Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 2392–2396, April 1986
Biochemistry

Structure and sequence of the human c-*sis*/platelet-derived growth factor 2 (*SIS/PDGF2*) transcriptional unit

(protooncogene mRNA/growth factor)

C. DURGA RAO, HISANAGA IGARASHI*, ING-MING CHIU, KEITH C. ROBBINS, AND STUART A. AARONSON

Laboratory of Cellular and Molecular Biology, Building 37, Room 1E24, National Cancer Institute, Bethesda, MD 20892

Communicated by H. Hanafusa, December 18, 1985

ABSTRACT **The structure of the normal human c-*sis*/platelet-derived growth factor 2 (*SIS/PDGF2*) transcript was determined by a combination of cDNA cloning, nuclease S1 mapping, and primer extension. Nucleotide sequence analysis revealed that the 3373-nucleotide SIS/PDGF2 mRNA contained only a 723-base-pair (bp) coding sequence for the PDGF2 precursor polypeptide. The coding sequence was flanked by long 5′ (1022 bp) and 3′ (1625 bp) untranslated regions. The 5′ noncoding region, as well as upstream flanking genomic sequences, contained clusters of specific short repeat sequences. A consensus transcriptional promoter sequence, TATAAA, was identified 24 bp upstream of the mRNA start site and an enhancer-like "TG element" was detected about 180 bp downstream from the site of polyadenylylation. These findings identify putative regulatory elements of the *SIS/PDGF2* gene.**

The simian sarcoma virus transforming gene, v-*sis*, encodes a polypeptide that is closely related in its predicted amino acid sequence and processing to human platelet-derived growth factor (PDGF) (for review, see ref. 1). Human PDGF preparations contain two related but distinct polypeptide chains, only one of which, PDGF2, is homologous to the v-*sis* gene product. Like human PDGF, the processed dimeric forms of the v-*sis/PDGF2* gene product have been shown to bind to the PDGF receptor, trigger its phosphorylation at tyrosine residues, and specifically stimulate DNA synthesis of cells possessing such receptors (2). Moreover, only those cell types possessing PDGF receptors are susceptible to growth alterations induced by the v-*sis* transforming gene (2). Expression of the normal human c-*sis/PDGF2* gene (*SIS/PDGF2*) coding sequence under conditions leading to synthesis of a PDGF2 dimer has also been shown to cause malignant transformation of NIH 3T3 cells (3). These findings have indicated that expression of this normal human growth factor in a cell responsive to its growth-promoting activity can lead to transformation.

The human *SIS/PDGF2* transcript is expressed in some glioblastomas and fibrosarcomas but not in the normal counterparts of such tumors (4). Since fibroblasts and glial cells possess PDGF receptors and are susceptible to the growth-stimulating activity of PDGF (5–7), it is possible that activation of *SIS* expression may play a role in the development of such tumors. Efforts to investigate control of human *SIS/PDGF2* gene expression in normal and tumor cells have focused on attempts to isolate full-length *SIS/PDGF2*-complementary DNA (8, 9) for use in localizing regulatory elements adjacent to the transcribed sequence. In the present study, we report the structure and sequence of the entire normal *SIS/PDGF2* transcript and localize putative regulatory elements of the gene.

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.

MATERIALS AND METHODS

Tissues, Cells, and Enzymes. Frozen full-term, normal-delivery human placenta was obtained from Advanced Biotechnologies. The EJ cell line has been described (10). A human ovarian carcinoma-derived tumor cell line, A2780 (N. W. Ellmore and S.A.A., unpublished work) was also utilized.

Construction and Screening of Placental cDNA Library. RNA was extracted by the method of Chirgwin *et al.* (11). Poly(A)$^+$ RNA was selected by passage over an oligo(dT)-cellulose column (Collaborative Research, Waltham, MA). The method of Okayama and Berg (12) was used to construct a placental cDNA library. A genomic 880-base-pair (bp) fragment representing the 3′ half of the 1.9-kilobase (kb) *Pst* I fragment from human λ-c-*sis* clone 8 (13) was used for screening according to the method described (14).

Nuclease S1 Analysis and Primer Extension. Nuclease mapping using poly(A)$^+$ RNA was carried out as described (15), except that the hybridizations were at 56°C and digestions were at a concentration of 200 units/μl. Synthetic oligonucleotide-primed cDNA was synthesized using avian myeloblastosis virus reverse transcriptase in the presence of ^{32}P-labeled deoxynucleoside triphosphates; synthetic oligonucleotides were from Pharmacia.

DNA Sequence Analysis. Nucleotide sequence was determined by the partial chemical degradation method of Maxam and Gilbert (16).

RESULTS

Isolation of Normal Human SIS/PDGF2 cDNA Clones. In the investigation of human *SIS* expression in normal tissues, we found placenta to be among the richest sources of SIS/PDGF2 mRNA. Thus, a human placental cDNA library was constructed and screened with a probe that contained human sequences located downstream of the v-*sis*-related regions of λ-c-*sis* clone 8 (13). This probe, designated pS880, detected SIS mRNA expressed in placenta but did not detect simian sarcoma virus RNA (data not shown). Among 120,000 recombinants analyzed, 15 initially were scored as positive with the pS880 probe. Analysis of recombinant clones revealed five independent cDNAs ranging in size from 1200 to 2600 bp. When these cDNAs were hybridized with a v-*sis* DNA probe, all but the shortest were readily detected. Restriction enzyme mapping demonstrated that the five clones were overlapping and contained restriction enzyme sites corresponding to those also found in exons of the human *SIS* protooncogene (13, 17, 18). Thus, all of the cDNA clones isolated represented the human *SIS/PDGF2* transcript. The longest cDNA clone was designated pSD1.

Abbreviations: PDGF, platelet-derived growth factor; bp, base pair(s); kb, kilobase(s).

*Present address: Department of Bacteriology, Nagasaki University School of Medicine, Nagasaki 852, Japan.

Identification of the SIS/PDGF2 mRNA Start Site in Human Genomic DNA. Difficulties have been encountered by ourselves and others (8, 9) in obtaining the entire human SIS/PDGF2 cDNA. Evidence that the upstream region of the *SIS/PDGF2* transcript is highly G+C-rich (9, 13) suggested that strong secondary mRNA structure impaired complete cDNA extension. To circumvent this problem, we chose to search for the SIS/PDGF2 mRNA start site in genomic DNA. Our strategy involved hybridization of either uniformly or end-labeled genomic DNA fragments with RNA from cells expressing the *SIS/PDGF2* transcript. The extent and location of transcribed regions were determined by electrophoretic analysis of hybrid molecules in denaturing gels after treatment with nuclease S1 (15).

We previously identified an upstream human *SIS* exon not related to v-*sis* within human genomic DNA. This segment, designated pc-*sis* RH-9, was localized ≈10 kb upstream from the first v-*sis*-related exon of the human *SIS/PDGF2* gene (3). The 5′ terminus of pSD1 was sequenced and found to match sequences within pc-*sis* RH-9, extending as far as 556 bp from its upstream extent (Fig. 1). More recent studies have shown that the region from position 9-761 of pc-*sis* RH-9 is also present in a SIS cDNA clone derived from normal human endothelial cell RNA (9). Using recombinant M13 clones containing the 5′ terminus of pc-*sis* RH-9 as probes for S1 analysis, we confirmed the presence of these sequences in normal placental SIS/PDGF2 mRNA (data not shown). All of these findings established that pSD1, as well as at least the 556-bp stretch of genomic DNA immediately upstream, were contained within the *SIS/PDGF2* transcript.

In an effort to define the full extent of the *SIS/PDGF2* transcript, we searched for human genomic DNA clones upstream from pc-*sis* RH-9. Using pc-*sis* RH-9 as a probe, a genomic clone, designated λ-c-*sis* clone 10, was isolated from a normal human fetal liver DNA library. As shown in Fig. 1, this clone contained sequences overlapping with pc-*sis* RH-9 as well as information upstream. A 735-bp DNA fragment that contained upstream information was isolated from λ-c-*sis* clone 10 for use as a molecular probe in nuclease S1 protection experiments (Fig. 1). Bands of 340, 330, 169, 161, and 156 nucleotides were detected following hybridization with EJ-cell mRNA and digestion with nuclease S1 (Fig. 1*A*). As a control, the same probe was completely digested by nuclease S1 after hybridization with RNA from the human tumor cell line A2780, known not to express the *SIS/PDGF2* transcript. These results localized the 5′ boundary of the upstream *SIS/PDGF2* exon initially identified by pc-*sis* RH-9 (3) ≈340 bp upstream of the *Apa* I site shown in Fig. 1.

To precisely localize the 5′ extent of this exon, we utilized a shorter, end-labeled probe (Fig. 1*B*) for simultaneous nuclease S1 and nucleotide sequencing studies. Major bands of 95 and 85 nucleotides survived nuclease S1 treatment after hybridization with mRNA from human placenta or from the human tumor cell line EJ, both of which express the *SIS/PDGF2* transcript (Fig. 1*B*). These results confirmed that the 5′ boundary of the upstream *SIS* exon is 95 bp to the left of the *Xma* III site shown in Fig. 1.

To determine whether the boundary identified by nuclease S1 experiments was the first *SIS* exon, we utilized the technique of primer extension. The 26-nucleotide primer synthesized for these studies represented the antisense sequence as indicated in Fig. 1. Using poly(A)$^+$ RNA of EJ cells or placenta as templates, we observed several extension products. In contrast, no products were generated using the same primer with poly(A)$^+$ RNA templates from A2780 cells, which lack the *SIS/PDGF2* transcript (Fig. 1*C*). The size of the longest extension product established that the exon border identified by nuclease S1 mapping studies represented the major upstream start site for human SIS/PDGF2 mRNA. There is evidence that many genes contain more than a single mRNA start site (19–21). Although premature termination of reverse transcription could account for some of the smaller

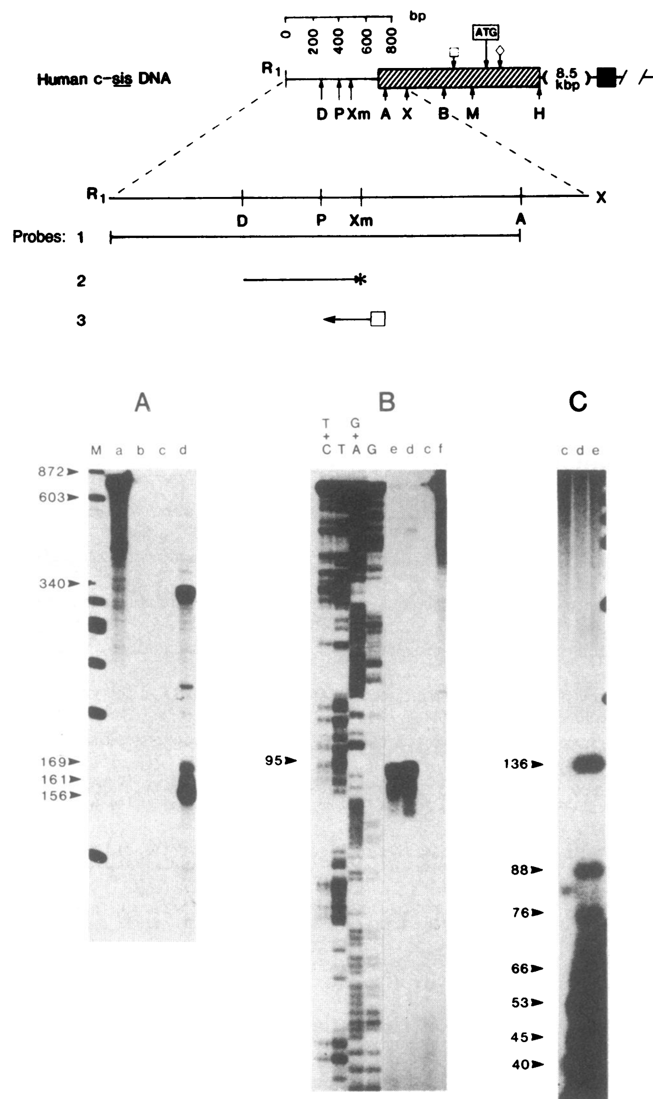

FIG. 1. Nuclease S1 and primer extension analysis of the *SIS/PDGF2* transcript. (*Upper*) Physical map of the 3′ region of λ-c-*sis*-clone 10 and the location of DNA segments used as probes. Restriction sites shown are *Eco*RI (R_1), *Dde* I (D), *Pvu* II (P), *Xma* III (Xm), *Apa* I (A), *Xho* I (X), *Bam*HI (B), *Mlu* I (M), and *Hin*dIII (H). Hatched box shows the location of pc-*sis* RH-9 (13). The symbols □ and ◇ represent the 5′ end of pSDI and 3′ end of *SIS/PDGF2* exon 1, respectively. Solid box represents the first v-*sis*-related exon. (*Lower*) Experimental results of nuclease S1 and primer-extension analyses. (*A*) The *Eco*RI–*Apa* I DNA fragment (probe 1) was inserted into an M13 vector and used to synthesize a uniformly labeled DNA. This probe was examined without further treatment by gel electrophoresis under denaturing conditions (lane a) or was hybridized with tRNA (lane b) or poly(A)$^+$ RNA from A2780 (lane c) or EJ (lane d) cells. Following digestion of hybrid molecules with nuclease S1, products were analyzed in denaturing 5% acrylamide gels. (*B*) The *Dde* I–*Xma* III DNA fragment (probe 2) was end labeled and denatured. Antisense DNA was isolated and sequenced (four lanes at left) or examined by electrophoresis without further treatment (lane f). The same probe was hybridized with poly(A)$^+$ RNA from A2780 (lane c) or EJ (lane d) cell lines or human placenta (lane e), treated with nuclease S1, and examined by gel electrophoresis. (*C*) A 26-base oligonucleotide primer (probe 3) was hybridized with poly(A)$^+$ RNA from A2780 (lane c) or EJ (lane d) cells or from human placenta (lane e). Hybrids were used as substrates for reverse transcriptase in the presence of labeled dNTPs. Extension products were analyzed by electrophoresis in denaturing polyacrylamide gels. M indicates marker lane. Molecular sizes are indicated in nucleotides.

primer-extension products observed (Fig. 1*C*), some of the sites detected corresponded to those identified by nuclease S1 mapping (Fig. 1*A*). Thus, such sites might reflect alternative mRNA start sites.

Nucleotide Sequence of the Full-Length *SIS/PDGF2* Transcript. We determined the nucleotide sequence of the entire *SIS/PDGF2* transcript, which we defined to encompass the pSD1 cDNA clone and contiguous genomic DNA sequences upstream of the 5′ mRNA start site as localized above (see Fig. 1). This composite 3373-nucleotide sequence is shown in Fig. 2.

We localized the sequence of the PDGF2 coding region from position 1023 to 1745 within the full-length transcript. The ATG at position 1023 was in-frame with the PDGF2 coding sequence and was preceded by two upstream termination codons at positions 905 and 921. The next in-frame methionine codon was localized within the known coding region of the molecule. Thus, the ATG at position 1023 is the most likely initiation codon for translation of the PDGF2 precursor. The arginine residue at amino acid position 3 of the predicted sequence of the PDGF2 precursor is followed by a stretch of 12 hydrophobic amino acids. This sequence meets

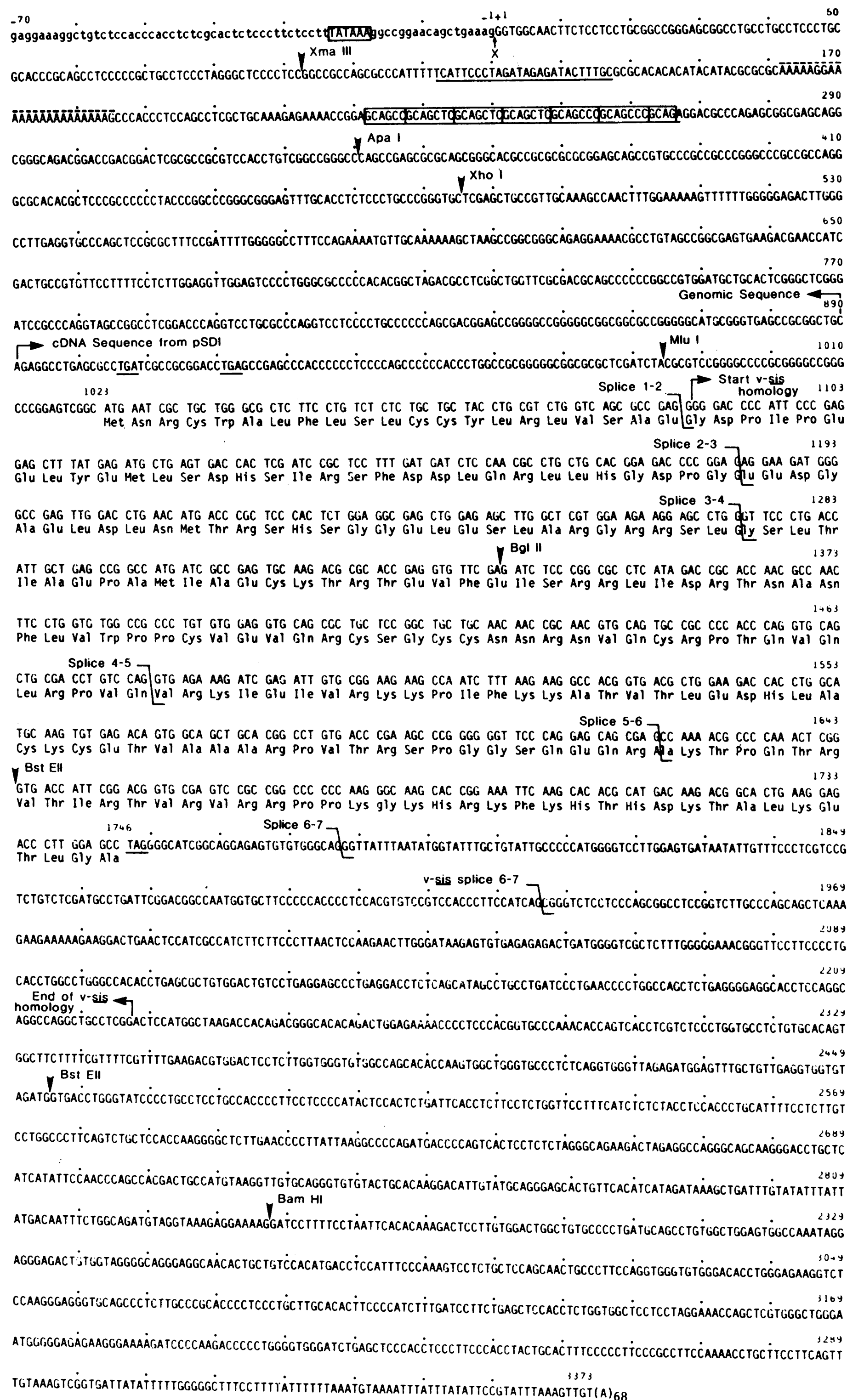

FIG. 2. Nucleotide sequence of the *SIS/PDGF2* transcript and flanking genomic DNA. Genomic sequences upstream of the mRNA start site are shown in lowercase letters. The consensus transcriptional promoter and repeated heptanucleotide sequence are boxed. The polypurine stretch (positions 162–185) is identified by a broken line above the sequence. The sequence complementary to the primer used in Fig. 1*C* is underlined. Junctions of human genomic DNA, pSD1, and the v-*sis*-related sequence are indicated by arrows. Locations at which splicing has occurred are indicated by ‾|__. The major transcriptional start site (position 1), as determined by both nuclease S1 mapping and primer extension, is indicated by an arrow labeled X. In-frame translation termination codons upstream of the PDGF2 open reading frame are underlined. Predicted amino acid sequence of the PDGF2 open reading frame is indicated below the nucleotide sequence. Note that the splice points of exons 6 and 7 for *SIS* (c-*sis*) and v-*sis* are different. A specific 149-bp human *SIS* sequence is absent in v-*sis* at the junction.

the criteria established for signal peptides which direct nascent proteins across the endoplasmic reticulum (22). Moreover, a site for N-linked glycosylation was localized at amino acid position 63. The remaining v-*sis*-related amino acid sequence of the SIS/PDGF2 precursor was identical to that previously reported for *SIS* exons (8, 13, 18), as well as for two partial cDNA clones isolated from the human T-lymphotropic virus Type I-infected cell line HUT-102 (8) and human endothelial cells (9).

The coding sequence was flanked by long 5′ and 3′ noncoding stretches of 1022 and 1625 bases, respectively. The 5′ untranslated region contained a highly G+C-rich stretch immediately upstream of the PDGF2 coding sequence. The entire 5′ noncoding sequence was around 70% in G+C content. The sequences from position 840 to 1022 and from position 331 to 450 were 85% and 90% G+C, respectively. Both the length and the high G+C content of the 5′ noncoding sequence are uncommon among mRNAs that have been analyzed to date. The length of the 3373-nucleotide transcript, even accounting for a polyadenylylated stretch, is shorter than the observed 4.0–4.2 kb size of the transcript under certain denaturing electrophoretic conditions (4, 23). It is likely that the anomalous electrophoretic mobility reflects strong secondary structures in the 5′ untranslated region of the transcript that are due to its high G+C content.

Another unusual feature of the 5′ noncoding sequence was a 7-bp sequence (GCAGCYC; Y = T or C) repeated five times and flanked by an incomplete repeat at either end. This sequence was located between positions 222 and 266 (Fig. 2). Finally, a polypurine tract of 23 nucleotides was observed between positions 162 and 185. Such stretches are known to be hypersensitive to nuclease S1 (24, 25). Three putative mRNA start sites identified as bands of 169, 161, and 156 nucleotides (Fig. 1*A*) by nuclease S1 mapping were localized to this homopurine tract. Thus, these may not be true mRNA start sites but may instead reflect the nuclease S1 hypersensitivity of homopurine stretches.

Within the 1625-nucleotide 3′ untranslated region of the *SIS/PDGF2* transcript, the highly conserved eukaryotic polyadenylylation signal AATAAA was not detected. These results were confirmed by analysis of nucleotide sequences of the 3′ regions of each of the other four cDNA clones obtained. Even though the size of the poly(A) stretch was unique for each of these cDNA clones, their polyadenylylation sites were identical. Eukaryotic mRNAs generally contain the highly conserved polyadenylylation signal anywhere within 50 bp upstream of the site of polyadenylylation (26). We identified an A+T-rich stretch in the proper location for a polyadenylylation signal. Whether this serves as an alternative polyadenylylation signal remains to be determined.

Sequence Analysis of 5′ and 3′ Flanking Genomic Regions. We next analyzed the nucleotide sequence of genomic DNA immediately upstream from the putative mRNA start site. A consensus transcriptional promoter sequence was observed 19 to 24 bp upstream from the putative mRNA start site (Fig. 2). These results confirm that this is indeed the major mRNA start site. Upstream from the promoter sequence, we were not able to detect a "CAAT box," which is commonly observed 80 bp upstream of the mRNA start sites in eukaryotic genes (27). There were several (C)CTCC repeats clustered both upstream and downstream of the transcriptional start site (Fig. 2) as well as in sequences further upstream (data not shown). Similar sequence repeats have been observed and implicated in the transcriptional regulation of the epidermal growth factor receptor gene (21).

Sequence analysis of the genomic 3′ flanking region revealed an authentic polyadenylylation signal AATAAA just 18 bp downstream of the site of polyadenylylation observed in each of our cDNA clones (Fig. 3). However, there was a stretch of 9 consecutive adenylate residues in genomic DNA

```
                                2483| Site                          2510
TAAAATTTATTTATATTCCGTATTTAAAGTTGTAAAAAAAAATAACCACAAAATAAAACC
ATTTTAAATAAATATAAGGCATAAATTTCAACATTTTTTTTTATTGGTGTTTTATTTTGG
                                                     Signal ↑  2570
AAATGAATCCGCCGGAGGTCTGTCTGTTGGCATCGTGCGTGACAATTAACCTTTCTGCCT
TTTACTTAGGCGGCCTCCAGACAGACAACCGTAGCACGCACTGTTAATTGGAAAGACGGA
                                                            2630
TGGCAGGATGTGCCGACAGCTTGCGGCGTGTTCCTCTCACTCTGGGAGCCTCAGGCGTGA
ACCGTCCTACACGGCTGTCGAACGCCGCACAAGGAGAGTGAGACCCTCGGAGTCCGCACT
                                                            2690
TCTCACACACTGGCGTGCACATACACACACACACACATACATGCTCACACATGCGTGCAC
AGAGTGTGTGACCGCACGTGTATGTGTGTGTGTGTGTATGTACGAGTGTGTACGCACGTG
                                                            2750
ATACACGCAGGCCTGCAACTTGGGGGAGGCCTCTGTCTGGCGGGAAGAAGAGACACACAG
TATGTGCGTCCGGACGTTGAACCCCCTCCGGAGACAGACCGCCCTTCTTCTCTGTGTGTC
                                                            2810
GCTACTCTGTTGGTCTTGGTCCTGGCACAGCTCCTGACACGTGAACTTGTGCGTGTCTCT
CGATGAGACAACCAGAACCAGGACCGTGTCGAGGACTGTGCACTTGAACACGCACAGAGA
                            2838
GGCAGTGACGAGAGATAAATTTCTGCAG
CCGTCACTGCTCTCTATTTAAAGACGTC
```

FIG. 3. Nucleotide sequence of 3′ flanking genomic DNA. The site of polyadenylylation as well as the downstream authentic polyadenylylation signal are shown. The putative enhancer-like TG-element is boxed. The nucleotide sequence of both strands is shown.

immediately downstream of the site where poly(A) tails were observed in each of the cDNA clones analyzed. Thus it is possible that cDNA synthesis could start here even though the true poly(A) stretch is further downstream. We also detected an imperfect "TG element" 150 bp downstream from the polyadenylylation site. The alternating copolymer tracts poly(dT·dG)poly(dC·dA), or TG elements, are highly conserved, randomly distributed components of eukaryotic genomes (28, 29). Such elements have been identified in both intragenic and intergenic regions of the eukaryotic genome and have been shown to possess enhancer-like activity owing to their ability to assume a left-handed conformation (28, 29).

DISCUSSION

In the present studies, we have defined the complete extent of the human *SIS/PDGF2* transcript by a combination of cDNA cloning, nuclease S1 mapping, and primer-extension analysis. The organization of the human *SIS/PDGF2* transcript as deduced from our findings is summarized in Fig. 4. The coding regions are encompassed within the first six exons. The great majority of the first exon, as well as the entire seventh exon, are comprised of noncoding sequences. The long lengths of the 5′ and 3′ untranslated regions of the transcript are unusual among eukaryotic genes so far studied. However, there is emerging evidence that at least some of the genes involved in the pathways by which growth factors stimulate normal cellular proliferation may exhibit long 5′ and/or 3′ noncoding sequences (30–32). The significance of these long noncoding sequences remains to be determined.

Within the simian sarcoma virus genome, the v-*sis* sequence starts at a position corresponding to the 5′ border of

FIG. 4. Structural organization of the *SIS/PDGF2* gene. Coding portions of exons are repesented by solid boxes, and noncoding ones, by open boxes. Exons are numbered 1–7. Introns are not drawn to scale. Numbers in parentheses refer to intron lengths in kb. The 3′ boundaries of exons are indicated by the respective nucleotide numbers.

human *SIS/PDGF2* exon 2. Like its viral counterpart, the normal human *SIS/PDGF2* coding sequence has the capacity to efficiently transform cells responsive to its growth-promoting activity (3). Our present findings establish that the major contribution of the helper virus to the structure and function of the v-*sis*-encoded transforming protein was its substitution of an initiation signal for translation and a signal peptide sequence residing within the N-terminal region of its *env* gene for directly analogous sequences present within the first exon of the normal *SIS/PDGF2* gene.

PDGF is found in the α granules of platelets and is presumably synthesized in their megakaryocyte precursors. More recent studies have revealed expression of the *SIS/PDGF2* gene in a variety of other tissues, including normal endothelial cells (33), placental trophoblasts (34), and activated monocytes (35). Accumulating evidence concerning the regulation of other genes with selective tissue expression, such as the gene for α-fetoprotein, has suggested the existence of multiple regulatory elements that may affect gene expression in individual tissues (R. Godbout and S. Tilghman, personal communication). Efforts to dissect the normal control of *SIS/PDGF2* gene expression during development and differentiation should be aided by our identification and isolation of regions flanking the transcribed sequences of the *SIS/PDGF2* gene.

Expression of the normal human *SIS/PDGF2* coding sequence in assay cells responsive to this growth factor can cause neoplastic transformation (3). Moreover, the SIS/PDGF2 mRNA has been detected in tumors such as glioblastomas and fibrosarcomas (4), whose normal counterparts possess PDGF receptors and are responsive to the growth-promoting action of the *SIS/PDGF2* gene product (5, 6). Our present findings show that the major mRNA start sites of *SIS/PDGF2* transcripts derived from normal placenta and from a human tumor cell line were identical, excluding the possibility that *SIS/PDGF2* expression in this tumor reflects the altered transcriptional initiation of this gene. Further investigation of the role of flanking sequences that may affect *SIS/PDGF2* gene expression in tumor cells that expressthe *SIS/PDGF2* transcript should help to elucidate the normal transcriptional control of this gene, as well as mechanisms that may activate it as an oncogene in certain tumor cells.

1. Robbins, K. C., Antoniades, H. N., Devare, S. G., Hunkapiller, M. W. & Aaronson, S. A. (1984) in *Cancer Cells*, eds. Levine, A. J., Vande Woude, G. F., Topp, W. C. & Watson, J. D. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Vol. 1, pp. 35–41.
2. Leal, F., Williams, L. T., Robbins, K. C. & Aaronson, S. A. (1985) *Science* **230**, 327–330.
3. Gazit, A., Igarashi, H., Chiu, I.-M., Srinivasan, A., Yaniv, A., Tronick, S. R., Robbins, K. C. & Aaronson, S. A. (1984) *Cell* **39**, 89–97.
4. 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. & Aaronson, S. A. (1982) *Nature (London)* **295**, 116–119.
5. Heldin, C. H., Westermark, B. & Wasteson, A. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3664–3668.
6. Bowen-Pope, D. F. & Ross, R. (1982) *J. Biol. Chem.* **257**, 5161–5171.
7. Huang, J. S., Huang, S. S., Kennedy, B. & Duel, T. F. (1982) *J. Biol. Chem.* **257**, 8130–8136.
8. Josephs, S. F., Ratner, L., Clarke, M. F., Westin, E. H., Reitz, M. S. & Wong-Staal, F. (1984) *Science* **225**, 636–639.
9. Collins, T., Ginsburg, D., Boss, J. M., Atkin, S. H. & Pober, J. S. (1985) *Nature (London)* **316**, 748–750.
10. Perucho, M., Goldfarb, M., Shimizu, K., Lama, C., Fogh, J. & Wigler, M. (1981) *Cell* **27**, 467–476.
11. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
12. Okayama, H. & Berg, P. (1982) *Mol. Cell. Biol.* **2**, 161–170.
13. Chiu, I.-M., Reddy, E. P., Givol, D., Robbins, K. C., Tronick, S. R. & Aaronson, S. A. (1984) *Cell* **37**, 123–129.
14. Hanahan, D. & Meselson, M. (1980) *Gene* **10**, 63–67.
15. Favaloro, J., Treisman, R. & Kamen, R. (1980) *Methods Enzymol.* **65**, 718–749.
16. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 498–560.
17. Josephs, S. F., Guo, C., Ratner, L. & Wong-Staal, F. (1984) *Science* **223**, 487–491.
18. Johnsson, A., Heldin, C.-H., Wasteson, A., Westermark, B., Devel, T. F., Huang, J. S., Seeburg, H., Gray, A., Ullrich, A., Scrace, G., Stroobant, P. & Waterfield, M. D. (1984) *EMBO J.* **3**, 921–928.
19. Reynolds, G. A., Basu, S. K., Osborne, T. F., Chin, D. J., Gil, G., Brown, M. S., Goldstein, J. L. & Luskey, L. L. (1984) *Cell* **38**, 275–285.
20. Sudhof, T. C., Goldstein, J. L., Brown, M. S. & Russel, D. W. (1985) *Science* **228**, 815–822.
21. Ishii, S., Xu, Y.-H., Stratton, R. H., Roe, B. A., Merlino, G. T. & Pastan, I. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4920–4924.
22. Kreil, G. (1981) *Annu. Rev. Biochem.* **50**, 317–348.
23. Westin, E. H., Wong-Staal, F., Gelmann, E. P., Dalla Favera, R., Papas, T. S., Lautenberger, J. A., Eva, A., Reddy, E. P., Tronick, S. R., Aaronson, S. A. & Gallo, R. C. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2490–2494.
24. Evans, T., Schon, E., Gora-Maslak, G., Patterson, J. & Efstratiadis, A. (1984) *Nucleic Acids Res.* **12**, 8043–8058.
25. Cantor, C. R. & Efstratiadis, A. (1984) *Nucleic Acids Res.* **12**, 8059–8072.
26. Manley, J. L., Yu, H. & Ryner, L. (1985) *Mol. Cell. Biol.* **5**, 373–379.
27. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383.
28. Hamada, H., Seidman, M., Howard, B. H. & Gorman, C. M. (1984) *Mol. Cell. Biol.* **4**, 2622–2630.
29. Hamada, H., Petrino, M. G., Kakunaga, T., Seidman, M. & Stollar, B. D. (1984) *Mol. Cell. Biol.* **4**, 2610–2621.
30. Battey, J., Moulding, C., Taub, R., Murphy, W., Stewart, T., Potter, H., Lenoir, G. & Leder, P. (1983) *Cell* **34**, 779–787.
31. Gray, A., Dull, T. J. & Ullrich, A. (1983) *Nature (London)* **303**, 722–725.
32. Dull, T. J., Gray, A., Hayflick, J. S. & Ullrich, A. (1984) *Nature (London)* **310**, 777–781.
33. Barrett, T. B., Gajdusek, C. M., Schwartz, S. M., McDougall, J. K. & Benditt, E. P. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6772–6774.
34. Gorestin, A. S., Betsholtz, C., Pfeifer-Ohlsson, S., Persson, H., Rydnert, J., Bywater, M., Holmgren, G., Heldrin, C.-H., Westermark, B. & Ohlsson, R. (1985) *Cell* **41**, 301–312.
35. Martinet, Y., Bitterman, P. B., Mornex, J.-F., Grotendorst, G. R., Martin, G. R. & Crystal, R. G. (1986) *Nature (London)* **319**, 158–160.