Volume 3 · Number 2 · 1993
GENE EXPRESSION

Molecular organization of the rat glia-derived nexin/protease nexin-1 promoter

Henrik Ernø and Denis Monard

Friedrich Miescher Institute, Basel, Switzerland

The first three exons and the promoter of rat glia-derived nexin, also called protease nexin-1 (GDN/PN-1), have been identified through analysis of rat genomic clones. A 1.6 kilobase (kb) fragment containing 105 base pairs of the first exon and 5′-flanking sequences was sequenced. The 5′-flanking sequence and the first exon were found to be GC-rich, indicating that the 5′ region of the rat GDN/PN-1 gene resides within a CpG island. A TATA box-like sequence, but no CAAT box, was found. The rat GDN/PN-1 promoter contains five SP1 consensus sites, four consensus sites for the MyoD1 transcription factor, and one binding site for the transcription factors NGFI-A, NGFI-C, Krox-20, and Wilms tumor factor. The presence of these consensus sequences is consistent with the known expression pattern of GDN/PN-1. Primer extension and RNase protection assays identified one transcriptional start site. The 1.6 kb promoter fragment cloned in a reporter plasmid was found to induce firefly luciferase expression in a cell-specific manner. A positive regulatory element is localized in the region −1545 to −389. In vitro CpG methylation blocked transcription from the GDN/PN-1 promoter in rat hepatoma cells but not in C6 rat glioma cells.

Glia-derived nexin (GDN), also known as protease nexin-1 (PN-1), is a potent serine protease inhibitor first purified as a protein promoting neurite outgrowth in mouse neuroblastoma cells, as well as in cultured chick sympathetic and rat hippocampal neurons (Farmer et al., 1990; Guenther et al., 1985; McGrogan et al., 1988; Monard et al., 1973; Zurn et al., 1988). The primary amino acid sequence showing domains of significant homology with other serine protease inhibitors qualifies GDN/PN-1 as a member of the serpin superfamily (Gloor et al., 1986; Sommer et al., 1987). In cell-free systems, GDN/PN-1 inhibits the activities of thrombin, urokinase, tissue plasminogen activator, and trypsin (Stone et al., 1987). Recent findings suggest that the mechanism of neurite-promoting activity is the binding and inhibition of thrombin by GDN/PN-1. Thrombin has been shown to cause neurite retraction through cleavage and activation of the thrombin receptor (Suidan et al., 1992). It was previously shown that glia-derived nexin and protease nexin-1 are identical proteins (McGrogan et al., 1988).

In the rat, GDN/PN-1 is expressed at high levels in the brain, where it is developmentally regulated: the highest mRNA expression is detected before postnatal day 12 (Gloor et al., 1986). However, GDN/PN-1 expression remains high in the olfactory system, where neuronal degeneration and regeneration continue throughout life (Reinhard et al., 1988). GDN/PN-1 is upregulated at both the mRNA and protein levels following lesion of the rat sciatic nerve (Meier et al., 1989). In the central nervous system, glial cells of the CA1 region of the gerbil

Received December 16, 1992; revision accepted March 24, 1993.

Correspondence: Denis Monard, Friedrich Miescher Institute, P. O. Box 2543, 4002 Basel, Switzerland Tel 41-61-697-66-58 Fax 41-61-697-39-76

 1052-2166/93/0302/163-12$2.00

hippocampus showed a similar upregulation of GDN/PN-1 following the death of pyramidal neurons caused by transient ischemia (Hoffmann et al., 1992). These findings indicate that the regulation of GDN/PN-1 expression is important both during development and in pathological situations.

In order to study the molecular mechanisms of the regulation of rat GDN/PN-1, two rat genomic libraries were screened with probes containing the 5′ noncoding sequence of the rat cDNA (Sommer et al., 1987). Three clones carrying the GDN/PN-1 promoter and exon I, and two additional clones carrying exons II and III were isolated. Primer extension and RNase protection assays demonstrated one transcriptional start site. Sequencing of the proximal part of the GDN/PN-1 promoter has revealed consensus sequences for different transcription factors, among them Sp1, MyoD1, Myc, NGFI-A, NGFI-C, Krox-20, and Wilms tumor factor, which is consistent with the known in vivo expression of GDN/PN-1 (Festoff et al., 1991; Hoffmann et al., 1992; Meier et al., 1989). Transient transfection experiments using firefly luciferase as a reporter gene demonstrated that the proximal GDN/PN-1 promoter (−1545 to +105) is active in all of the different cell types tested. Rat C6 glioma cells stably transfected with the GDN/PN-1 promoter controlling the luciferase gene showed very high levels of luciferase activity, while Rat-1 fibroblasts showed much lower activity. This indicates that the promoter fragments tested operate in a glia-specific manner. FTO2B rat hepatoma cells that do not contain GDN/PN-1 mRNA expressed high levels of luciferase after transient transfection with the reporter plasmid. However, after in vitro CpG methylation of the plasmid containing the GDN/PN-1 promoter and the reporter gene, no luciferase activity could be detected in FTO2B cells, indicating that GDN/PN-1 promoter activity is regulated by CpG methylation. CpG methylation does not block transcription in C6 rat glioma cells. Gel shift experiments with nuclear extracts from rat liver detect proteins that bind methylated sequences of the GDN/PN-1 promoter, but not unmethylated sequences.

Materials and methods

Isolation of genomic clones

A lambda rat genomic library (Stratagene, cat. no. 943501) was screened using a 280 bp fragment containing 100 bp of the 5′ noncoding region, 160 bp of the coding sequence of the rat GDN/PN-1 cDNA sequence (Sommer et al., 1987), and 20 bp of the pGem3 polylinker (Fig. 1A).

A second genomic library (Stratagene, cat. no. 945501) was screened with a probe prepared as follows. An 83 nt oligonucleotide corresponding to the 5′ noncoding sequence of the rat GDN/PN-1 cDNA (J. Sommer, unpublished results) was primed with a 9 nt oligonucleotide corresponding to the 3′ end of the 83 nt oligonucleotide and extended with Klenow enzyme in the presence of [^{32}P]dCTP. The resulting probe corresponded to the first exon of GDN/PN-1 upstream of the splice site, as shown in Fig. 1A. The positive phages were purified by three sequential rounds of screening and plaque purification. The clones were mapped using conventional restriction enzyme analysis and Southern blot analysis.

Subcloning and DNA sequencing

Fragments of the genomic clones were subcloned into the pGEM3 and Bluescript SK(+) plasmids using conventional techniques (Sambrook et al., 1989). Putative GDN/PN-1 promoter fragments were cloned into the luciferase vector pPALU (Artelt et al., 1991), replacing the SV40 promoter. The constructions are shown in Figure 4A.

Sequencing was performed by the dideoxynucleotide chain termination method with the Sequenase Kit (United States Biochemicals) as specified by the supplier.

Primer extension analysis

Total RNA from C6 rat glioma cells was prepared using the guanidinium isothiocyanate method as described (Sambrook et al., 1989). Oligonucleotides were end-labeled with ^{32}P, hybridized overnight with 20 μg total RNA, extended with reverse transcriptase for 1 hour at 37°C, and analyzed on a 6% sequencing gel as described (Sambrook et al., 1989).

RNase protection assay

A 1.6 kb EcoR I fragment containing 105 bp of the first exon and upstream sequences of the rat GDN/PN-1 gene was cloned into the Bluescript SK(+) plasmid and linearized with Afl II, which cuts once within the GDN/PN-1 promoter sequence. The resulting fragment was gel-purified, phenol-extracted, and used as a tem-

plate for the T7 RNA polymerase under the conditions specified by the manufacturer (New England Biolabs). The probe was digested with RNase-free DNase I, phenol-extracted, and hybridized overnight with 5 μg total RNA at 45°C. The hybrid molecules were digested with RNase A and RNase T1 for 60 minutes at 25°C, proteinase K-digested, extracted with phenol, and analyzed on a 6% sequencing gel.

Transfection

COS-7 African Green Monkey kidney cells, C6 rat glioma cells, P2T rat Schwannoma cells, Rat-1 fibroblasts, L6 rat myoblasts cells, and FTO2B rat hepatoma cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS), except for P2T cells, which grew in the presence of 5% FCS.

Transfections were carried out at a cell density of 60–80% confluency using the lipofecting agent DOTAP as described by the supplier (Boehringer Mannheim).

DNA uptake was determined by isolating the transfected plasmid at day two after transfection by Hirt extraction (Hirt, 1967) and transformation of competent HB101 bacteria.

Stable transfectants were obtained by co-transfection with a hygromycin B-resistant plasmid and selection with hygromycin B (0.2 mg/ml). Surviving clones were pooled and assayed for luciferase expression.

Methylation

Methylation of plasmids and oligonucleotides (see below) was performed by using the Sss I methylase according to the supplier (New England Biolabs). Methylation of the plasmids and oligonucleotides was analyzed by digestion with the methylation-sensitive enzyme Hpa II, followed by gel electrophoresis analysis.

Luciferase assays

To determine the promoter activity of the genomic clones, cells were transfected with the luciferase constructs (Fig. 4A) and grown for 48 hours. Cell extracts were prepared as follows. Cells were washed once with PBS; scraped off the dish in 1 ml of 0.1 M potassium phosphate, pH 7.5, and 1 mM DTT; freeze-thawed three times; and spun down in a microfuge at room temperature. The supernatant (1 ml) was collected and 100 μl used for the luciferase assay as described (de Wet et al., 1987). Luciferase activity was normalized to the amount of protein, determined by using the Bio-Rad protein assay as described by the manufacturer.

Preparation of nuclear extracts and gel shift assays

Nuclear extracts from male Wistar rat liver were prepared as described (Dignam et al., 1983) and fractionated on a heparin sepharose column. Bound protein was eluted stepwise with increasing KCl concentration from 0.1 to 1 M. For each gel shift assay, labeled oligonucleotide and 2 μg nuclear extract protein was incubated with 2 μg poly(dI-dC) and analyzed on a native 3% polyacrylamide gel. A pair of oligonucleotides (sense and antisense) from the GDN/PN-1 promoter sequence, position −283 to −245 (5′-GACCCACGTGCAGCTCTACCCCGGGTTGGC GCCG-3′), was used as a probe.

Results

Isolation and cloning of the rat GDN/PN-1 genomic clones

To define the 5′ region of rat GDN/PN-1 gene and study its regulation, five rat genomic clones covering the 5′ end of the rat GDN/PN-1 gene were isolated from two rat lambda genomic libraries. Two genomic clones (lambda gem8 and lambda fix6, shown in Fig. 1B) were isolated using the 280 bp 5′ fragment of the rat cDNA clone (Sommer et al., 1987) as a probe. The lambda gem8 phage contained a 16 kb genomic insert; the lambda fix6 phage contained 14 kb of genomic insert. Restriction enzyme analysis showed these two clones to be partially overlapping. DNA sequencing and comparison with the known cDNA sequence showed that both of these clones contained exons II and III of the rat GDN/PN-1 gene. Comparison between the cDNA and the genomic sequences showed 100% homology from a position 22 bp upstream of the ATG and downstream to codon 89, indicating the presence of an intron within the 5′ noncoding sequence of the cDNA (Fig. 1B). Further mapping and sequencing of the two clones showed the presence of exon III, extending from codon 90 to codon 161.

To isolate exon I and the promoter region of the rat GDN/PN-1 gene, a second genomic library was screened with the 83 base exon I probe (see Materials and Methods). Three independent positive clones, lambda Y, lambda

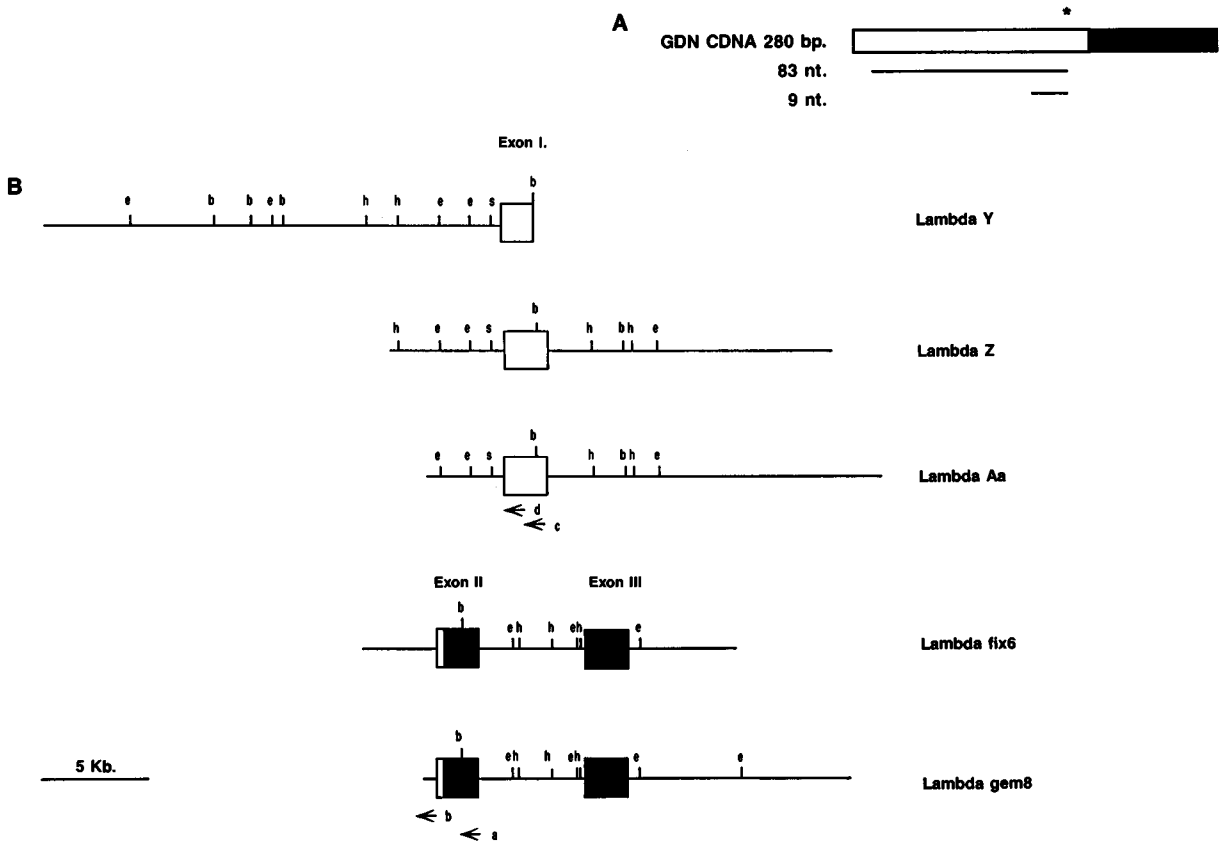

Figure 1. Maps of probes and isolated genomic clones. Open boxes represent 5′ noncoding sequences, filled boxes coding sequences. Exons are not drawn to scale. **A.** The 280 bp probe used for isolating the genomic clones containing the second and third exons of rat GDN/PN-1, and localization of the 83 nt and 9 nt oligonucleotides used for preparing the probe to isolate the rat GDN/PN-1 promoter clones. The asterisk shows the position of the splice site. **B.** Restriction map of isolated genomic clones. b: BamH I sites; e: EcoR I sites; h: Hind III; and s: Sal I sites. The localization of the primers used for primer extension are indicated by arrows. a is primer "exon II.1," b is primer "intron I.1," c is primer "exon I.1," and d is primer "exon I.2."

Z, and lambda Aa, were identified (Fig. 1B). All clones contained the noncoding exon I of the GDN/PN-1 gene. The lambda Y clone contained the first 105 bp of exon I upstream of the BamH I site and 21 kb of 5′ upstream sequence. The lambda Aa and lambda Z clones contained the whole exon I, as well as 5′- and 3′-flanking sequences (Fig. 1B). Mapping and Southern blot analysis showed that these clones did not overlap the clones lambda gem8 and lambda fix6 carrying exons II and III (data not shown). These results, in addition to Southern blot analysis of rat genomic DNA (not shown), indicated that the first intron was larger than 20 kb.

To identify the putative promoter region, a 1.6 kb EcoR I–BamH I fragment from the lambda Y clone was subcloned in the Bluescript SK(+) plasmid and sequenced. The nucleotide sequence of the first exon in the three clones (lambda Y, lambda Z, and lambda Aa) was identical to the GDN/PN-1 cDNA sequence, with the exception of a few mismatched bases within the first 15 bp of the cDNA sequence. This was probably due to cloning artifacts created during cloning of the GDN/PN-1 cDNA.

The rat GDN/PN-1 transcriptional start site, exon I upstream of the BamH I site, and the 5′-flanking sequences are shown in Figure 2. The promoter sequence from −1 to −300 and exon I were found to be very GC-rich (80%), containing a high number of CpG dinucleotides. In addition, CpGs and GpCs occurred with the same frequency, thus conforming with the cri-

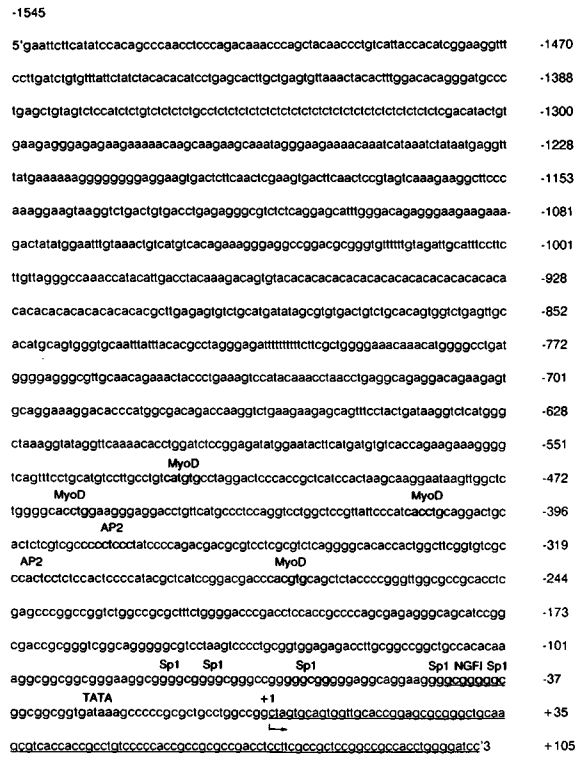

Figure 2. Sequence of rat GDN/PN-1 promoter. The sequence shown extends from the internal BamH I site in the first exon to the upstream EcoR I site shown in Figure 1B. The transcriptional start site is indicated by an arrow; the TATA box, Sp1, AP-2, and MyOD sites are shown in boldface. Sp1/NGFI/Sp1 indicates the overlapping bindings sites for Sp1 and NGFI-A/C, Krox-20, and Wilms tumor factor; the latter binding site is in boldface and is underlined. The exon sequence is underlined.

teria for a CpG island, common to promoter regions of many housekeeping genes (Bird, 1986). A number of potential transcriptional regulatory sites were identified in the promoter sequence. A TATA box-like element (TGATAAA) was found at position −28. The rat GDN/PN-1 promoter contained five high-affinity sites (GGGCGG) for the transcription factor Sp1 (Briggs et al., 1986) at positions −81, −76, −65, −46, and −40. Two potential AP-2 sites (Mitchell et al., 1987) were identified, with the consensus sequence [CCC(C/A)N(G/C)(G/C)(G/C)] at positions −384 and −305. Four consensus MyoD/E-box (CANNTG) sites (Lassar et al., 1989) are present at positions −524, −467, −410, and −279. The MyoD consensus sequences have also been shown to bind the Myc protein and other members of the MyoD transcription factor and E-box binding factor families (Lassar et al., 1989). A perfect consensus sequence (GCGGGGGCG) for the NGF early response genes NGFI-A (Christy and Nathans 1989), NGFI-C (Crosby et al., 1991), Krox 20 (Joseph et al., 1988), and Wilms tumor factor (Rauscher et al., 1990) was found at position −44, overlapping the fourth and fifth Sp1 sites. Furthermore, an unusual sequence was found at −1356 to −1312; it consisted of a CT dinucleotide repeat followed by a stretch of predominantly Gs and As (a so-called homopyrimidine-homopurine stretch). An AC dinucleotide repeat was found between positions −961 and −910.

Mapping of the transcriptional start point

To define the transcriptional start point of the rat GDN/PN-1 mRNA, primer extension and RNase protection experiments were performed. The primer extension experiments revealed one 130 nt extension product using the primer "exon II.1" (5′-agaagggaaaatgccaattcatgtttccttccaaca agga-3′; Fig. 3A, lane 5), which corresponds to the antisense sequence (Sommer et al., 1987) from codon 8 to 18 bp upstream of the ATG.

The primer "exon I.1" (5′-tcgcttggatccccaggt ggcggccggagcggcgaagagg-3′), corresponding to nucleotides +110 to +70 of the 5′ noncoding sequence, produced several bands, with a major band at 72 nt (Fig. 3A, lane 8). This band corresponded to the 130 nt extension product seen with the primer "exon II.1"; the four lower bands appeared to be premature termination products, since they did not correspond with the "exon II.1" primer bands or the band seen in the RNase protection experiment.

The primer "exon I.2" (5′-acaggcggtggtgac gcttgcgacccgcgctcgcctgcaacca-3′), corresponding to the position +5 to +52, produced 2 bands, indicating extension products of 2 or 3 nucleotides (Fig. 3A, lane 7). This was 4 to 5 nucleotides less than expected from the results seen with the other two primers. This discrepancy might be due to hybridization problems between the primer "exon I.2" and the 5′ end of the mRNA due to the presence of the cap or other secondary structures in the GDN/PN-1 mRNA.

No extension products were seen with the primer "intron I.1" (5′-tcagaaacatatcactcgtttgca aactaaaaataatagt-3′; Fig. 3A, lane 6), corresponding to the sequence just upstream of exon 2, thus confirming the presence of an intron within the 5′ noncoding mRNA leader sequence.

The RNase protection assay (Fig. 3B), carried

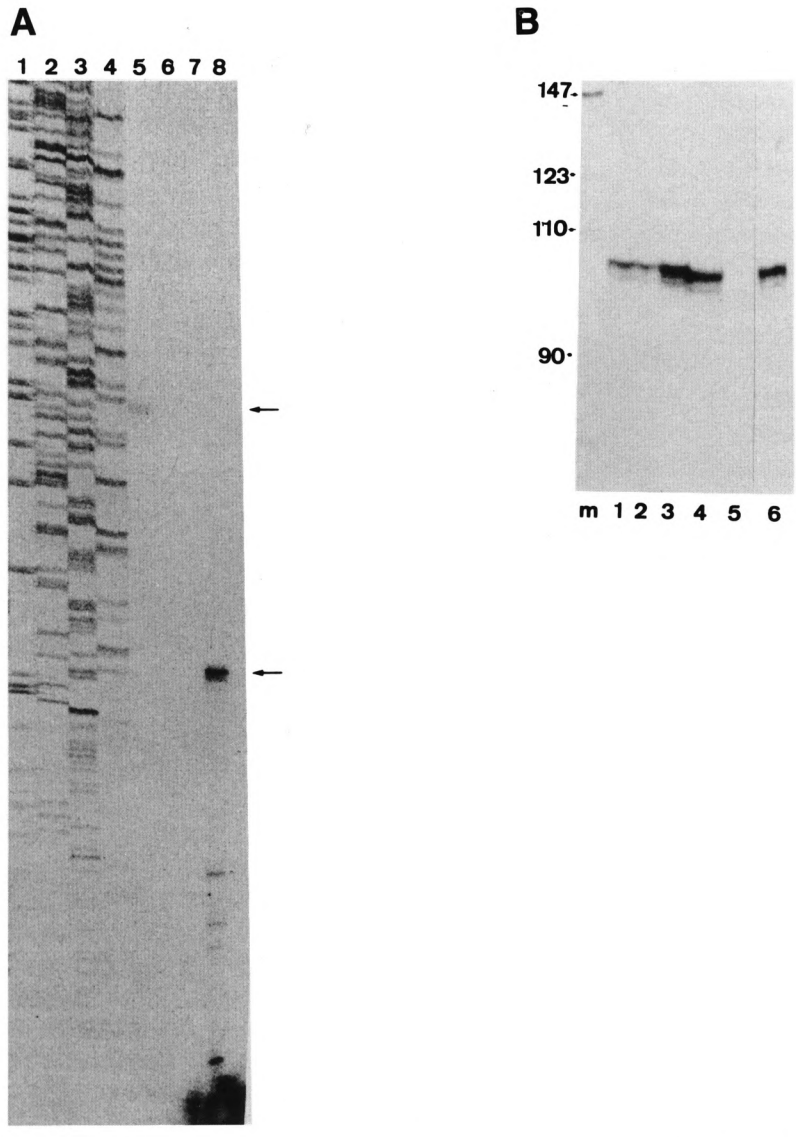

Figure 3. Primer extension (**A**) and RNase protection assays (**B**). Arrows indicate the presence of transcripts. **A.** Sequencing ladder (lane 1–4). Extension products with primer "exon II.1" (lane 5), "intron I.1" (lane 6), "exon I.2" (lane 7), and "exon I.1" (lane 8). Total RNA from C6 rat glioma cells was used in all cases. **B.** pBR Hga I marker numbers given in base pairs (lane m); 5 μg RNA from C6 (lane 1) and P2T (lane 2); 20 μg RNA from C6 (lane 3) and P2T (lane 4); negative control with yeast RNA (lane 5); RNA from rat Schwann cells (lane 6).

out with RNA from C6 rat glioma, P2T rat Schwannoma, and rat Schwann cells, showed two bands in all three cell lines: a major band at approximately 105 bases and a minor band at approximately 104 bases. Taken together, these results indicate that there is one transcriptional start point for the GDN/PN-1 gene.

Activity of the rat GDN/PN-1 promoter

To test whether the isolated clones contained any promoter activity, the three constructs described in Figure 4A were transfected into COS-7 cells and assayed for transient luciferase expression. The construct p500luc, containing 389 bp of the 5′-flanking region and 105 bp of the first exon, showed 1,000,000 cpm luciferase activity over background levels of 350 cpm. The p1600luc plasmid, containing 1545 bp of flanking sequence and 105 bp of exon 1, showed an additional 2.5-fold increase over that measured with p500luc, indicating that a positive regulatory element is present between positions −389 and −1545. No significant luciferase activity was detected after transfection with the control plasmid p-1600luc, in which the GDN/PN-1 promoter fragment was cloned in the reverse orientation, or in mock-transfected cells (Fig. 4B).

GDN/PN-1 mRNA expression varied substantially in different cell types, as shown in Figure 5. C6 glioma cells and P2T Schwannoma cells expressed GDN/PN-1 mRNA at very high levels, while Rat-1 fibroblasts and L6 myoblasts contained low levels of GDN/PN-1 mRNA. As in rat liver (Gloor et al., 1986), FTO2B hepatoma cells

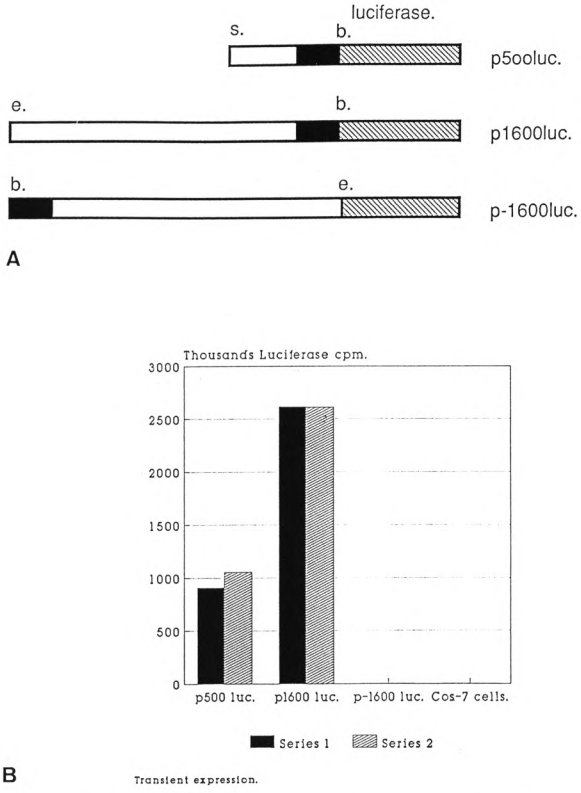

Figure 4. Luciferase assays. **A.** The transfected GDN/PN-1 promoter constructs. The p1600luc plasmid contains the promoter fragment from the internal BamH I site (b) of the first exon to the upstream EcoR I (−1545) site (e) shown in Figure 1B. The plasmid p-1600luc contains the same fragment but in the opposite orientation. The plasmid p500luc contains the promoter fragment from the BamH I (b) site upstream to the Sal I site (s; −389) indicated in Figure 1B. Open boxes represent 5′-flanking sequences, filled boxes 5′ noncoding sequences, and hatched boxes the luciferase gene. **B.** Luciferase activity in transiently transfected COS-7 cells. Series 1 and 2 represent two independent experiments.

contained no detectable GDN/PN-1 mRNA. This finding was confirmed by PCR analysis of total RNA from C6, Rat-1, and FTO2B cells. No PCR product was detected with FTO2B cells, while both C6 and Rat-1 cells were positive (data not shown).

To test whether the promoter fragments would confer cell-specific transcription, we transfected C6 glioma cells, P2T Schwannoma cells, L6 myoblasts, Rat-1 fibroblasts, and FTO2B hepatoma cells with the promoter plasmids mentioned above. As seen in Table 1A, all of the cell lines expressed high amounts of luciferase activity. The unexpectedly high amount of luciferase expression in the FTO2B hepatoma cells, a finding also seen with the H4 rat hepatoma cell line (data not shown), demonstrates that the promoter fragments tested are sufficient to confer a basal level of transcription, but that additional sequences or DNA methylation may be needed for cell-specific gene expression. Compared to the p500luc construct, a 2- to 4-fold increase in luciferase activity was obtained with the p1600luc construct in all cells, indicating that the putative positive element located between −389 and −1545 operates without cell specificity.

P2T Schwannoma cells and C6 glioma cells expressed very high amounts of GDN/PN-1 mRNA, while Rat-1 fibroblasts expressed low amounts (Fig. 5). However, in the transient transfection assays, Rat-1 cells express higher luciferase amounts (Table 1A). To determine whether this was due to differences in transfection efficiencies or promoter activity, we determined the DNA uptake in the various cell lines by transfection of the p1600luc plasmid followed by Hirt extraction of the plasmid two days after transfection. This showed that the DNA uptake in the glial cell lines C6 and P2T was much lower than in the other cell lines (C6 11 colonies, P2T 49 colonies, FTO2B 240 colonies, Rat-1 161 col-

Table 1A. Activity of the rat GDN/PN-1 promoter in different rat cell lines.

GDN/PN-1 promoter constructs were transfected into the indicated cell lines as described in Materials and Methods. The values shown represent luciferase activity found in the cell extracts.

	Plasmid	
Cell line	*p1600luc cpm*	*p500luc cpm*
C6 glioma	2,233	1,065
P2T Schwannoma	8,255	2,680
L6 myoblast	25,053	17,708
RAT-1 fibroblast	8,244	3,050
FTO2B hepatoma	10,747	4,537

Table 1B. Activity of the rat GDN/PN-1 promoter in stable transfected rat cell lines.

The GDN/PN-1 promoter construct p1600luc was transfected into the indicated cell lines, and pools of transfected cells were selected as described in Materials and Methods. The values shown represent luciferase activity found in the cell extracts.

Cell line	*cpm*
C6 glioma	253,270
RAT-1 fibroblast	24,837

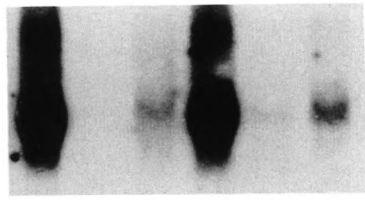

Figure 5. Northern blot showing GDN/PN-1 mRNA expression in different cell types. C: C6 glioma cells; F: FTO2B hepatoma cells; L: L6 myoblast; P: P2T Schwannoma cells; R: Rat-1 fibroblast; and Rb: rat brain. 10 μg total RNA were used in each lane.

onies, and L6 309 colonies). Transfection with the pSVII β-galactosidase plasmid and staining with X-gal two days after transfection gave similar results.

To determine whether the 1600 bp of the GDN/PN-1 promoter was sufficient to give high luciferase expression in glial cells, we selected pools of stable transfectants with the p1600luc plasmid of C6 glioma cells and Rat-1 fibroblasts. As seen in Table 1B, C6 glioma cells express tenfold higher luciferase amounts than Rat-1 fibroblasts. The copy number in the two stable lines was determined by hybridizing a genomic Southern blot with a luciferase cDNA probe. The copy number in the C6 cell line was 2–3 times higher than the copy number in the Rat-1 line.

The higher expression of luciferase in the stable C6 glioma transfectants demonstrates that the promoter sequence from −1545 to +105 operates in a cell-specific manner. Taken together with the finding that the sequence from −1545 to −389 can be deleted without any cell-specific effects (Table 1A), this leads to the conclusion that sufficient information for high glial expression is found within the fragment from −389 to +105. However, sequences located further upstream or in the first intron may contribute to cell-specific gene expression.

Effect of CpG methylation of the GDN/PN-1 promoter on gene expression

The high number of CpG dinucleotides in the GDN/PN-1 promoter suggested that DNA methylation may play a role in GDN/PN-1 gene regulation. To test this hypothesis, the p1600luc plasmid was methylated in vitro with the CpG methylase SssI prior to transfection. The results shown in Table 2 demonstrate that CpG methylation of GDN/PN-1 promoter resulted in a complete cell-specific block of GDN/PN-1 promoter activity in FTO2B hepatoma cells. By comparison, methylation of the plasmid pPALU, carrying the SV40 early promoter, reduced the luciferase activity by 90% of that seen with unmethylated pPALU. A decrease in the levels of transcription from the SV40 promoter after methylation has been attributed to the formation of inactive chromatin over the whole plasmid (Adams et al., 1990). This suggests that the complete silencing of the GDN/PN-1 promoter must be due to additional factors—a hypothesis supported by the observation that methylated p1600luc and pPALU plasmids both show only a 40% reduction of luciferase activity in C6 glioma cells (Table 2). These findings suggest that methylation per se is *insufficient* to silence the GDN/PN-1 promoter, and that additional cell-specific repressors are necessary for silencing the GDN/PN-1 promoter (−1545 to +105) in FTO2B hepatoma cells.

To test whether CpG methylation would have any effect on DNA–protein interaction, band shift assays were performed with methylated and non-methylated oligonucleotides and fractionated rat liver nuclear extracts (as well as crude nuclear extract from FTO2B hepatoma cells). Figure 6 shows the result of such an experiment: both in the 0.1 M (lane 2 and 9) and 0.2 M (lane 3 and 10) KCl fractions, a band of high molecular weight is seen using methylated

Table 2. Effect of in vitro CpG methylation by Sss I methylase.

The GDN/PN-1 promoter plasmid p1600luc and the construct with the SV40 promoter (pPALU) were methylated in vitro and transfected into the indicated cell lines as described in Materials and Methods. The values shown represent the luciferase activity found in the cell extracts.

	C6 glioma		*FTO2B hepatoma*	
Plasmid	*cpm*	*decrease (%)*	*cpm*	*decrease (%)*
p1600luc	3,124		208,905	
p1600luc CpG Meth.	1,818	42	0	100
SV 40luc	34,890		94,098	
SV 40luc CpG Meth.	19,620	44	8,280	91

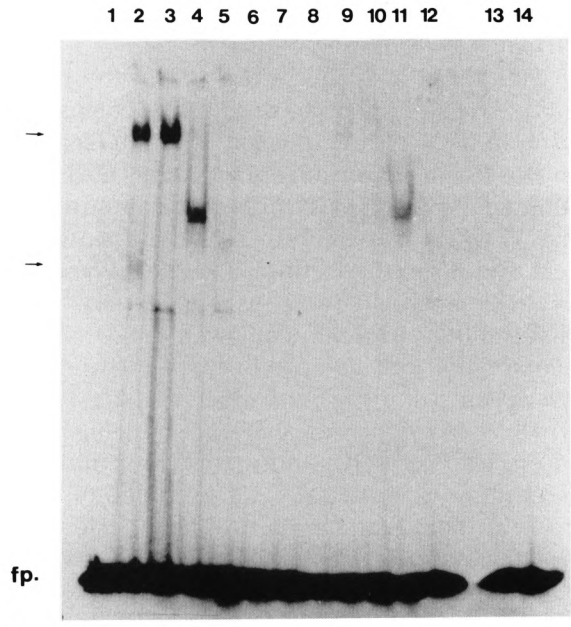

Figure 6. Gel shift assay of methylated and unmethylated oligonucleotides (see Materials and Methods) from the GDN/PN-1 promoter. Fractionated nuclear extracts from rat liver were incubated with methylated (lanes 1–7) and unmethylated DNA (lanes 8–14). Fractionated nuclear extract was added as follows: no extract (lanes 1 and 8), 0.1 M KCl (lanes 2 and 9), 0.2 M KCl (lanes 3 and 10), 0.3 M KCl (lanes 4 and 11), 0.4 M KCl (lanes 5 and 12), 0.5 M KCl (lanes 6 and 13), 0.6 to 1 M KCl (lanes 7 and 14). Arrows indicate the presence of methylation-specific mobility shifts, fp. the free probe.

and non-methylated probes. However, the signal is much stronger with the methylated probe. Whether this reflects the same proteins binding both methylated and unmethylated DNA—but with higher affinity for CpG methylated DNA—or two different set(s) of proteins with the same electrophoretic mobility is not clear at the moment. Additionally, CpG methylated DNA is bound by a protein of low molecular weight in the 0.1 M KCl fraction (lane 2). No significant differences are seen in the other fractions. These results support a model for cell-specific silencing of the GDN/PN-1 promoter in FTO2B cells and liver by proteins that bind specifically to CpG methylated sequences within the GDN/PN-1 promoter region.

Discussion

Three different overlapping genomic clones, all containing the first exon and the promoter of GDN/PN-1, were isolated and characterized. RNase protection assay and primer extension showed that the rat GDN/PN-1 gene had one major transcriptional start point in all cells tested. The promoter-containing fragments are able to direct basal transcription and give high expression in glial cells.

None of the isolated promoter clones overlapped the genomic clones that contained the second and third exons, indicating that the first intron is probably larger than 20 kb (Fig. 1). This is in accord with preliminary data published on the genomic organization of the human GDN/PN-1 gene (McGrogan et al., 1990). In addition, the position of splice sites of exons I, II, and III are conserved between the rat and the human genes.

The GDN/PN-1 promoter contains several unusual sequence features. The high CG content and the high number of CpGs demonstrate that the GDN/PN-1 promoter is localized within a CpG island (Bird, 1986). This is often the case for the promoters of housekeeping genes. Housekeeping promoters are also characterized by the absence of CAAT and TATA boxes and multiple transcriptional start sites. In the case of the rat GDN/PN-1 promoter, no CAAT box was found, but a TGATAAA sequence was found 28 bp upstream of the major transcriptional start point. Whether or not this sequence functions as a binding site for the TFIID transcription factor remains unproven, but it has been shown that yeast TFIID can also bind to TATA elements with non-consensus DNA sequences (Hahn et al., 1989). In contrast to housekeeping genes, GDN/PN-1 is not ubiquitously expressed; GDN/PN-1 mRNA levels are high in the brain (Fig. 5), low in the heart, and not detectable in the liver (Gloor et al., 1986). On the basis of these criteria, the GDN/PN-1 promoter does not fall into the class of traditional housekeeping promoters.

The GDN/PN-1 promoter sequence contains a CT dinucleotide repeated 21 times, followed by a GA-rich sequence containing numerous direct and mirror repeats. Such sequences are found to form unusual structures, such as H-DNA or cruciforms (for sequences with small mirror repeats; Mirkin et al., 1987), and several non-Watson-Crick structures have been proposed for CT repeats (reviewed by Chomet, 1991). Homopyrimidine-homopurine sequences exhibit S1 nuclease hypersensitivity and are fre-

quently found in the promoter region of eucaryotic genes, such as rat N-CAM and others (Chen et al., 1990), and can influence gene regulation by changes in chromatin structure and DNA–protein interaction (Wells et al., 1988). An AC repeat stretches from −956 to −905. Such repeats are known to form Z-DNA structures, which also exhibit S1 nuclease hypersensitivity (Thomas et al., 1990). It is interesting to note that deletion of the sequences from −1545 to −389, which contain these unusual sequences, results in a 2- to 4-fold decrease in promoter activity in different cell lines.

The difference in GDN/PN-1 mRNA levels between P2T Schwannoma cells and C6 glioma cells on one side and L6 myoblast and Rat-1 fibroblasts on the other indicates that cell-specific factors must regulate the GDN/PN-1 promoter. The target sites for these factors must primarily be localized in the GC-rich proximal (−389 to +105) part of the promoter sequence, since the upstream fragment from −1545 to −389 can be deleted without any cell-specific effects, although sequences in the intron or in the far upstream region may play a role in cell-specific gene expression. Several consensus sites for transcription factors are found in the proximal part of the GDN/PN-1 promoter, including five Sp1 sites. Sp1 is an ubiquitous transcription factor found in most tissues, although at different concentrations. Sp1 is found at higher concentration in the brain than in the liver (Saffer et al., 1991), an expression pattern compatible with the expression of rat GDN/PN-1 mRNA.

Of special interest to us was the finding that the GDN/PN-1 promoter contains a binding site for the NGFI-A and NGFI-C factors, Wilms tumor factor, and Krox-20. NGFI-A and NGFI-C are induced in immediate response to NGF treatment of PC12 pheochromocytoma cells (Crosby et al., 1991; Joseph et al., 1988). Moreover, both NGFI-A and NGFI-C are induced after seizure in the brain and/or addition of growth factors (Sukhatme et al., 1988), a finding which is interesting in the light of the increase of GDN/PN-1 mRNA after lesions in both the central and peripheral nervous system. Wilms tumor factor is a repressor found in the kidney and functions as an antagonist to NGFI-A (Madden et al., 1991). In the adult mouse, GDN/PN-1 mRNA is expressed at a low level in the kidney (F. M. Botteri, personal communication); Wilms tumor factor may therefore be involved in GDN/PN-1 regulation.

Strickland and co-workers (Pecorino et al., 1991) have identified a brain-specific GC-binding protein (BGC) that binds to a GC-rich element in the tissue plasminogen activator (tPA) promoter. Since the rat GDN/PN-1 promoter is also highly GC-rich and exhibits an expression pattern similar to that reported for tPA (Pecorino et al., 1990), one has to consider the possibility that the BGC protein is involved in GDN/PN-1 gene regulation. The promoter of the amyloid precursor protein gene (Pollwein et al., 1992), which is also highly expressed in brain, has been found to bind a GC-binding transcription factor different from Sp1. The presence of four consensus binding sites for the MyoD1 transcription factor (Lassar et al., 1989) is consistent with the expression of GDN/PN-1 mRNA in L6 rat myoblasts and the finding that GDN/PN-1 mRNA and protein are expressed in rat muscle (Festoff et al., 1991), where they are developmentally regulated (H. S. Suidan, unpublished data).

It has recently been shown that GDN/PN-1 is expressed and secreted at very high levels in the murine seminal gland and under transcriptional control by androgens (Vassalli et al., 1993). No androgen response element (Beato, 1989) is found in the sequenced part of the rat GDN/PN-1 promoter, but it may be located further upstream or in the intron. Another possibility is that the transcriptional upregulation by testosterone is mediated by other transcription factors.

Methylation of CpG dinucleotides has been suggested as one mechanism for cell-specific silencing of genes (Chomet, 1991). CpG methylation has been shown to alter chromatin structure in several systems, and changes in methylation correlate with changes in gene expression. Gene transcription was found to be inhibited by in vitro methylation by the SssI methylase of the MyoD1 promoter (Zingg et al., 1991) and mouse metallothionein I (mMTI) promoter (Levine et al., 1991). Recently it was shown by Levine et al. (1992) that CpG methylation of the preinitiation domain of the mMTI promoter suppresses gene transcription. In this case, suppression was probably mediated by binding of a methylation-dependent DNA-binding protein.

In many cell lines, gene inactivation by methylation of CpG islands has been proposed to

account for the loss of cell-specific gene expression. In the case of the GDN/PN-1 promoter, the ability of CpG methylation in vitro to completely abolish gene expression in FTO2B cells but not in C6 glioma cells suggests that cell-specific transcriptional repressors are present in liver cells. Whether there is a difference in the methylation state between the endogenous GDN/PN-1 promoters in C6 glioma and FTO2B hepatoma cells remains to be shown. A number of transcription factors have been isolated that either bind only to methylated sequences at specific sites or to methyl-CpGs, regardless of sequence (Boyes et al., 1991; Pawlak et al., 1991). Whether the CpG methylation-dependent factors binding to the GDN/PN-1 sequences are new factors or are identical to those previously described remains to be shown.

Comparison of the rat GDN/PN-1 promoter with the promoters of plasminogen activator inhibitor 1 (Riccio et al., 1988) and plasminogen activator inhibitor 2 (Kruithof et al., 1988) does not show significant homology, a finding consistent with the different expression patterns of the three serpins. The organization of the GDN/PN-1 promoter described here represents an important initial step for further investigation of the complex regulation of GDN/PN-1 expression.

Acknowledgments

We would like to thank Dr. P. Naef, Basel, and Dr. D. Ganten, Heidelberg, for the gift of genomic libraries. We are also grateful for the gift of RNA from cultured rat Schwann cells (A. Bleuel), preparation of oligonucleotides (W. Zuercher and F. Fischer), and isolation of phages lambda gem8 and lambda fix 6 (A. Ultsch). The P2T Schwannoma cells were a kind gift of Dr. S. Pfeiffer (Hartford, CT). We would like to thank Dr. J.-P. Jost for his suggestions and constant support, E. Fries for excellent technical help, and Drs. F. M. Botteri, J.-P. Jost, H. S. Suidan, and P. Caroni for their comments and critical reading of the manuscript.

H. Ernø was supported in part by a grant from the Danish Research Academy.

The rat GDN/PN-1 promoter sequence reported in this paper has been submitted to the EMBL Data Library and has been assigned accession number X71791 RNGDNPN1.

The costs of publishing this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC Section 1734 solely to indicate this fact.

References

R. L. P. Adams, M. Bryans, A. Rinaldi, and H. M. I. Yesufu (1990), Philos Trans R Soc Lond Biol 326, 189–198.

P. Artelt, R. Grannemann, C. Stocking, J. Friel, J. Bartsch, and H. Hauser (1991), Gene 99, 249–254.

M. Beato (1989), Cell 56, 335–344.

A.P. Bird (1986), Nature 321, 209–213.

J. Boyes and A. P. Bird (1991), Cell 64, 1123–1134.

M. R. Briggs, J. T. Kadonaga, S. P. Bell, and R. Tjian (1986), Science 234, 47–52.

S. P. Chomet (1991), Curr Opin Cell Biol 3, 438–443.

A. Chen, A. Reyes, and R. Akeson (1990), Mol Cell Biol 10, 3314–3324.

B. Christy and D. Nathans (1989), Proc Natl Acad Sci USA 86, 8737–8741.

S. D. Crosby, J. J. Puetz, K. S. Simburger, T. J. Fahrner, and J. Milbrandt (1991), Mol Cell Biol 11, 3835–3841.

J. R. de Wet, K. V. Wood, M. DeLuca, D. R. Helinski, and S. Subramani (1987), Mol Cell Biol 7, 725–737.

J. D. Dignam, R. M. Lebowitz, and R. G. Roeder (1983), Nucleic Acids Res 11, 1475–1489.

L. Farmer, J. Sommer, and D. Monard (1990), Dev Neurosci 12, 73–80.

B. W. Festoff, J. S. Rao, and D. Hantai (1991), J Cell Physiol 147, 76–86.

S. Gloor, K. Odink, J. Guenther, H. Nick, and D. Monard (1986), Cell 47, 687–693.

J. Guenther, H. Nick, and D. Monard (1985), EMBO J 4, 1963–1966.

S. Hahn, S. Buratowski, P. A. Sharp, and L. Guarente (1989), Proc Natl Acad Sci USA 86, 5718–5722.

B. Hirt (1967), J Mol Biol 26, 365–372.

M. C. Hoffmann, C. Nitsch, A. L. Scotti, E. Reinhard, and D. Monard (1992), Neuroscience 49, 397–408.

L. J. Joseph, M. M. Le Beau, G. A. Jamieson, D. Acharya, T. B. Show, J. D. Rowley, and V. P. Sukhatme (1988), Proc Natl Acad Sci USA 85, 7164–7168.

E. K. O. Kruithof and E. Cousin (1988), Biochem Biophys Res Commun 156, 383–388.

A. B. Lassar, J. N. Buskin, D. Lockshon, R. L. Davis, S. Apone, S. D. Hauschka, and H. Weintraub (1989), Cell 58, 823–831.

A. Levine, G. L. Cantoni, and A. Razin (1991), Proc Natl Acad Sci USA 88, 6515–6518.

A. Levine, G. L. Cantoni, and A. Razin (1992), Proc Natl Acad Sci USA 89, 10119–10123.

S. L. Madden, D. M. Cook, J. F. Morris, A. Gashler, V. K. Sukhatme, and F. J. Rauscher (1991), Nature 253, 1550–1153.

M. McGrogan, J. Kennedy, M. P. Li, C. Hsu, R. W. Scott, C. C. Simonsen, and J. B. Baker (1988), Bio-Technology 6, 172–177.

M. McGrogan, J. Kennedy, F. Golini, N. Ashton, F.

Dunn, K. Bell, E. Tate, R. W. Scott, and C. C. Simonsen (1990), in Serine Proteases and Their Serpin Inhibitors in the Nervous System (B. W. Festoff, ed.), Plenum Press, New York, pp. 147–161.

R. Meier, P. Spreyer, R. Ortmann, A. Harel, and D. Monard (1989), Nature 342, 548–550.

P. J. Mitchell, C. Wang, and R. Tjian (1987), Cell 50, 847–861.

S. M. Mirkin, V. I. Lyamichev, K. N. Drushlyak, V. N. Dobrynin, S. A. Filippov, and M. D. Franck-Kamenetskii (1987), Nature 330, 495–497.

D. Monard, F. Solomon, M. Rentsch, and R. Gysin (1973), Proc Natl Acad Sci USA 70, 1894–1897.

A. Pawlak, M. Bryans, and J.-P. Jost (1991), Nucleic Acids Res 19, 1029–1034.

L. T. Pecorino, A. L. Darrow, and S. Strickland (1991), Mol Cell Biol 11, 3139–3147.

P. Pollwein, C. L. Masters, and K. Beyreuther (1992), Nucleic Acids Res 20, 63–68.

F. J. Rauscher III, J. F. Morris, O. E. Tournay, D. M. Cook, and T. Curran (1990), Science 250, 1259–1262.

E. Reinhard, R. Meier, W. Halfter, G. Rovelli, and D. Monard (1989), Neuron 1, 387–394.

A. Riccio, L. R. Lund, R. Sartorio, A. Lania, P. A. Andreasen, K. Dano, and F. Blasi (1988), Nucleic Acids Res 16, 2805–2824.

J. D. Saffer, S. P. Jackson, and M. B. Annarella (1991), Mol Cell Biol 11, 2189–2199.

J. Sambrook, E. F. Fritsch, and T. Maniatis (1989), Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

J. Sommer, S. M. Gloor, G. F. Rovelli, J. Hofsteenge, H. Nick, R. Meier, and D. Monard (1987), Biochemistry 26, 6407–6410.

S. R. Stone, H. Nick, J. Hofsteenge, and D. Monard (1987), Arch Biochem Biophys 252, 237–244.

H. S. Suidan, S. R. Stone, B. A. Hemmings, and D. Monard (1992), Neuron 8, 363–375.

V. P. Sukhatme, X. Cao, L. C. Chang, C. H. Tsai-Morris, D. Stamenkovich, P. C. P. Ferreira, D. R. Cohen, S. A. Edwards, T. B. Shows, T. Curran, M. M. LeBeau, and E. D. Adamson (1988), Cell 53, 37–43.

M. J. Thomas, T. M. Freeland, and J. S. Strobl (1990), Mol Cell Biol 10, 5378–5387.

J.-D. Vassalli, J. Huarte, D. Bosco, A.-P. Sappino, N. Sappino, A. Wohlwend, H. Ernø, D. Monard, and D. Belin (1993), EMBO J 12, 1871–1878.

R. D. Wells, D. A. Collier, J. C. Hanvey, M. Shimizu, and F. Wohlrab (1988), FASEB J 2, 2939–2949.

J.-M. Zingg, P. G. Alva, and J.-P. Jost (1991), Nucleic Acids Res 19, 6433–6439.

A. D. Zurn, H. Nick, and D. Monard (1988), Dev Neurosci 10, 17–24.