Volume 14 Number 2 1986 Nucleic Acids Research

Structure and nucleotide sequence of the 5′ region of the human and feline c-*sis* proto-oncogenes

Ans M.W.Van den Ouweland, Anton J.M.Roebroek, Jack A.Schalken[1], Chris A.A.Claesen[2], Henri P.J.Bloemers and Wim J.M.Van de Ven

Molecular Oncology Section, Department of Biochemistry, University of Nijmegen, Geert Grooteplein Noord 21, 6525 EZ Nijmegen and Departments of [1]Urology and [2]Organic Chemistry, University of Nijmegen, Nijmegen, The Netherlands

Received 21 October 1985; Revised and Accepted 16 December 1985

ABSTRACT

Comparative analysis of cosmid clones containing the human and feline c-sis genetic regions revealed the similar structural organization of these areas in the two species. The areas shared seven different genetic regions in and around the c-sis locus and one of these was related to v-sis. Another region, 1.9 kbp in size and located about 8 kbp upstream of the v-sis homologous region in the human genome, also hybridized to the main c-sis transcriptional product of 3.5 kb. Comparison with a recently described c-sis cDNA clone (Collins et al., Nature 316, 748-750 (1985)) revealed that the 1.9 kbp DNA region contained a large 5′ c-sis exon of at least 1050 bp. In this exon, the presumed initiation site of the predicted PDGF-2 containing precursor protein was located and appeared to be preceded by a large untranslated region. In the region immediately upstream of this exon, a TATA box and a consensus sequence for a potential Sp1 binding site were found at similar positions in both species. This region also exhibited promoter activity when tested in an assay in which coding sequences of bacterial chloramphenicol acetyltransferase (CAT; acetyl-CoA: chloramphenicol 3-O-acetyltransferase, EC 2.3.1.28) were placed under its control. The five other DNA regions were found upstream and downstream of the human c-sis transcription unit and also in an intron. Four of them contained repetitive sequences.

Hybridization analysis of human and feline c-sis containing cosmid clones with a mixed synthetic nucleotide probe, which corresponded to sequences encoding amino acid residues 2-7 of chain 1 of platelet-derived growth factor (PDGF-1), suggested that the c-sis cosmid clones did not include PDGF-1-specific genetic sequences.

INTRODUCTION

The human c-sis proto-oncogene contains coding sequences for a precursor protein, part of which is similar if not identical to chain 2 of platelet-derived growth factor (PDGF-2) (1,2). The initial link between PDGF-2 and sis was provided by partial amino acid sequence analysis of PDGF and nucleotide sequence analysis of the v-sis oncogene of simian sarcoma virus (3-5). Genetic sequences encoding PDGF-1 remain to be identified. Expression of c-sis is found in endothelial cells (6-8), human placenta (9), mouse embryo tissue (10), human lung tumors (11) and a number of tumor cell lines (12).

As the main transcriptional product, a mRNA species of about 3.5 kb (6,8,11,12), with coding potential for a protein of about 27 kd, was described (8,13). The precise function of the predicted c-sis encoded protein remains to be defined but it probably exhibits mitogen activity since it resembles PDGF with respect to amino acid sequence and also antigen and receptor binding characteristics (1,2,14-16). The responsiveness of fibroblasts, mesenchymal cells and glial cell lines to the mitogen PDGF and also the stimulation by PDGF of DNA synthesis during wound healing is well established (17-19).

It was recently reported that expression of the c-sis gene is modulated during endothelial cell differentiation in vitro (7). Furthermore, it was shown in transfection experiments, that expression in mouse embryo fibroblasts of human cDNA and genomic DNA constructs of c-sis resulted in a similar cell transformation (20,21) as observed upon transfection of the viral oncogene v-sis (22). Genetic sequences within or in close proximity of the locus may be involved in the control of such modulation of expression and they may also be relevant to the malignant potential of the locus. In an attempt to identify such regulatory sequences, we have compared the genetic organization of the c-sis gene and its flanking sequences in two different species, man and cat. In this report, we present the identification and characterization of seven genetic regions that human and cat share in and around the c-sis transcription unit.

MATERIALS AND METHODS

Cosmid and plasmid clones, E.coli strains and cell lines: Isolation of the human and feline c-sis cosmid clones was described previously (11). pAO68, pAO70, pAO73, pAO121, pAO151, pAO154 and pAO155 are subclones of the human c-sis locus (see also Fig. 2). pAO70, pAO73 and pAO121 were described elsewhere (11). pAO68 consists of a 3.8 kbp EcoRI/EcoRI DNA fragment subcloned in pSVBR91. pAO154 and pAO155 consist, respectively, of a 1.3 kbp BamHI/EcoRI and a 4.3 kbp BamHI/XhoI restriction fragment subcloned in pUC18. pAO151 consists of a 1.4 kbp EcoRI/HindIII restriction fragment subcloned in pAT153. pAO144 is a recombinant of pUC18 containing a 3.5 kbp HindIII/XhoI restriction fragment of the feline c-sis clone MB65. pSV2 was described by Gorman et al. (23) and pSuperCAT, a derivative of pSV2, was obtained from Dr. B. Dekker, University of Leiden, The Netherlands. pAO165 and pAO166 are recombinants of pSuperCAT containing an 0.4 kbp PstI/PstI human DNA fragment that is located immediately upstream of the 5′ c-sis exon

sequences. The insert in pAO166 has the same orientation relative to the chloramphenicol acetyltransferase (CAT) coding sequences as to the c-sis coding sequences in the human genomic DNA. The orientation of the insert in pAO165 is opposite to that in pAO166. For the DNA sequence analysis, the DNA fragments were inserted into the polylinker site of M13mp8-11 (24). E.coli strain HB101 was used as a host for the human and feline c-sis subclones. The recombinant M13 bacteriophages were propagated in E.coli strain JM101. Cell lines used in this study included HeLa (American Type Culture Collection (ATCC) CCL 2), Vero (ATCC CCL 81), HUT 102 (12) and Neuro 2A (ATCC CCL 131).

Preparation of DNA probes and hybridization: Preparation of DNA probes and their labeling by nick translation was carried out as described (25). The specific activity of the probes used in the hybridization studies was (2-5)$\times 10^8$ cpm/µg. Agarose gel electrophoresis, Southern blotting and hybridization analysis was performed as described (25). The mixed oligonucleotide probe was synthesized as described (26) and the probe was labeled with γ-[^{32}P]-ATP (3000 Ci/mmol) according to the procedure described by Maxam and Gilbert (27) to a specific activity of about 2×10^8 cpm/µg. Hybridization analysis with the mixed oligonucleotide probe was performed at 35 °C for 16 h in 6×SSC, 5×Denhardt's solution, 100 µg/ml denatured salmon testis DNA and 2×10^6 cpm/ml labeled mixed probe. Following hybridization, filters were washed in 5 mM EDTA pH 8.0 and 6×SSC during one hour periods at 37 °C, 40 °C, and 45 °C, successively, and dried. Autoradiography was performed by exposure to XAR-2 film (Kodak) with Dupont lightning plus intensifying screens.

DNA sequence analysis: Sequencing of DNA fragments was according to the dideoxy method of Sanger et al. (28). The gels readings were recorded, edited and compared using the Staden programs (29).

RNA isolation, Northern blotting and hybridization: Total cellular RNA was isolated using the lithium-urea procedure described by Auffray and Rougeon (30). Ten µg of mRNA, purified by oligo(dT)-cellulose affinity chromatography, was glyoxalated, fractionated on 1.0 % agarose gels (31) and transferred to Hybond-N (Amersham). Hybridization of the Northern blots was carried out as described by Church and Gilbert (32).

Assay for chloramphenicol acetyl transferase activity: The assay for chloramphenicol acetyltransferase (CAT) activity was performed as described by Gorman et al. (23).

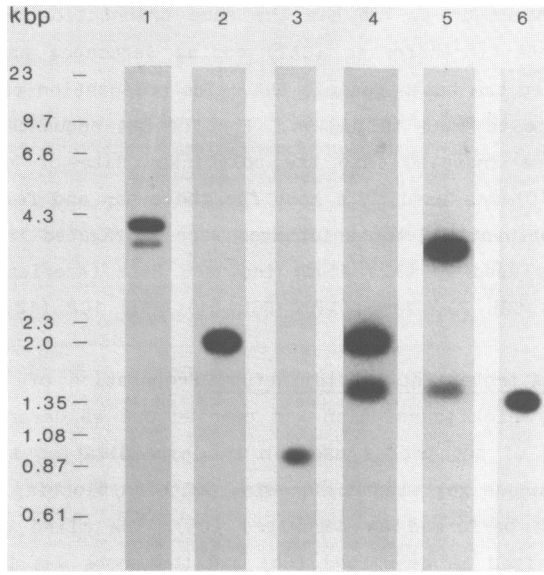


Fig. 1. Identification of common genetic sequences in the inserts of the human and feline c-sis cosmid clones. MB65 (lanes 1, 2 and 3) and MB70 (lanes 4, 5 and 6) were digested with restriction endonucleases EcoRI and KpnI (lane 1), BamHI and XbaI (lane 2), HindIII and BamHI (lane 3), HindIII, BamHI and XhoI (lanes 4 and 5) and EcoRI and XbaI (lane 6). As molecular probes, the inserts of the following human c-sis subclones were used: pA0155 (lane 1), pA0151 (lane 2), pA0154 (lane 3), pA0121 (lane 4), pA070 (lane 5) and pA068 (lane 6). Molecular weight markers include λ DNA digested with restriction endonuclease HindIII and DNA of ΦX174 digested with HaeIII.

RESULTS

Identification and characterization of seven genetic regions shared by the human and feline c-sis loci.

As an initial approach to identify potential regulatory regions within or in close proximity of the c-sis locus, we compared previously described human and feline c-sis containing cosmid clones (11) by Southern blot analysis. We reasoned that genetic regions which modulated expression of c-sis and which might be relevant to the malignant potential of the locus were likely to be conserved during evolution. A number of DNA fragments could indeed be identified when ^{32}P-labeled feline c-sis cosmid clones were hybridized with the human c-sis containing clones (data not shown). Following molecular cloning of 11 DNA fragments from the human c-sis containing cosmid clones and using them as probes, seven different genetic regions which were shared by man and cat could be identified (Fig. 1). One of these regions

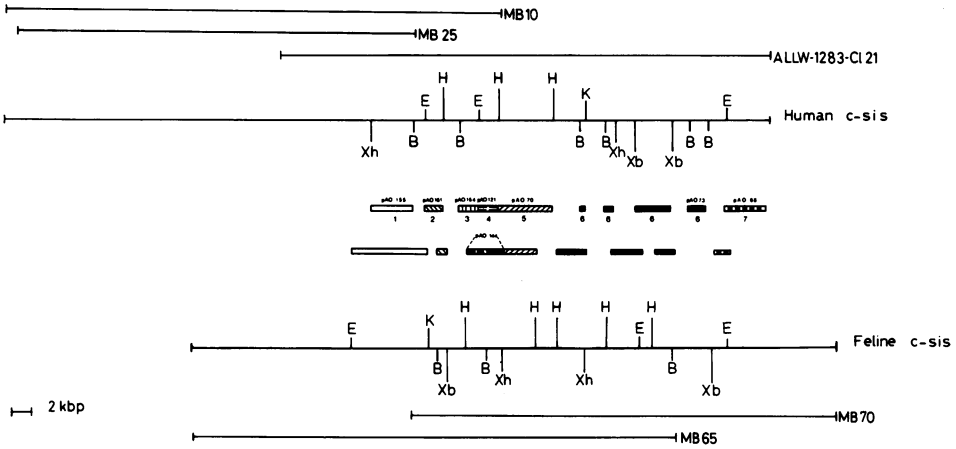


Fig. 2. Homologous regions in and around the human and feline c-sis locus. In the upper part of the figure, the size and relative positions of the DNA inserts of three human cosmid clones (ALLW-1283-Cl21, MB10 and MB25) are shown. Directly below these inserts, a schematic restriction endonuclease map of the human c-sis containing DNA region is presented. At the bottom of the figure, the size and relative positions of the cellular DNA inserts of two feline cosmid clones (MB65 and MB70) are given. Immediately above them, a schematic restriction endonuclease map of the feline c-sis containing DNA region is presented. Between the two restriction maps, the seven regions that man and cat share in this area are depicted as shaded boxes. Boxes of corresponding regions are shaded similarly and overlapping regions are shaded accordingly. The human regions are numbered (1-7). Human region 6, which represents the v-sis homologous region, is subdivided in four v-sis homologous fragments which all belong to region 6. The positions of the DNA inserts of recombinant plasmids are indicated by the names of the plasmids above the shaded boxes.
B, BamHI; E, EcoRI; H, HindIII; K, KpnI; Xb, XbaI; Xh, XhoI. (For a detailed restriction endonuclease map see Ref. 11).

(region 6) was homologous to v-sis (see also ref. 11) and a unique c-sis-specific probe (pA073) was prepared from the 3′ part of this region. In Fig. 2, the localization of the seven DNA regions within the human and feline c-sis containing genetic area is depicted.

To define these seven regions in and around the c-sis locus of the two species in more detail, we have further analyzed the molecular cloned human DNA fragments. It appeared from Southern blot analysis that region 1, 2, 3, 6 and 7 contained repetitive sequences (data not shown). The presence of Alu-repeats in region 6 (33) was confirmed. In region 1, 2, 3 and 7, no hybridization was observed when total human DNA was used as a probe. The repetitive sequences in these regions, therefore, do not belong to the

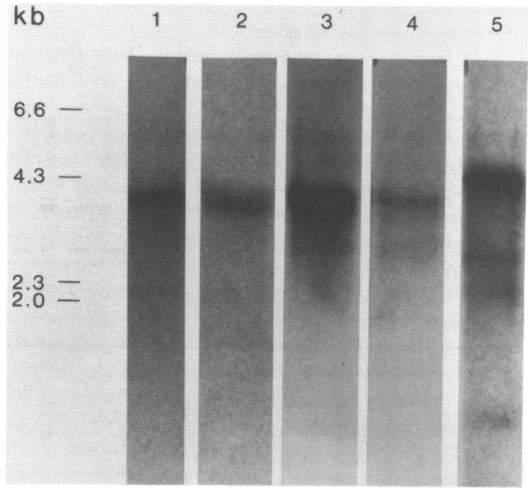


Fig. 3. Identification of c-sis transcripts in a number of cell lines. Poly(A)-selected RNA was isolated from HeLa (lanes 1 and 2), Vero (lane 3), HUT 102 (lane 4) and Neuro 2A (lane 5) cells and screened (10 µg aliquots) in Northern blot analysis with the inserts of pA0121 (lane 1), pA073 (lanes 2, 3, 4 and 5) as molecular probes. Molecular weight markers include λ DNA digested with the restriction endonuclease HindIII.

category of highly repetitive sequences (data not shown). Region 4 contained only unique genetic sequences. In Northern blot analysis of poly(A)-selected RNA isolated from HeLa cells, a 3.5 kb transcript was observed using pA073 (c-sis-specific) or pA0121 (containing region 4 as insert) as molecular probes (Fig. 3, lanes 1 and 2). Transcripts of similar size were also detected with these probes in a HTLV-I infected human T cell line (HUT102) (Fig. 3, lane 4) and the Vero cell line (Fig. 3, lane 3). The c-sis-related transcript in murine neuroblastoma cells was slightly larger (3.4) (Fig. 3, lane 5). Region 5 also contained only unique DNA sequences but it did not hybridize to the 3.5 kb c-sis transcript (data not shown).

DNA sequence analysis of human region 4 and its feline counterpart.

To resolve the question to what extent the genetic sequences in the unique human region 4 and the corresponding region in the cat genome were homologous, we have performed DNA sequence analysis. The complete nucleotide sequence of human region 4 and most of its feline counterpart were compared (Fig. 4). A remarkable homology between the human and feline DNA could be observed. Nucleotides printed in capitals represent sequences also found in a recently published cDNA clone of human c-sis expressed in endothelial

cells (8) and they can therefore be considered as exon sequences. As can be seen in Fig. 4, a continuous stretch of at least 1050 nucleotides constitute a single new exon. Homology between man and cat in this exon was 90 %. Together with the six already established c-sis exons (13) this new exon could form a mRNA transcript of about 3.5 kb, a size similar to that of the main c-sis transcript observed in a number of human cells (6,8,11,12) (see also Fig. 3). At its 3' end, a consensus splice junction was found. The 5' end of the exon remains to be defined. Within the human and feline exon sequences, four initiation codons were present. Three of them did not seem to be functional since they were followed almost immediately by a stop codon. The fourth ATG codon, which was found 63 nucleotides upstream of the 3' splice junction, is in-phase with the PDGF-2 open reading frame in the following exons (8,13). The 63 bp coding region of the new exon contained genetic sequences for a hydrophobic amino terminal sequence of the predicted precursor protein of PDGF-2.

Within the 5' portion of the human exon, a triplet of repeats (GCAGCTC) was followed by a doublet of a similar repeat (GCAGCCC). Interestingly, one of the repeats of the triplet was deleted in the c-sis cDNA clone published by Collins et al. (8). Five other differences, all point mutations, between human region 4 and this c-sis cDNA clone could be observed, namely two single nucleotide substitutions (an A instead of a G at position 683 and a T instead of an C at position 1383), two deletions (position 580 and 1151) and one insertion (a C at position 1262) (Fig. 4).

The exon sequences in human region 4 were preceded by a TATAAA sequence, the consensus promoter sequence for RNA transcription of eucaryotic genes (Fig. 4) (35). At the same position in the feline DNA, the same TATAAA sequence was found (Fig. 4). The observation that the 3.5 kb c-sis transcripts observed in a number of cells have the same length as the combined overlapping sequences in a number of recently published cDNA clones (8,13) is in support of the possibility that transcription of the c-sis locus starts in close proximity of this TATAAA containing region. In addition to the TATAAA consensus sequence, some other promoter-specific sequences were present upstream of the TATA box (see Discussion).

To test this potential promoter region for promoter activity, the 0.4 kbp PstI/PstI human DNA fragment, which contained the TATAAA box and the other promoter-specific sequences, was subcloned in pSuperCAT and the recombinant pA0166 was transfected into HeLa cells. As negative controls, pSuperCAT itself or pA0165 (pSuperCAT containing the 0.4 kbp PstI/PstI DNA fragment in

Fig. 4. Comparison of the nucleotide sequence of human sis region 4 with sequences of its feline counterpart. Exon sequences, based on sequence data of the sis cDNA clone isolated by Collins et al. (8), are printed in capitals. The first three ATG codons are underlined and asterisks (*) mark termination codons that are in-phase with a preceding ATG. The amino acid sequence of

the 63 bp open reading frame is given in the conventional one letter code. The putative TATA boxes of the human and feline locus are placed in a box. A closed circle (●) and an open square (□) mark the 5′ ends of, respectively, the human genomic c-sis subclone isolated by Gazit et al. (21) and the cDNA clone described by Ratner et al. (13). Two sets of small repetitive sequences are indicated with arrows (⟶,- - -⟶). The consensus sequence for the potential Sp1 binding sites and the CCGCCC sequence similar to the one found in the SV40 early promoter region are underlined.

an orientation opposite to that in pAO166) was used. As a positive control, pSV2 was included in the experiments (23). Measurements of CAT activity in HeLa extracts were made 48 hr after introduction of the DNA into the HeLa cells. Analysis of thin layer chromatograms revealed that there was a considerable amount of CAT enzymatic activity in extracts of HeLa cells transfected with pAO166. No such activity was detected in extracts from cells transfected with pSuperCAT or pAO165. These results indicated that the 0.4 kbp PstI/PstI human DNA fragment had the capacity to function as a promoter.

Analysis of the human and feline c-sis containing cosmid clones for the presence of PDGF-1 related sequences.

The observation that a number of DNA regions in and around the human and feline c-sis coding region, which encodes a protein similar if not identical to PDGF-2, appeared to be conserved, raised the possibility of the presence of genetic sequences related to those encoding PDGF-1. To test this possibility, we prepared a mixed synthetic oligonucleotide probe based upon the amino-terminal amino acid sequences (residues 2-7) (3,4,36) of PDGF-1. The composition of the synthetic probe is depicted in Fig. 5A. Hybridization analysis was positive in the case of a 5.0 kbp HindIII/HindIII DNA fragment of cosmid clone ALLW-1283-C121 (Fig. 5B, lane 3). None of the other human (Fig. 5B, lanes 1 and 2) or feline (data not shown) cosmid clones hybridized to the mixed synthetic probe. Upon cloning and further analysis of the 5.0 kbp HindIII/HindIII DNA fragment, the matching sequences were localized in a 1.0 kbp Sau3A/HindIII DNA fragment (Fig. 5C). Nucleotide sequence analysis of this fragment revealed a match of 15 out of 17 nucleotides of one of the probes in the mixture (Fig. 5D). However, this sequence appeared to be no part of a region encoding the amino-terminal portion of PDGF-1.

DISCUSSION

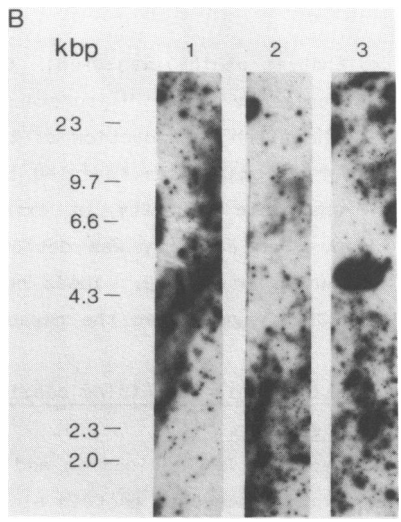
In this report, we describe the structure and the nucleotide sequence of the 5′ region of the human and feline c-sis proto-oncogenes and provide the

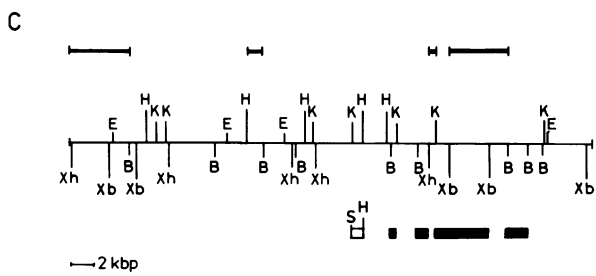
A

Chain 2 NH_2- S L G S L T I A E P A M I A E

Chain 1 NH_2- S I E E A V P A V

Mixed probe 3'-d(TA_T^{Pu} CTPy CTPy CGN CAN GG)-5'





D

Sau3A HpaII

GATCTTGGGACAGCTTCTTCAGGGTGCCAGGCCTCCTTTCCCATCTCTGAAGTGAGCTGTCCACCTGGAGGCCTGCGGAACCTGTGCCCAGGAAAACCAGGCTCCGG

::::::::::::::::

T T T

GGNACNGC TC TCAAT

C C G

Fig. 5. Identification of genetic sequences homologous to the synthetic mixed probe. (A) Composition of the mixed probe. The N-terminal 15 amino acids of PDGF-2 and 9 amino acids of PDGF-1 are aligned for maximum homology (3,36). The composition of the mixed probe is indicated under the amino acid sequence data. (B) Southern blot analysis of human c-sis cosmid clones with the mixed probe. Clones MB10 (lane 1), MB25 (lane 2) and ALLW-1283-C121

(lane 3) were digested with the restriction endonuclease HindIII. Hybridization of the Southern blots was as described under Materials and Methods. (C) Restriction endonuclease map of the human c-sis cosmid clone ALLW-1283-C121. The upper heavy bars represent highly repetitive DNA sequences. The next line represents a schematic restriction endonuclease map of the human c-sis locus. At the bottom of the figure, black boxes indicate the relative positions of v-sis homologous sequences and the open box the DNA region that hybridized with the mixed synthetic probe.
B, BamHI; E, EcoRI; H, HindIII; K, KpnI; S, Sau3A; Xb, XbaI; Xh, XhoI. (D) Nucleotide sequence of the first 107 bp of the Sau3A/HindIII DNA fragment illustrating the homology with the mixed synthetic probe.

first characteristics of their promoter regions. Comparative analysis of the human and feline c-sis containing DNA areas resulted in the identification of seven DNA regions which the two species have in common. These regions are highly homologous in the two species and their genetic organizations are colinear. One of the regions, region 6, constitutes the human and feline v-sis cellular homolog which we have described before (11). In region 4, the 5′ region of the transcription unit of the c-sis proto-oncogene is located. We conclude this from the following observations. Combination of the overlapping inserts of the human sis cDNA clones (8,13) constitutes a stretch of DNA of about 3.5 kbp which resembles the size of the c-sis mRNA species detectable in a number of cell lines. This means that the large exon described in this paper contains almost all the 5′ genetic sequences of the messenger although the precise position of the cap site remains to be identified. Further support for our conclusion is based upon the fact that a TATA box is present in the DNA region immediately upstream of the large exon. Furthermore, this region exhibited promoter activity in an assay for transcriptional activity. The TATAAA consensus promoter sequence for RNA transcription of eukaryotic genes is present at the same positions in the c-sis loci of the two species. Since the distance in many eukaryotic genes between the TATA box and the cap site of the mRNA is about 30 nucleotides, it can be concluded that from the 5′ sis cDNA clone described by Collins et al. (8) only a few nucleotides of the 5′ region of the mRNA are excluded. In both species, additional promoter-specific consensus sequences are found upstream of the TATA box. A (G+C)-rich region is located about 80 nucleotides upstream of the putative RNA transcriptional starting region and the characteristic CCGCCC sequence (position 259-264 in Fig. 4) is present within this region. In the SV40 early promoter region, six copies of this sequence were found in a similar (G+C)-rich region (37) and in the myc proto-oncogene the hexanucleotide is located in close proximity to the TATA box (38).

At about 200 nucleotides upstream of the TATA box a GC box consensus sequence (GGGGCGGGAC in the human and GGGGCGGGGC in the feline DNA) for the binding site of the promoter-specific factor Sp1 (39) is present. A number of GC box-containing promoters have been described (39) and they include cellular as well as viral genes. However, it remains to be established whether the GC box in the sis locus is involved in Sp1 binding.

A striking characteristic of the 5' c-sis exon is its long noncoding region (at least 987 bp). The messenger of the insulin-like growth factor II precursor has a similar long 5' untranslated region (40). It is possible that the long 5' untranslated region of the c-sis locus plays a regulatory role but this needs further studies. The presence of four ATG initiation codons was also remarkable. Most likely, translation starts at the fourth initiation site (position 1417). This ATG codon meets the criteria postulated by Kozak (41,42) and it also is the beginning of an open reading frame of 723 bp (8,13) that encodes the predicted PDGF-2 precursor protein of about 27 kd.

The other five genetic regions which man and cat have in common in and around the c-sis transcription unit are not yet fully characterized. It is clear from our data that at least one of these regions lies in an intron and that its sequences are more conserved than other intron regions. This intron region is not present in the main c-sis transcript (data not shown). Furthermore, it is of interest to note that some of these regions contain repetitive DNA sequences. The presence of such repetitive sequences in the proximity of cellular coding sequences was reported before (43) and could have functional implications.

A still unresolved issue pertains to where the genetic sequences that encode the PDGF-1 chain are located. We have tested the possibility that they are present in the DNA areas described in this study. Southern blot analysis with a PDGF-1-specific synthetic mixed probe did not reveal any potential region. However, it should be noted that in the case that a splice site would be located in the sequences recognized by the mixed probe, a match could have escaped detection. Our results suggest that the PDGF-1 encoding sequences are located somewhere else, even on another chromosome.

ACKNOWLEDGEMENTS

We thank Carla Onnekink, Jan van Groningen and Anja ter Avest for their excellent technical assistance and Lambert Dorssers for helpful discussions. This work was supported by the Netherlands Cancer Foundation, Contract No. NUKC Bioch. 85-9.

REFERENCES

1. Chiu, I.M., Reddy, E.P., Givol, D., Robbins, K.C., Tronick, S.R. and Aaronson, S.A. (1984) Cell 37, 123-129.
2. Johnsson, A., Heldin, C.H., Wasteson, A., Westermark, B., Deuel, T.F., Huang, J.S., Seeburg, P.H., Gray, A., Ullrich, A., Scrace, G., Stroobant, P. and Waterfield, M.D. (1984) EMBO J. 5, 921-928.
3. Doolitle, R.F., Hunkapillar, M.W., Hood, L.E., Devare, S.G., Robbins, K.C., Aaronson, S.A. and Antoniades, H.N. (1983) Science 221, 275-277.
4. Waterfield, M.D., Scrace, G.T., Whittle, N., Stroobant, P., Johnsson, A., Wasteson, A., Westermark, B., Heldin, C.H., Huang, J.S. and Deuel, T.F. (1983) Nature 304, 35-39.
5. Robbins, K.C., Antoniades, H.N., Devare, S.G., Hunkapiller, M.W. and Aaronson, S.A. (1983) Nature 305, 605-608.
6. Barrett, T.B., Gajdusek, C.M., Schwartz, S.M., McDougall, J.K and Benditt, E.P. (1984) Proc. Natl. Acad. Sci. USA 81, 6772-6774.
7. Jaye, M., McConathy, E., Drohan, W., Tong, B., Deuel, T. and Maciag, T. (1985) Science 228, 882-885.
8. Collins, T., Ginsburg, D., Boss, J.M., Orkin, S.H. and Pober, J.S. (1985) Nature 316, 748-750.
9. Goustin, A.S., Betsholtz, C., Pfeifer-Ohlsson, S., Persson, H., Rydnert, J., Bywater, M., Holmgren, G., Heldin, C.H., Westermark, B., Ohlsson, R. (1985) Cell 41, 301-312.
10. Slamon, D.J. and Cline, M.J. (1984) Proc. Natl. Acad. Sci. USA 81, 7141-7145.
11. Van den Ouweland, A.M.W., Breuer, M.L., Steenbergh, P.H., Schalken, J.A., Bloemers, H.P.J. and Van de Ven, W.J.M. (1985) Biochim. Biophys. Acta 825, 140-147.
12. Eva, A., Robbins, K.C., Andersen, P.R., Srinivasan, A., Tronick, S.R., Reddy, E.P., Ellmore, N.W., Galen, A.T., Lautenberger, J.A., Papas, T.S., Westin, E.H., Wong-Staal, F., Gallo, R.C. and Aaronson, S.A. (1982) Nature 295, 116-119.
13. Ratner, L., Josephs, S.F., Jarrett, R., Reitz, M.S. and Wong-Staal, F. (1985) Nucl. Acids Res. 13, 5007-5018.
14. Betsholtz, C., Westermark, B., Ek, B. and Heldin, C.H. (1984) Cell 39, 447-457.
15. Nister, M., Heldin, C.H., Wasteson, A. and Westermark, B. (1984) Proc. Natl. Acad. Sci. USA 81, 926-930.
16. Stiles, C.D. (1983) Cell 33, 653-655.
17. Ross, R., Glomset, J., Kariya, B. and Harker, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1207-1210.
18. Scher, C.D., Shepard, R.C., Antoniades, H.N. and Stiles, C.D. (1979) Biochim. Biophys. Acta 560, 217-241.
19. Heldin, C.H., Westermark, B. and Wasteson, A. (1980) J. Cell. Physiol. 105, 235-246.
20. Clarke, M.F., Westin, E., Schmidt, D., Josephs, S.F., Ratner, L., Wong-Staal, F., Gallo, R.C. and Reitz, M.S., Jr. (1984) Nature 308, 464-467.
21. Gazit, A., Igarashi, H., Chiu, I.M., Srinivasan, A., Yaniv, A., Tronick, S.R., Robbins, K.C. and Aaronson, S.A. (1984) Cell 39, 89-97.
22. Hannik, M. and Donoghue, D.J. (1984) Science 226, 1197-1199.
23. Gorman, C.M., Merlino, G.T., Willingham, M.C., Pastan, I. and Howard, B.H. (1982) Proc. Natl. Acad. Sci. USA 79, 6777-6781.
24. Messing, J. and Vieira, J. (1982) Gene 19, 269-276.
25. Schalken, J.A., Van den Ouweland, A.M.W., Bloemers, H.P.J. and Van de Ven, W.J.M. (1985) Biochim. Biophys. Acta 824, 104-112.
26. Claesen, C., Tesser, G.I., Dreef, C.E., Marugg, J.E., Van der Marel, G.A. and Van Boom, J.H. (1984) Tetrahedron Letters 25, 1307-1310.

27. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
28. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
29. Staden, R. (1982) Nucl. Acids Res. 10, 2951-2961.
30. Auffray, C. and Rougeon, F. (1980) Eur. J. Biochem. 107, 303-314.
31. McMaster, G.K. and Carmichael, G.G. (1977) Proc. Natl. Acad. Sci. USA 74, 4835-4838.
32. Church, G.M. and Gilbert, W. (1984) Proc. Natl. Acad. Sci. USA 81, 1991-1995.
33. Favera, R.D., Gelmann, E.P., Gallo, R.C. and Wong-Staal, F. (1981) Nature 292, 31-35.
34. Van Zoelen, E.J.J., Van de Ven, W.J.M., Franssen, H.J., Oosterwaard, T.M.J., Van der Saag, P.T., Heldin, C.H. and De Laat, S.W. (1985) Mol. Cell. Biol. 5, 2289-2297.
35. Breathnach, R. and Chambon, P. (1981) Ann. Rev. Biochem. 50, 349-383.
36. Antoniades, H.N. and Hunkapiller, M.W. (1983) Science 220, 963-965.
37. Fromm, M. and Berg, P. (1982) J. Mol. Appl. Genet. 1, 457-481.
38. Ishii, S., Xu, Y.H., Stratton, R.H., Roe, B.A., Merlino, G.T. and Pastan, I. (1985) Proc. Natl. Acad. Sci. USA 82, 4920-4924.
39. Dynan, W.S. and Tjian, R. (1985) Nature 316, 774-778.
40. Dull, T.J., Gray, A., Hayflick, J.S. and Ullrich, A. (1984) Nature 310, 777-781.
41. Kozak, M. (1984) Nucl. Acids Res. 12, 857-872.
42. Kozak, M. (1984) Nature 308, 241-146.
43. Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M. Slightom, J.L., Blechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980) Cell 21, 653-668.