Volume 15 Number 7 1987 Nucleic Acids Research

Sequence of the porcine transforming growth factor-beta precursor

Rik Derynck and Lucy Rhee

Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080, USA

Submitted January 7, 1987 Accession no. Y00111

A λ gt-10-based cDNA library derived from polyadenylated porcine ovary mRNA (1) was screened by hybridization to the human transforming growth factor-β (TGF-β) cDNA insert of λ BC_1 (2). Two strongly hybridizing recombinant cDNA phage were sequenced (3,4) and their cDNA sequences were in absolute agreement. The longest cDNA insert contained in λ pigβ1 is 1,604 bp and encodes the complete porcine TGF-β precursor (Fig. 1). Comparison of the deduced amino acid sequences for the human (2), murine (5) and porcine TGF-β precursors shows that the N-terminal third of the precursor is virtually identical while the middle third is somewhat less conserved. The C-terminal third of the precursor which comprises the mature TGF-β monomer sequence is again very conserved. The RGDL sequence (residues 244-247) which could play a role in cellular adhesion is present in all three precursors (6,7). The polypeptide sequence of the mature human and porcine TGF-β is identical. The porcine precursor sequence contains only two potential N-glycosylation sequences instead of the three corresponding sequences in the human and murine TGF-β precursors.

ACKNOWLEDGEMENTS We thank Dr A.Mason for providing the cDNA library, Dr E.Chen for his support in the sequence analysis.

REFERENCES 1. Mason,A.J. et al. (1985) Nature 318, 659-663. 2. Derynck,R. et al. (1985) Nature 316, 701-705. 3. Sanger,F. et al. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467. 4. Messing,J. et al. (1981) Nucl. Acids Res. 9, 309-321. 5. Derynck,R. et al. (1986) J. Biol. Chem. 761, 4377-4379. 6. Pierschbacher,M.D. and Ruoslahti,E. (1984) Nature 309, 30-33. 7. Pierschbacher,M.D. and Ruoslahti,E. (1984) Proc. Natl. Acad. Sci. USA 81, 5985-5988.

```
   1 CGACTCCCGCGAAGACTTGACCCCAAAGCTCGGGCGCACCCCCCTGCACACTTCCCCACTCTCAGCCTCTCTGCTGAGCCCCTGCGCATCCAAGGACCCT
 101 TCTCGGATCCGGGAGACGGAATCTGTCTCAGACCTGCCTCAGCTTTCCTATTCAAGACCACCCACCTCTGGTACCAGATCTCGCCCATCTCGGTTTTTTC
 201 CGTAGGATACCGAGAACCCACCCATCAGAGCCTCCCTTCCACCTCTGCTCCTCCGTTCTCCCTGAGAGCCTCAACTTCCCTCCCACCCCAGATCCTCCTA
 301 CCTTTTCTGGGGAGACCCCTCAGCCCCTGTAGGGGCGGGGCCTCCCTCTTCCTACCCCAGCCGGCTCGCACTCTCGGCTGTGCCGGGGGGCGCCGCCTCC

       Met Pro Pro Ser Gly Leu Arg Leu Leu Pro Leu Leu Leu Pro Leu Leu Trp Leu Leu Val Leu Thr Pro Gly
 401 CCC ATG CCG CCT TCG GGG CTG CGG CTC TTG CCG CTG CTG CTG CCG CTG CTG TGG CTG CTA GTG CTG ACG CCT GGC

       Arg Pro Ala Ala Gly Leu Ser Thr Cys Lys Thr Ile Asp Met Glu Leu Val Lys Arg Lys Arg Ile Glu Ala Ile
 476 CGG CCG GCC GCC GGA CTG TCC ACC TGC AAG ACC ATC GAC ATG GAG CTG GTG AAG CGG AAG CGC ATC GAG GCC ATT

  50 Arg Gly Gln Ile Leu Ser Lys Leu Arg Leu Ala Ser Pro Pro Ser Gln Gly Asp Val Pro Pro Gly Pro Leu Pro
 551 CGC GGC CAG ATT CTG TCC AAG CTT CGG CTC GCC AGC CCC CCG AGC CAG GGG GAC GTG CCG CCC GGC CCG CTG CCT

       Glu Ala Val Leu Ala Leu Tyr Asn Ser Thr Arg Asp Arg Val Ala Gly Glu Ser Val Glu Pro Glu Pro Glu Pro
 626 GAG GCC GTA CTG GCT CTT TAC AAC AGT ACC CGC GAC CGG GTA GCC GGG GAA AGT GTC GAA CCG GAG CCC GAG CCA

 100 Glu Ala Asp Tyr Tyr Ala Lys Glu Val Thr Arg Val Leu Met Val Glu Ser Gly Asn Gln Ile Tyr Asp Lys Phe
 701 GAG GCG GAC TAC TAC GCC AAG GAG GTC ACC CGC GTG CTA ATG GTG GAA AGC GGC AAC CAA ATC TAT GAT AAA TTC

       Lys Gly Thr Pro His Ser Leu Tyr Met Leu Phe Asn Thr Ser Glu Leu Arg Glu Ala Val Pro Glu Pro Val Leu
 776 AAG GGC ACC CCC CAC AGC TTA TAT ATG CTG TTC AAC ACG TCG GAG CTC CGG GAA GCG GTG CCG GAA CCT GTA TTG

 150 Leu Ser Arg Ala Glu Leu Arg Leu Leu Arg Leu Lys Leu Lys Val Glu Gln His Val Glu Leu Tyr Gln Lys Tyr
 851 CTC TCT CGG GCA GAG CTG CGC CTG CTG AGG CTC AAG TTA AAA GTG GAG CAG CAC GTG GAG CTA TAC CAG AAA TAC

       Ser Asn Asp Ser Trp Arg Tyr Leu Ser Asn Arg Leu Leu Ala Pro Ser Asp Ser Pro Glu Trp Leu Ser Phe Asp
 926 AGC AAT GAT TCC TGG CGC TAC CTC AGC AAC CGG CTG CTG GCC CCC AGT GAC TCA CCG GAG TGG CTG TCC TTT GAT

 200 Val Thr Gly Val Val Arg Gln Trp Leu Thr Arg Arg Glu Ala Ile Glu Gly Phe Arg Leu Ser Ala His Cys Ser
1001 GTC ACC GGA GTT GTG CGG CAG TGG CTG ACC CGC AGA GAG GCT ATA GAG GGT TTT CGC CTC AGT GCC CAC TGT TCC

       Cys Asp Ser Lys Asp Asn Thr Leu His Val Glu Ile Asn Gly Phe Asn Ser Gly Arg Arg Gly Asp Leu Ala Thr
1076 TGT GAC AGC AAA GAT AAC ACA CTC CAC GTG GAA ATT AAC GGG TTC AAT TCT GGC CGC CGG GGT GAC CTG GCC ACC

 250 Ile His Gly Met Asn Arg Pro Phe Leu Leu Leu Met Ala Thr Pro Leu Glu Arg Ala Gln His Leu His Ser Ser
1151 ATT CAC GGC ATG AAC CGG CCC TTC CTG CTC CTC ATG GCC ACC CCG CTG GAG AGG GCC CAG CAC CTG CAC AGC TCC

       Arg His Arg Arg Ala Leu Asp Thr Asn Tyr Cys Phe Ser Ser Thr Glu Lys Asn Cys Cys Val Arg Gln Leu Tyr
1226 CGG CAC CGC CGA GCC CTG GAT ACC AAC TAC TGC TTC AGC TCC ACG GAG AAG AAC TGC TGC GTG CGG CAG CTC TAC

 300 Ile Asp Phe Arg Lys Asp Leu Gly Trp Lys Trp Ile His Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Leu Gly
1301 ATT GAC TTC CGG AAG GAC CTG GGC TGG AAG TGG ATT CAT GAA CCC AAG GGC TAC CAT GCC AAT TTC TGC CTG GGG

       Pro Cys Pro Tyr Ile Trp Ser Leu Asp Thr Gln Tyr Ser Lys Val Leu Ala Leu Tyr Asn Gln His Asn Pro Gly
1376 CCC TGT CCC TAC ATC TGG AGC CTA GAC ACT CAG TAC AGC AAG GTC CTG GCT CTG TAC AAC CAG CAC AAC CCG GGC

 350 Ala Ser Ala Ala Pro Cys Cys Val Pro Gln Ala Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly Arg Lys Pro
1451 GCG TCG GCG GCG CCG TGC TGC GTG CCG CAG GCG CTG GAG CCA CTG CCC ATC GTG TAC TAC GTG GGC CGC AAG CCC

       Lys Val Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys Lys Cys Ser
1526 AAG GTG GAG CAG CTG TCC AAC ATG ATC GTG CGT TCC TGC AAG TGC AGC TGA  GGCCCCGCCCCGCCCACAGCCCCGCCCAC
```