MOLECULAR AND CELLULAR BIOLOGY, May 1989, p. 2251–2253
0270-7306/89/052251-03$02.00/0

Vol. 9, No. 5

The Alternatively Spliced Exon of the Platelet-Derived Growth Factor A Chain Encodes a Nuclear Targeting Signal

DAVID W. MAHER,[1,2] BRUCE A. LEE,[2,3] AND DANIEL J. DONOGHUE[1,2]*

Departments of Chemistry[1] and Biology[3] and Center for Molecular Genetics,[2] University of California—San Diego, La Jolla, California 92093

Received 9 November 1988/Accepted 24 January 1989

We have previously shown that the SIS/platelet-derived growth factor B chain contains a nuclear targeting signal near its C terminus. Here we show that the platelet-derived growth factor A chain also contains a nuclear targeting signal encoded by an exon which is subject to alternative splicing. This sequence is capable of targeting a nonsecreted form of the A chain to the nucleus and can also target the cytoplasmic proteins dihydrofolate reductase, chloramphenicol acetyltransferase, and pyruvate kinase to the nucleus.

Platelet-derived growth factor (PDGF) is a potent mitogen for cells of mesenchymal origin. Human PDGF is composed of two related polypeptide chains (A and B) linked by disulfide bonds (19). Each chain is also able to form homodimers, as exemplified by the B-chain homodimers produced in simian sarcoma virus-transformed cells (9) and by the A-chain homodimers produced in various tumor cell lines (2). The A-chain protein appears to have either of two different C termini which result from alternative splicing of the RNA (7, 21). The predicted PDGF A-chain protein encoded by a human glial tumor cDNA (2) (referred to here as the long form) contains a sequence near its C terminus (RPRESGKKRKRKR) that is similar to the nuclear targeting signal (NTS) we had previously identified in the SIS/B-chain protein (RRPPKGKHRK). Subsequently, A-chain cDNAs isolated from endothelial cells (4, 7, 21) and from the glioma cell line from which the original A-chain clone was isolated (18) were shown to encode a shorter protein (referred to as the short form), with the C terminus coded for by a different exon. No differences in biological activity between the long and short forms of the PDGF A chain have yet been found (1).

To determine whether the A chain contains a NTS, eucaryotic expression vectors containing the long A chain as well as nonsecreted forms of the long and short A chains were constructed. Plasmid D-1, containing the longer PDGF A-chain cDNA derived from a human glioma, was obtained (2). The A-chain-coding region from D-1 was then inserted into a vector derived from pRSV (8) downstream of a Rous sarcoma virus promoter, creating pDM109. The short form of the A chain lacking the NTS was created through loop-out mutagenesis in m13 by using an oligonucleotide corresponding to the exon 5-exon 7 junction which occurs in the short form of the A chain (GAGGACACGGATGTGAGGTG AGG). The native A-chain protein contains a signal sequence at its N terminus, allowing its secretion. Since signal sequences function cotranslationally, it was necessary to remove this in order to examine the C terminus for its ability to direct import to the nucleus. Nonsecreted forms of the different A chains were created by joining a consensus start codon for eucaryotic initiation (12, 14) to the A-chain-coding sequences at a *Taq*I site so that the resulting proteins would resemble the mature A chains after propeptide cleavage. These nonsecreted forms were also placed in a Rous sarcoma virus vector, yielding pDM134 (nonsecreted long form) and pDM135 (nonsecreted short form). The plasmids were transfected into the African green monkey kidney cell line CV-1, and 36 to 48 h posttransfection, the cells were assayed by indirect immunofluorescence (15) to determine the location of the protein. The native long form of the A chain showed the characteristic endoplasmic reticulum and Golgi staining patterns of a secreted protein (Fig. 1B). The nonsecreted short form gave whole-cell staining (Fig. 1C), with the protein distributed throughout the cytoplasm and nucleus (the sizes of the nonsecreted forms of the A chain are such that the proteins can diffuse through the nuclear pores [3]). The nonsecreted long form showed intense nuclear and nucleolar staining (Fig. 1D), similar to that observed with nonsecreted forms of the SIS protein. This demonstrates that the C terminus of the long form of the A chain contains a NTS which is not present in the alternatively spliced short form.

To show that the NTS in the PDGF A chain was sufficient to target non-nuclear proteins to the nucleus, gene fusions were created in which portions of exon 6 of the A chain (encoding the NTS) were fused to the 3′ end of the coding regions of the pyruvate kinase (PK), chloramphenicol acetyltransferase (CAT), and dihydrofolate reductase (DHFR) genes. Plasmids encoding specially constructed CAT and DHFR genes which contained blunt restriction sites immediately at the 3′ end of the coding region were obtained (12). The 3′ end of the long form of the A chain containing the NTS was added directly to the ends of CAT and DHFR genes by using the *Stu*I site (residue 970) in the PDGF A-chain gene, and the fusions were placed under the Rous sarcoma virus promoter, creating pDM117 (CAT-NTS) and pDM118 (DHFR-NTS). The PK-NTS fusion (pAL125) was created by joining the *Bst*XI site (residue 1567) of the PK-coding sequence in pAL79 (15) to the *Sty*I site (residue 973) in exon 6 of the A chain by using *Eco*RI linkers to maintain the correct reading frame. The hybrid genes were expressed in CV-1 cells, and the transfected cells were examined by indirect immunofluorescence for the location of the hybrid proteins. CAT (Fig. 2B), DHFR (Fig. 2D), and PK (Fig. 2F) were efficiently targeted to the nucleus by the A-chain NTS. Additionally, CAT was concentrated in the nucleoli, as were the nonsecreted forms of the A and B chains. It appears that this short sequence is capable of not only nuclear targeting but also nucleolar targeting, although we cannot rule out an interaction between the NTSs and

* Corresponding author.

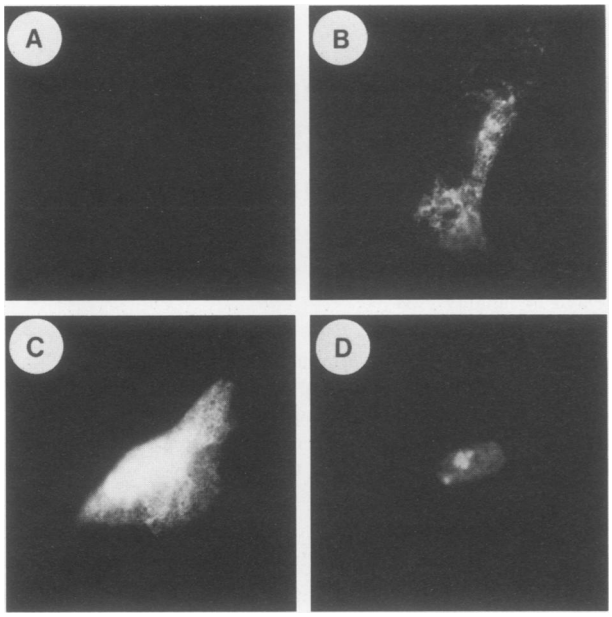

FIG. 1. Indirect immunofluorescence of secreted and nonsecreted forms of the PDGF A chain expressed under the Rous sarcoma virus promoter in CV-1 cells. (A) Mock-transfected cells; (B) the native long form of the A chain showing diffuse endoplasmic reticulum and Golgi staining; (C) a nonsecreted short form of the A chain showing whole-cell staining; (D) a nonsecreted long form of the A chain showing pronounced nucleolar and nuclear staining. Plasmids transfected into CV-1 cells by the calcium phosphate procedure were pDM109 (B), pDM135 (C), and pDM134 (D). An affinity-purified rabbit antiserum raised against a bacterially synthesized A-chain fusion protein served as the primary antibody, and rhodamine-conjugated goat anti-rabbit antiserum was used as the secondary antibody.

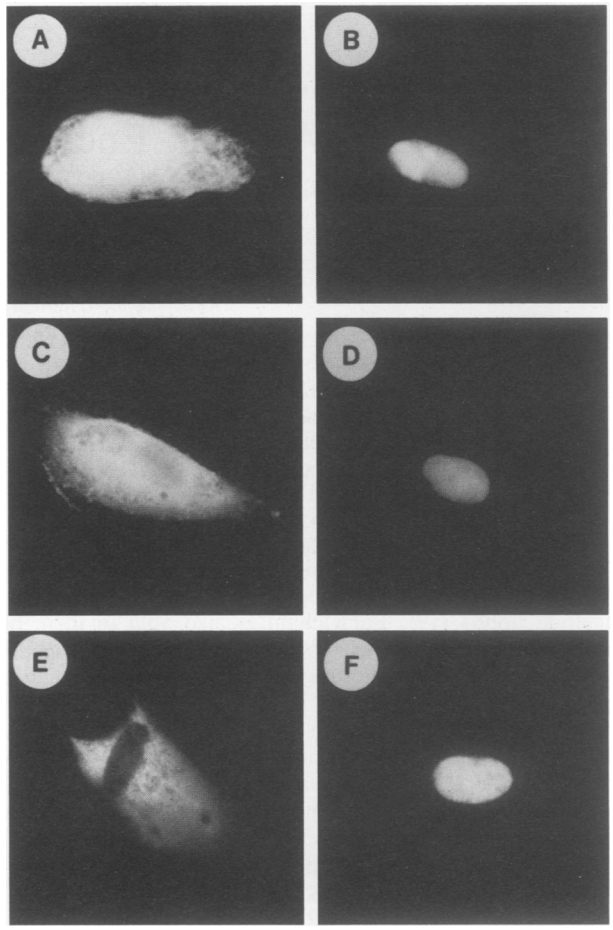

FIG. 2. Targeting of CAT, DHFR, and PK to the nucleus by the A-chain NTS. (A) Wild-type CAT; (B) CAT-NTS fusion; (C) wild-type DHFR; (D) DHFR-NTS fusion; (E) wild-type PK; (F) PK-NTS fusion. Wild-type CAT and DHFR are small enough to diffuse through the nuclear pores; PK is not (3). Plasmids transfected into CV-1 cells by the calcium phosphate procedure were pSV2CAT (A), pDM117 (B), pSV2DHFR (C), pDM118 (D), pAL79 (E), and pAL125 (F). pSV2CAT (11), pSV2DHFR (20), and pAL79 (containing PK [15]) contain the wild-type genes under the control of the simian virus 40 early promoter. The CAT-NTS and PK-NTS fusions have the sequence PRESGKKRKRKRLKPT appended to the C terminus of each protein, and the predicted molecular mass is increased from 25.6 to 27.6 kilodaltons for CAT and from 21.6 to 23.6 kilodaltons for DHFR. The PK-NTS fusion adds 7 amino acids from linkers and 15 from exon 6 (NTIGIPARESGKKR KRKRLKPT) to the C terminus of PK and increases the predicted molecular mass from 57.8 to 60.3 kilodaltons. Indirect immunofluorescence was performed by using a monoclonal antibody against CAT or polyclonal antisera raised in rabbits against DHFR or PK as the primary antibody, respectively, and the corresponding rhodamine-conjugated secondary antibody.

sequences elsewhere in CAT or PDGF that causes the nucleolar localization.

Different exons code for the C termini of the long and short forms of the A chain. Exon 6 contains 69 base pairs and encodes the 18 amino acids at the C terminus of the long form of the A chain (containing the NTS) but is spliced out of the A-chain mRNA in the short form (4, 7, 18, 21). Exon 6 was previously believed to be nonconserved between the A and B chains (4, 18). Figure 3 reflects the similarity between the A and B chain NTSs encoded by exon 6 in both genes. These sequences contain a proline(s) flanked by positively charged amino acids at the N terminus and a glycine followed by a run of positively charged residues at the C terminus. Additionally, we noted a similarity between the four residues at the C terminus of the long A chain, LKPT, and a sequence in the B chain, LKET. This LKXT sequence is not required for the nuclear targeting of the B chain (15) and is probably not required for the targeting of the A chain.

That secreted growth factors may have a direct role in the nucleus has been a controversial topic (6, 13). Early studies indicated that many growth factors were degraded after binding to and entering their target cells (13). But basic fibroblast growth factor and insulin seem clearly to exert some of their effects directly in the nucleus. Basic fibroblast growth factor, a competence factor like PDGF, has recently been observed to accumulate in the nucleoli of cells and to directly stimulate ribosome biosynthesis (5). Insulin injected into *Xenopus laevis* oocytes was found to directly stimulate RNA, protein, and glycogen synthesis (16). PDGF itself has been found by one group to accumulate in the nuclei of cells bearing its receptor (17), and SIS-related proteins have been found in the nuclei of cells transformed by simian sarcoma virus (22). It is possible that all or part of PDGF containing the NTS is transported to the nucleus after binding and internalization and may, like basic fibroblast growth factor, play a direct role in stimulating the $G_0 \rightarrow G_1$ transition.

We have identified NTSs in both chains of PDGF. The C termini of the A and B chains previously were not thought to be similar. We have shown that the two C termini have a

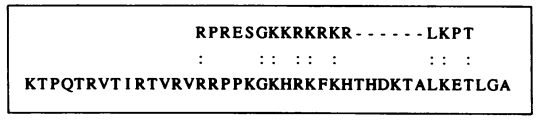

FIG. 3. Alignment of amino acids exclusively encoded by exon 6 of the human PDGF A-chain (top) and B-chain (bottom) genes demonstrating the conservation of NTSs as well as the LKXT sequence. The amino acids from each exon were aligned by using the program pDayhoff as described elsewhere (10).

conserved amino acid sequence and that each is able to act as a NTS. It is possible that these are merely cryptic NTSs; however, it seems unlikely that the ability of these sequences to cause nuclear localization was fortuitously conserved along with some other presumably conserved function for the C terminus. A role for PDGF in the nucleus remains to be determined. The alternative splicing of the NTS in the PDGF A chain may represent a level of control over this putative nuclear role.

We thank Christer Betsholtz for the A-chain clone D-1 and Steve Gould and Suresh Subramani for CAT and DHFR plasmids and antisera. We also thank Bruce Roberts for PK plasmids and antisera, Jeff Gray for his assistance with the alignment, and S. Jon Singer and Immo Scheffler for use of their microscope facilities.

This work was supported by Public Health Service grant CA 40573 from the National Institutes of Health. B.A.L. gratefully acknowledges support from a National Science Foundation graduate fellowship.

LITERATURE CITED

1. **Beckmann, M. P., C. Betsholtz, C.-H. Heldin, B. Westermark, E. Di Marco, P. P. Di Fiore, K. C. Robbins, and S. A. Aaronson.** 1988. Comparison of biological properties and transforming potential of human PDGF-A and PDGF-B chains. Science **241**:1346–1349.
2. **Betsholtz, C., A. Johnsson, C.-H. Heldin, B. Westermark, P. Lind, M. S. Urdea, R. Eddy, T. B. Shows, K. Philpott, A. L. Mellor, T. J. Knott, and J. Scott.** 1986. cDNA sequence and chromosomal localization of human platelet-derived growth factor A-chain and its expression in tumour cell lines. Nature (London) **320**:695–699.
3. **Bonner, W. M.** 1975. Protein migration into nuclei. J. Cell Biol. **64**:421–430.
4. **Bonthron, D. T., C. C. Morton, S. H. Orkin, and T. Collins.** 1988. Platelet-derived growth factor A chain: gene structure, chromosomal location, and basis for alternative mRNA splicing. Proc. Natl. Acad. Sci. USA **85**:1492–1496.
5. **Bouche, G., N. Gas, H. Prats, V. Baldin, J.-P. Tauber, J. Teissie, and F. Amalric.** 1987. Basic fibroblast growth factor enters the nucleolus and stimulates the transcription of ribosomal genes in ABAE cells undergoing $G_0 \rightarrow G_1$ transition. Proc. Natl. Acad. Sci. USA **84**:6770–6774.
6. **Burwen, S. J., and A. L. Jones.** 1987. The association of polypeptide hormones and growth factors with the nuclei of target cells. Trends Biochem. Sci. **12**:159–162.
7. **Collins, T., D. T. Bonthron, and S. H. Orkin.** 1987. Alternative RNA splicing affects function of encoded platelet-derived growth factor A chain. Nature (London) **328**:621–624.
8. **De Wet, J. R., K. V. Wood, M. DeLuca, D. R. Helinski, and S. Subramani.** 1987. Firefly luciferase gene: structure and expression in mammalian cells. Mol. Cell. Biol. **7**:725–737.
9. **Doolittle, R. F., M. W. Hunkapiller, L. E. Hood, S. G. Devare, K. C. Robbins, S. A. Aaronson, and H. N. Antoniades.** 1983. Simian sarcoma virus *onc* gene, v-*sis*, is derived from the gene (or genes) encoding a platelet-derived growth factor. Science **221**:275–277.
10. **Feng, D.-F., and R. F. Doolittle.** 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. J. Mol. Evol. **25**:351–360.
11. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
12. **Gould, S. G., G.-A. Keller, and S. Subramani.** 1987. Identification of a peroxisomal targeting signal at the carboxy terminus of firefly luciferase. J. Cell Biol. **105**:2923–2931.
13. **James, R., and R. A. Bradshaw.** 1984. Polypeptide growth factors. Annu. Rev. Biochem. **53**:259–292.
14. **Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44**:283–292.
15. **Lee, B. A., D. W. Maher, M. Hannink, and D. J. Donoghue.** 1987. Identification of a signal for nuclear targeting in platelet-derived-growth-factor-related molecules. Mol. Cell. Biol. **7**: 3527–3537.
16. **Miller, D. S.** 1988. Stimulation of RNA and protein synthesis of intracellular insulin. Science **240**:506–509.
17. **Rakowicz-Szulczynska, E. M., U. Rodeck, M. Herlyn, and H. Koprowski.** 1986. Chromatin binding of epidermal growth factor, nerve growth factor, and platelet-derived growth factor in cells bearing the appropriate surface receptors. Proc. Natl. Acad. Sci. USA **83**:3728–3732.
18. **Rorsman, F., M. Bywater, T. J. Knott, J. Scott, and C. Betsholtz.** 1988. Structural characterization of the human platelet-derived growth factor A-chain cDNA and gene: alternative exon usage predicts two different precursor proteins. Mol. Cell. Biol. **8**: 571–577.
19. **Ross, R.** 1987. Platelet-derived growth factor. Annu. Rev. Med. **38**:71–79.
20. **Subramani, S., R. Mulligan, and P. Berg.** 1981. Expression of the mouse dihydrofolate reductase complementary deoxyribonucleic acid in simian virus 40 vectors. Mol. Cell. Biol. **1**: 854–864.
21. **Tong, B. D., D. E. Auer, J. Jaye, M. M. Kaplow, G. Ricca, E. McConathy, W. Drohan, and T. F. Deuel.** 1987. cDNA clones reveal differences between human glial and endothelial cell platelet-derived growth factor A-chains. Nature (London) **328**: 619–621.
22. **Yeh, H.-J., G. F. Pierce, and T. F. Deuel.** 1987. Ultrastructural localization of a platelet-derived growth factor/v-*sis*-related protein(s) in cytoplasm and nucleus of simian sarcoma virus-transformed cells. Proc. Natl. Acad. Sci. USA **84**:2317–2321.