG AAC TTC GAC
Val Ser Ala Ser Pro Val Leu Thr Leu Pro Asn Asp Ile Gln Val Gln Glu Asn Phe Asp 20

                                        90                                 120
CTG TCT AGG ATC TAC GGG AAA TGG TTC CAC GTG GCC GTG GGC TCC ACC TGC CCC TGG CTG
Leu Ser Arg Ile Tyr Gly Lys Trp Phe His Val Ala Val Gly Ser Thr Cys Pro Trp Leu 40

                                       150                                 180
AAG AGG TTC AAG GAC AAG ATG ATG ATG GGC ACG CTG ATG CTG GGA GAG GGG GCG ACG GAG
Lys Arg Phe Lys Asp Lys Met Met Met Gly Thr Leu Met Leu Gly Glu Gly Ala Thr Glu 60

                                       210                                 240
AGG GAG ATC AGC GTG ACC AAG ACT CAC CGG AGG AAA GGT ATC TGT GAG GTG ATC TCT GGG
Arg Glu Ile Ser Val Thr Lys Thr His Arg Arg Lys Gly Ile Cys Glu Val Ile Ser Gly 80

                                       270                                 300
GCT TAT GAG AAA ACA AGC ACT GAT GGA AAG TTC CTC TAT CAC AAA TCC AAA TGG AAC ATC
Ala Tyr Glu Lys Thr Ser Thr Asp Gly Lys Phe Leu Tyr His Lys Ser Lys Trp Asn Ile 100

                                       330                                 360
ACC ATG GAG TCC TAT GTG GTC CAC ACC AAC TAT GAT GAG TAT GCC ATA TTT CTG ACC AAG
Thr Met Glu Ser Tyr Val Val His Thr Asn Tyr Asp Glu Tyr Ala Ile Phe Leu Thr Lys 120

                                       390                                 420
AAG TTC AGC CGC CGC CAT GGA CCA ACC CTT ACT GCC AAG CTC TAC GGA CGG GAG CCG CAG
Lys Phe Ser Arg Arg His Gly Pro Thr Leu Thr Ala Lys Leu Tyr Gly Arg Glu Pro Gln 140

                                       450                                 480
CTT CGG GAA AGC CTG CTG GAG GAG TTC AGG GAG GTT GCC CTG GGC GTG GGC ATC CCG GAG
Leu Arg Glu Ser Leu Leu Glu Glu Phe Arg Glu Val Ala Leu Gly Val Gly Ile Pro Glu 160

                                       510                                 540
GAC TCC ATC TTT ACG ATG CCC GAC AGA GGA GAG TGT GTC CCT GGG GAG CAG GAG CCT GAG
Asp Ser Ile Phe Thr Met Pro Asp Arg Gly Glu Cys Val Pro Gly Glu Gln Glu Pro Glu 180

                 Ava I
                   a                   570                                 600
CCC ACC CTA CTC TCG AGA GCC CGG CGG GCC GTG CTG CCC CAG GAA GAG CAA GGA TCA GGA
Pro Thr Leu Leu Ser Arg Ala Arg Arg Ala Val Leu Pro Gln Glu Glu Glu Gly Ser Gly 200
                             *   *
                                          --> I
                                       630                                 660
GCT GGA CAA CCA GTA GCA GAT TTC AGC AAG AAA GAA GAT TCC TGC CAG CTG GGC TAC TCC
Ala Gly Gln Pro Val Ala Asp Phe Ser Lys Lys Glu Asp Ser Cys Gln Leu Gly Tyr Ser 220

                                       690                                 720
CAA GGC CCT TGC CTG GGC ATG ATC AAG AGG TAT TTC TAT AAT GGC TCA TCC ATG GCC TGC
Gln Gly Pro Cys Leu Gly Met Ile Lys Arg Tyr Phe Tyr Asn Gly Ser Ser Met Ala Cys 240

                                       750                                 780
GAG ACC TTC CAC TAT GGT GGC TGC ATG GGG AAC GGC AAT AAC TTC GTC TCG GAG AAG GAG
Glu Thr Phe His Tyr Gly Gly Cys Met Gly Asn Gly Asn Asn Phe Val Ser Glu Lys Glu 260

                          --> II
                                       810                                 840
TGT CTG CAG ACC TGC CGG ACT GTG GAG GCC TGC AGT CTC CCC ATC GTC TCC GGC CCC TGC
Cys Leu Gln Thr Cys Arg Thr Val Glu Ala Cys Ser Leu Pro Ile Val Ser Gly Pro Cys 280

                                       870                                 900
CGA GGT TTT TTC CAG CTC TGG GCG TTT GAT GCC GTG CAG GGG AAG TGT GTT CTC TTC AAC
Arg Gly Phe Phe Gln Leu Trp Ala Phe Asp Ala Val Gln Gly Lys Cys Val Leu Phe Asn 300

                                       930                                 960
TAT GGG GGC TGC CAG GGC AAC GGC AAC CAG TTC TAC TCG GAG AAG GAG TGC AAA GAG TAC
Tyr Gly Gly Cys Gln Gly Asn Gly Asn Gln Phe Tyr Ser Glu Lys Glu Cys Lys Glu Tyr 320

                                       990                              c
TGC GGC GTC CCC GGT GAA GAG GAT GAA GAG CTG CTG CGC TCC TCC AAC TGA CCAGTCCGCAG
Cys Gly Val Pro Gly Glu Glu Asp Glu Glu Leu Leu Arg Ser Ser Asn END
                                        t
GCCACAGGGCAGCAGGAGGGCCACGGCCAGCGCCTGCCCGGTGCCCCATGGCAGGTTCCAATAAAAACCAAATCGTAGC
CTCCTGAAATTCCACGTCCTGACTGTTCATCATTAAGTGTAATGAGATGGGGAGGGGAGCGGGGGACAAGCTGGGGTG
GGGGCCCGAGTAACCCCAGCATCCCCAGAAGTGAAAAATGTCTGTGTTGGAAATGTAATAGAACTCTCCTCCATACGTG
AAATTGGCTATGCAAATTATGAAACATAAATCACCCTTCTGTCGCTTAAG
```

284, 285, 292, 310 and 314; Fig. 2) from the amino acid sequence of PI-30 [30].

Pregnancy-associated and developmental expression of α_1-M/HI-30 mRNA

Northern blot analysis of total RNA from pig liver probed with ^{32}P-labeled serum albumin cDNA detected a single mRNA of approx. 2.5 kb as shown in Fig. 4, Fig. 5A and B. This cDNA was subsequently used in Northern blot hybridization as a control for RNA loading differences. To study temporal expression of α_1-M/HI-30 mRNA in pregnant animals, total RNA from sows was hybridized to the ^{32}P-labeled cDNA probes for α_1-M/HI-30 and serum albumin. The results, shown as a bar graph in Fig. 4, demonstrated that pig α_1-M/HI-30 mRNA levels remained virtually unchanged during pregnancy. Interestingly, also no stimulation of α_1-M/HI-30 mRNA expression was observed in estrogen-primed animals (E_2, Fig. 6). To delineate developmental expression of this mRNA, total liver RNA from fetal and postpartum pigs was analyzed by RNA blotting. The results shown in Fig. 5A and 5B revealed that L-chain mRNA expression was considerably higher in fetal and neonatal liver and declined to low levels in pregnant adult pigs. Densitometric scanning of blots from two independent experiments indicated 5–8-fold higher levels of expression in fetal liver compared to that of acult pigs (see bar graphs in Fig. 5A and B).

The α_1-M/HI-30 cDNA probe used in blot hybridization identified two mRNA bands corresponding to an estimated size of 1.3 kb and 1.6 kb (Fig. 4 and 5A). The size of the two mRNAs remained constant during pregnancy and development and no changes in their relative abundance were observed. Previous reports have not indicated the presence of two human α_1-M/HI-30 mRNAs. Therefore, to characterize the two variant mRNAs we constructed a cDNA probe specific for the 3′ UT of sti.1.10, the longest cDNA in our clones. This

```
                                                            25
pig         V S A S P V L T L P N D I Q V Q E N F D L S R I Y
human       - - - G P - P - P - D N - - - - - - - N I S - I -
rat               S N - P - L - D Z - - - - X - - N A A
rabbit            S P - P - L - D D - - - - - - - E L S - I -
guinea p.                 L - D D - - - - - - - D E S - M -

                                                            50
pig         G K W F H V A V G S T C P W L K R F K D K M M M G
human       - - W Y N L - I - - - - - - - - K I M - R - T V S
rat
rabbit      - - X Y N L - V
guinea p.   - - L Y S L - I - - -

                                                            75
pig         T L M L G E G A T E R E I S V T K T H R R K G I C
human       - - V - - - - - - - A - - - M - S - R W - - - V -

                                                           100
pig         E V I S G A Y E K T S T D G K F L Y H K S K W N I
human       - E T - - - - - - - D - - - - - - - - - - - - - -

                                                           125
pig         T M E S Y V V H T N Y D E Y A I F L T K K F S R R
human       - - - - - - - - - - - - - - - - - - - - - - - - H

                                                           150
pig         H G P T L T A K L Y G R E P Q L R E S L L E E F R
human       - G - - I - - - - - - - A - - - - E T - - Q D - -
rat         - A - - I - - - - - - - E - - - - D S - - Q E - -

                                                           175
pig         E V A L G V G I P E D S I F T M P D R G E C V P G
human       V - - Q G - - - - - D - - F T - A - - - - - - - -
rat         E - - L S - - - - - N - - V F - A - - - - - - - -

                                    188
pig         E Q E P E P T L L S R A R
human       E Q - P - - I - I P - V -
rat         D R - V -
```

Fig. 3. Inter-species comparison for porcine α_1-microglobulin. Pig, human and rat amino acids are derived from cDNA sequences whereas that of rabbit and guinea-pig are based on amino terminal peptide sequence analysis. The presence of a dash (-) denotes sequence homology. Blank spaces indicate where no published sequence were available, (x) where the amino acid residue was not determined and (z) where a deletion has been made for maximum homology. The data for the human cDNA are taken from Ref. 1; for rat from positions 4 to 22 [26] and from positions 126 to 181 [25] and for rabbit and guinea-pig [26].

Fig. 4. Pregnancy-associated α_1-M/HI-30 mRNA expression. Total RNA was prepared from pregnant animals at indicated time points. RNA blots were prepared and probed with α_1-M/HI-30 and serum albumin-specific ^{32}P-labeled cDNAs as indicated. The two bands representing α_1-M/HI-30 mRNA are located below 18 S ribosomal RNA marker, whereas the single mRNA band between the 18S and 28S belongs to serum albumin (cf Fig. 5). The bar charts are obtained by densitometer scanning of autoradiograms (see RNA isolation, 'Materials and Methods'). * Term in the pig is 115 days.

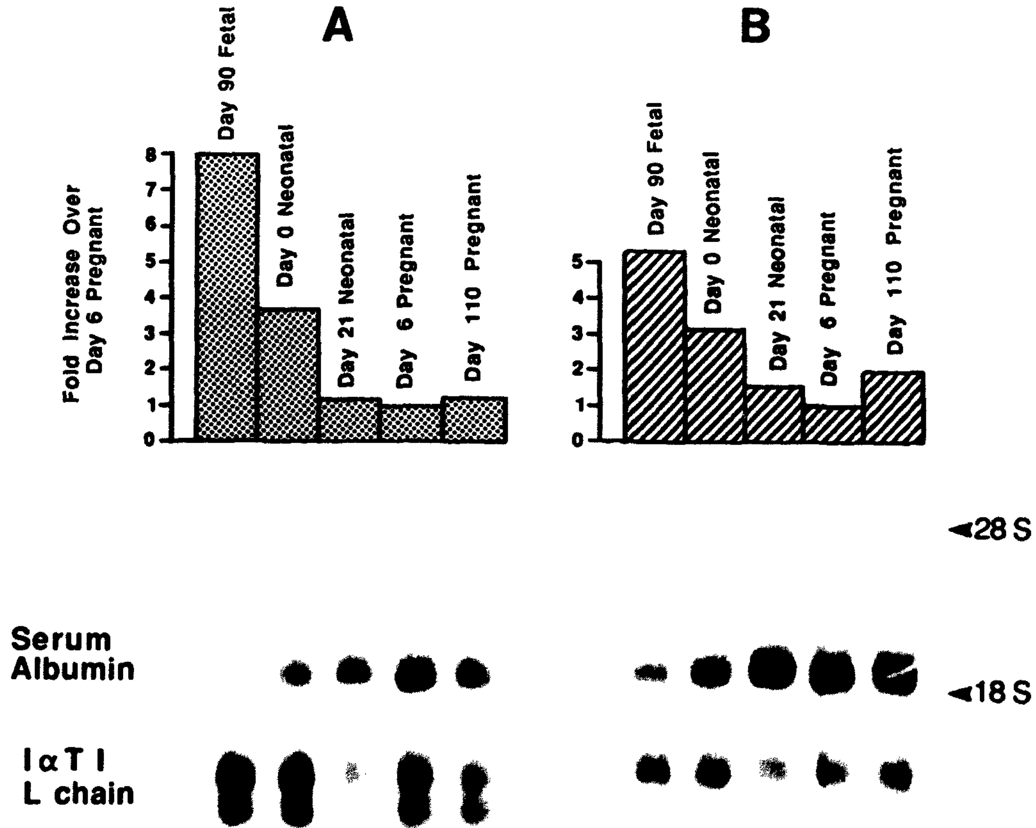

Fig. 5. Temporal regulation of α_1-M/HI-30 mRNA in pig liver. Total pig liver RNA (40 μg) was analyzed by RNA blotting. Hybridization probes were: (A) ^{32}P-labeled α_1-M/HI-30 cDNA or (B) ^{32}P-labeled AT 125 (see 'Results'). Both blots were also hybridized with the serum albumin cDNA probe as the reference gene (see legend to Fig. 4 and text). Liver RNA was obtained from individual animals at the development stages indicated. The two RNA blots were replicas prepared from the same RNA but hybridized separately. Bar charts are explained in legend to Fig. 4.

was achieved as described in 'Materials and Methods' (see also Fig. 1A), yielding a 125-nucleotide cDNA fragment (AT 125). This diagnostic probe hybridized only to a single mRNA of approx. 1.6 kb with no hybridization to the 1.3 kb mRNA even after prolonged autoradiographic exposure (Fig. 5B), indicating that the 1.3 kb cDNA (sti.1.10) was derived from 1.6 kb mRNA, whereas the other two α_1-M/HI-30 cDNAs (sti.1 and sti.1.4) originated from 1.3 kb mRNA. Using RNA dot blot analysis, liver was shown to be the major site of porcine α_1-M/HI-30 mRNA synthesis with lower levels observed in the stomach (Fig. 6). Few or no signals were detected for mammary, spleen and uterus at various developmental stages. Similar findings were also obtained by Northern blot analysis (not shown).

Discussion

We have identified three cDNA fragments corresponding to pig α_1-M/HI-30 and serum albumin mRNAs (sti.1, sti.2 and sti.3) all contained within a single immunoreactive bacteriophage (CLT8c). Cross-reactivity of CLT8c with the antiserum raised to human IGF-I may have been due to shared epitopes between pig serum albumin and/or α_1-M/HI-30 with human IGF-I, or to impurities present in antiserum preparations. However, the former possibility is unlikely as these proteins are unrelated both functionally and structurally, based on amino acid sequence comparison and hydropathicity analyses (data not shown). It is also likely that high serum concentration of IαTI protein

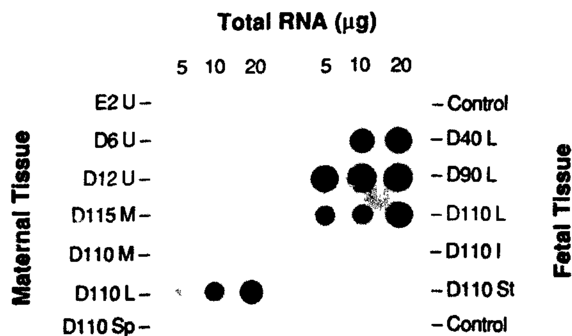

Fig. 6. Tissue distribution of IαTI messenger RNA. Total RNA was prepared from either fetal or maternal pig tissues at various stages of development and denatured with formaldehyde. The RNA was applied in 5, 10 and 20 μg quantities to Biotrans membrane and hybridized to a ^{32}P-labeled cDNA probe. Sample RNA blotted onto each well was adjusted to a total of 20 μg, composed of the liver tissue RNA at the given concentration plus yeast tRNA as a carrier. The dot blot was later hybridized to ribosomal DNA probes (not shown) to verify equal RNA loading. Control, yeast tRNA; D, day; E_2, pseudopregnant; I, intestine; L, liver; M, mammary; Sp, spleen; St, stomach and U, uterus.

(0.5 mg/ml) [28] which is also consistent with high mRNA representation (0.5% in lambda gt11 expression library) was responsible for production of cross-reacting anti-IαTI L-chain antibody in IGF-I antiserum preparations. This antiserum could have subsequently reacted to the recombinant phage expressing α_1-M/HI-30/β-galactosidase fusion protein during our library screening. We did not investigate whether physical association of α_1-M and immunoglobulin A [10] played a role in these events. Presence of two unrelated cDNAs encoding serum albumin and the L-chain of IαTI was also likely to have been due to a recombination event and an artefact during library construction. It is not uncommon for highly abundant mRNA species to be ligated together and cloned into the same lambda phage.

As this work was in progress a cDNA clone was described for porcine serum albumin, the sequence of which [23] was identical to those we isolated (data not shown). We detected a single mRNA species of approx. 2.5 kb consistent with the size of mature serum albumin mRNA in other species [31]. The α_1-M/HI-30 cDNA sequence described here identifies the pig as only the second species for which the DNA sequence is now available. Like its human homolog [1,14], the porcine cDNA also encodes the HI-30 [30] and α_1-M [6]. However, in pig α_1-M/HI-30 appears to be encoded by at least two variant mRNAs. Therefore, we conclude that these mRNAs are either the result of different polyadenylation or represent unprocessed nuclear RNA.

Three single point variations were detected in sti.1.10, one of which (position 555) may represent a restriction fragment length polymorphism (RFLP) since it introduced a new *Ava*I site into the domain encoding α_1-M without changing the amino acid usage. The presence of these single-base substitutions may indicate allelic differences in α_1-M/HI-30 gene, although it is possible that they were generated during the cDNA library construction. RFLP's in α_1-M/HI-30 gene have been previously documented [32], though not involving the nucleotides reported here. Overall, there was 80% sequence conservation between deduced amino acids of pig α_1-M/HI-30 mRNA and its human counterpart. Homology appears to vary when different domains are compared; pig α_1-M shows 79% homology to human [1], 78%, 77% and 68% with partially reported domains of rat, rabbit and guinea-pig, respectively [25,26]. Human α_1-M has been shown to contain tripeptide Asn-x-Ser/Thr at residues 20–22 [26], the consensus sequence for N-linked glycosylation [33]. Our cDNA assigns Asp-Lys-Ser to these positions thus excluding the possibility of glycosylation at this site, similar to rat, rabbit and guinea-pig [26]. However, the nucleotide sequence described here confirms the presence of a second N-linked glycosylation at positions 99–101 similar to that reported for human [1,26]. With respect to the proteinase inhibitory domains (comparison is not shown), porcine polypeptide is 85% homologous with sheep [34], 83% with human and bovine [1,29] and 81% with the horse inhibitor [30]. It is noteworthy that the deduced amino acid sequence of L-chain (HI-30) in our cDNA and that obtained by amino acid sequencing of the pig [30] are in major agreement (93% homology).

The light and heavy chain(s) in IαTI and PαTI are cross-linked by a chondroitin sulfate-like carbohydrate structure [35], reported to be glycosaminoglycans (GAGs) [36]. The GAG-like structure is believed to be originated from serine-10 of human HI-30 [8]. We also note that the amino acids flanking the serine-10 (residues 198–201 in Fig. 2) in pig α_1-M/HI-30 are identical to human sequence except for an alanine to glycine substitution indicating that similar carbohydrate structures may be present in porcine HI-30. Taken together, these findings imply that IαTI L-chain gene has remained relatively conserved throughout evolution. Recently, it was shown that the two tandem domains of HI-30 were each encoded by a single exon that was not split by intervening sequences [37] resembling aprotinin gene organization to whose family HI-30 belongs. We have also noted the presence of two repeat sequences in the two proteinase inhibitory domains which may indicate a gene duplication event involving a single exon.

We previously demonstrated that uterine but not liver IGF-I mRNA levels underwent drastic changes during pregnancy [15]. Feeding of colostrum devoid of proteinase inhibitors results in decreased intestinal absorption in vivo of albumin, IgG, β-lactoglobulin and growth factors in the neonatal pig [38,39]. Moreover, HI-30 is present in mammary secretions along with trypsin and chymotrypsin, suggesting a role for this

inhibitor during lactation [40]. However, little is known regarding modulation of IαTI L-chain gene expression and contribution of this inhibitor to colostrum-derived proteinase inhibitory pool in pregnant sows. Examination of uterus and mammary tissues of sows at various stages of pregnancy did not detect a significant difference in steady state IαTI L-chain mRNA levels, similar to findings for anther rat liver specific gene(s) encoding α_1-inhibitor III [41]. Our data also indicates that hepatic regulation of IαTI light chain mRNA in the pig unlike rat α_2-macroglobulin protease inhibitor [42] is not influenced by pregnancy. We also studied developmental expression of α_1-M/HI-30 mRNA at five different ages. Quatitative analysis of hybridization signals by a laser densitometer indicated that α_1-M/HI-30 mRNA levels were 5–8-fold higher in the fetus and neonate than adult animal. Analysis of two different RNA preparations and the use of 3′-UT-specific probe confirmed the general pattern of age dependence in L-chain mRNA expression. Decreased levels of proteinase inhibitory capacity in adult pigs compared to fetal tissues have been reported elsewhere [43] verifying the RNA expression studies reported here. Characterization of porcine α_1-M/HI-30 cDNA sequence and its mRNA expression reveals evolutionarily conserved nature of these genes but indicates specific differences in pattern of gene expression. These findings should help accelerate efforts to underst:nd the role of these important family of inhibitors

Acknowledgements

This work was carried out while the author was on a NIH funded postdoctoral fellowship at the laboratory of Frank A. Simmen, Ph.D. at the Ohio Agricultural Research and Development Center (OARDC), The Ohio State University. Salaries and research support were also provided by State and Federal funds appropriated to the OARDC. Frank Simmen was the recipient of the NICHHD (New Investigator Grant HD-22004). The author would like to acknowledge Judy Riggenbach for RNA preparation and gel electrophoresis, Linda Foster for plasmid DNA preparations, Beverly Fisher (The Ohio State University), Claudia M. Hagedon and Alexandria V. Chistyakova (The University of Michigan) for expert secretarial assistance. I am indebted to Drs Douglas N. Foster and Clague P. Hodgson (OARDC) for the critical reading of this manuscript and many useful discussions.

References

1 Kaumeyer, J.F., Polazzi, J.O. and Kotick, M.P. (1986) Nucleic Acids Res. 14, 7839–7850.
2 Pervaiz, S. and Brew, K. (1985) Science 228, 335–337.
3 Steinbuch, M. and Loeb, J. (1961) Nature 192, 1196.
4 Morii, M. and Travis, J. (1985) Biol. Chem. Hoppe-Seyler 366, 19–21.
5 Reisinger, P.W.M., Hochstrasser, K., Albrecht, G.J., Lempart, K. and Salier, J.P. (1985) Biol. Chem. Hoppe-Seyler 366, 479–483.
6 Schreitmüller, T., Hochstrasser, K., Reisinger, P.W.M., Wachter, E. and Gebhard, W. (1987) Biol. Chem. Hoppe-Seyler 368, 963–970.
7 Gebhard, W., Schreitmüller, T., Hochstrasser, K. and Wachter, E. (1989) Eur. J. Biochem. 181, 571–576.
8 Enghild, J., Thogersen, I.B., Pizzo, S.V. and Salvesen, G. (1989) J. Biol. Chem. 264, 15975–15981.
9 Tejler, L. and Grubb, A.O. (1976) Biochim. Biophys. Acta 439, 82–94.
10 Truedsson, L. and Grubb, A. (1988) Scand. J. Immunol. 27, 201–208.
11 Lögdberg, L. and Akerström, B. (1981) Scand. J. Immunol. 13, 383–390.
12 Méndez, E., Fernandez-Luna, J.L., Grubb, A. and Leyva-Cobin, F. (1986) Proc. Natl. Acad. Sci. USA 83, 1472–1475.
13 Bourguignon, J., Diarra-Mehrpour, M., Sesboüé, R., Frain, M., Sala-Trepat, J.M., Martin, J.P. and Salier, J.P. (1985) Biochem. Biophys. Res. Commun. 131, 1146–1153.
14 Gebhard, W., Schreitmüller, T., Hochstrasser, K. and Wachter, E. (1988) FEBS Letts. 229, 63–67.
15 Tavakkol, A., Simmen, F.A. and Simmen, R.C.M. (1988) Mol. Endocrinol. 2, 674–681.
16 Furlanetto, R.W., Underwood, L.E., Van Wyk, J.J. and D'Ercole, A.J. (1977) J. Clin. Invest. 60, 648–657.
17 Huynh, T.V., Young, R.A. and Davis, R.W. (1985) in DNA Cloning, A Practical Approach (Glover, D.M., ed.), Vol. 1, pp. 49–78, IRL Press, Oxford.
18 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
19 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
20 Henikoff, S. (1984) Gene 28, 351–359.
21 Simmen, F.A., Simmen, R.C.M. and Reinhart, G. (1988) Devel. Biol. 130, 14–25.
22 Lund, P.K., Moats-Staats, B.M., Hynes, M.A., Simmons, J.G., Jansen, M., D'Ercole, A.J. and Van Wyk, J.J. (1986) J. Biol. Chem. 261, 14539–14544.
23 Weinstock, J. and Baldwin, G.S. (1988) Nucleic Acids Res. 16, 9045.
24 Traboni, C. and Cortese, R. (1986) Nucleic Acids Res. 14, 6340.
25 Kastern, W., Björck, L. and Akerström, B. (1986) J. Biol. Chem. 261, 15070–15074.
26 Akerström, B., Lögdberg, L., Babiker-Mohamed, H., Lohmander, S. and Rask, L. (1987) Eur. J. Biochem. 170, 143–148.
27 Gebhard, W. and Hochstrasser, K. (1986) in Proteinase Inhibitors (Barret, A.J. and Salvesen, G., eds.), pp. 389–401. Elsevier/North-Holland, Amsterdam.
28 Hochstrasser, K. and Wachter, E. (1979) Hoppe-Seylers Z. Physiol. Chem. 360, 1285–1296.
29 Hochstrasser, K., Albrecht, G.J., Schönberger, O.L. and Wachter, E. (1983) Hoppe-Seylers Z. Physiol. Chem. 364, 1689–1696
30 Hochstrasser, K., Wachter, E., Albrecht, G.J. and Reisinger, P. (1985) Biol. Chem. Hoppe-Seyler 366, 473–47?.
31 Yamauchi, A., Imai, E., Noguchi, T., Tanaka, T., Yamamoto, S., Mikami, H., Fukuhara, Y., Fujii, M., Orita, Y. and Kamada, T. (1988) Amer. J. Physiol. 254, E676–679.
32 Léveillard, T., Bourguignon, J., Sesboüé, R., Hanauer, A., Salier, J.P., Diarra-Mehrpour, M. and Martin, J.P. (1988) Nucleic Acids Res. 16, 2744.
33 Hubbard, S.C. and Ivatt, R.J. (1981) Annu. Rev. Biochem. 261, 555–583.

34 Rasp, G., Hochstrasser, K., Wachter, E. and Reisinger, P.W.M. (1987) Biol. Chem. Hoppe-Seyler 368, 727–731.
35 Jessen, T.E., Faarvang, K.L. and Ploug, M. (1988) FEBS Lett. 230, 195–200.
36 Balduyck, W., Mizon, C., Loutfi, H., Richet, C., Roussel, P. and Mizon, J. (1986) Eur. J. Biochem. 158, 417–422.
37 Vetr, H., Kogler, M. and Gebhard, W. (1989) FEBS Lett. 245, 137–140.
38 Carlsson, L.C.T., Weström, B.R. and Karlsson, B.W. (1980) Biol. Neonate 38, 309–320.
39 Weström, B.R., Carlsson, L.C.T. and Karlsson, B.W. (1979) in Protein Transmission Through Living Membranes (Hemmings, W.A., ed.), pp. 225–232 Elsevier/North-Holland, Amsterdam.
40 Weström, B.R., Svendsen, J. and Karlsson, B.W. (1982) Biol. Neonate 42, 185–194.
41 Thompson, A.W., Wong, F.K. and Horne, C.H. (1980) Inves. Cell Pathol. 3, 231–236.
42 Aiello, L.P., Shia, M.A., Robinson, G.S., Pilch, P.F. and Farmer, S.R. (1988) J. Biol. Chem. 263, 4013–4022.
43 Ohlsson, B.G., Westrom, B.R. and Karlsson, B.W. (1986) Biol. Neonate 49, 292–300.