Oncogene (1999) 18, 5836–5842

http://www.stockton-press.co.uk/onc

SHORT REPORT

Molecular cloning and characterization of the 5′ region of the mouse *trk*A proto-oncogene

María P Sacristán[1], Juana G de Diego[2], Magdalena Bonilla[1] and Dionisio Martín-Zanca*[,1]

[1]*Instituto de Microbiología Bioquímica, CSIC/Universidad de Salamanca, Avda. Campo Charro s/n, 37007 Salamanca, Spain;* [2]*Departamento de Bioquímica y Biología Molecular, Universidad de Salamanca, Spain*

The *trk*A proto-oncogene encodes a high-affinity NGF receptor that is essential for the survival, differentiation and maintenance of many neural and non-neural cell types. Altered expression of the *trk*A gene or *trk*A receptor malfunction have been implicated in neurodegeneration, tumor progression and oncogenesis. We have cloned and characterized the 5′ region of the mouse *trk*A gene and have identified its promoter. *trk*A promoter sequences are GC-rich, lack genuine TATA or CAAT boxes, and are contained within a CpG island which extends over the entire first coding exon. The mouse *trk*A transcription start site is located 70/71 bp upstream to the AUG translation initiation codon. Sequence analysis showed that the gene encoding the insulin receptor-related receptor, IRR, is located just 1.6 kbp upstream to the *trk*A gene and is transcribed in the opposite direction. We have used *trk*A-CAT transcriptional fusions to study *trk*A promoter function in transient transfection experiments. RNase protection assays and CAT protein ELISA analyses showed that a 150 bp long DNA segment, immediately upstream to the start site, is sufficient to direct accurate transcription in *trk*A-expressing cells. Dissection of this fragment allowed us to identify a 13 bp *cis*-regulatory element essential for both promoter activity and cell-type specific expression. Deletion of this 13 bp segment as well as modification of its sequence by site-directed mutagenesis led to a dramatic decline in promoter activity. Gel mobility shift assays carried out with double-stranded oligonucleotides containing the 13 bp element revealed several specific DNA-protein complexes when nuclear extracts from *trk*A-expressing cells were used. Supershift experiments showed that the Sp1 transcription factor was a component of one of these complexes. Our results identify a minimal *trk*A gene promoter, located very close to the transcription start site, and define a 13 bp enhancer within this promoter sequence.

Keywords: *trk*A; proto-oncogene; promoter; NGF receptor; CpG island; IRR

The *trk*A proto-oncogene was originally isolated from a colon carcinoma as a transforming oncogene activated by a somatic rearrangement that fused a non-muscle tropomyosin gene with the receptor tyrosine kinase-encoding *trk*A gene (Martin-Zunca *et al*., 1986, 1989). Similar mechanisms are responsible for the malignant activation of *trk*A in a significant fraction of papillary thyroid carcinomas (Greco *et al*., 1997). *trk*A is the prototype of a family of genes, which includes *trk*B and *trk*C, encoding high affinity receptors for the neurotrophins of the nerve growth factor family (Barbacid, 1995). Neurotrophins are responsible for the survival, differentiation and maintenance of specific populations of neurons in the developing and adult nervous system (Davies, 1994). Expression of *trk*A is developmentally regulated at the transcriptional level; thus, *trk*A mRNA first appears at day 9.5 of mouse embryonic development (E9.5) and is restricted to neural crest-derived neurons which are condensing to form sensory ganglia, where *trk*A expression reaches a maximum at E13.5 (Martin-Zanca *et al*., 1990). Expression of *trk*A in sympathetic neurons starts at E13.5 and reaches highest levels at E18, coinciding with the time of their maximum response to NGF (Ernfors *et al*., 1992). In the central nervous system *trk*A mRNA has been found in NGF-responsive cholinergic neurons of the basal forebrain and neostriatum and in some non-cholinergic neurons in several nuclei at different regions of the brainstem (Holtzman *et al*., 1995). *trk*A expressing neurons closely match those that are NGF-dependent for survival and differentiation, suggesting that *trk*A could mediate NGF actions *in vivo*. Consistently, *trk*A knock-out mice are insensitive to heat and pain, and show extensive loss of small-diameter neurons in the trigeminal (TGG) and dorsal root (DRG) ganglia, which are known to convey nociception and thermoreception stimuli. They are also devoid of sympathetic neurons, and have reduced numbers of cholinergic projections to the hippocampus and cortex (Fagan *et al*., 1996). Recently, sensitive *in situ* hybridization and RNase protection analyses have unveiled widespread expression of *trk*A in many non-neuronal tissues, including kidney, thymus, testes, stomach and muscle (Lomen-Hoerth and Shooter, 1995; Wheeler *et al*., 1998). It has also been shown that NGF-triggered *trk*A signaling is involved in many autocrine and paracrine circuits playing important roles in the growth, survival and/or differentiation of different cell types, including mast cells and B memory cells (Tam *et al*., 1997; Torcia *et al*., 1996; skin melanocytes (Yaar *et al*., 1994); and cells from the ovary (Mayerhofer *et al*., 1996); prostate (Dalal *et al*., 1997); and pancreas (Miralles *et al*., 1998). Loss of *trk*A function is responsible for a human disease, congenital insensitivity to pain with anhidrosis (CIPA), characterized by lack of reaction to noxious stimuli (Indo *et al*., 1996). In human neuroblastomas, expression of *trk*A is a good prognostic marker, suggesting that lack of *trk*A expression contributes to

*Correspondence: D Martin-Zanca
Received 15 October 1998; revised 11 May 1999; accepted 11 May 1999

malignancy, perhaps because it results in the loss of signaling pathways important for growth arrest and/or differentiation of neural crest-derived cells (Brodeur *et al*., 1997). Finally, *trk*A reduced expression has been reported in specific cholinergic neuronal populations which are affected in Alzheimer's disease (Boissiere *et al*., 1997). The IRR gene encodes a receptor tyrosine kinase, very similar to the insulin receptor, whose pattern of expression is essentially identical to that of *trk*A. Thus, IRR is co-expressed, at the cellular level, with *trk*A in sensory and sympathetic ganglia in the peripheral nervous system (Reinhardt *et al*., 1994), as well as in some neurons in nuclei of the basal forebrain and brainstem (Tsuji *et al*., 1996).

Unveiling the mechanisms controlling *trk*A gene expression will enhance our understanding on the development and function of the nervous as well as those of other systems, and will shed light into the genesis of some human diseases, including neurodegeneration and cancer. To this end, we have cloned and characterized the promoter region of the mouse *trk*A gene. A 15 kbp genomic clone, encompassing the 5′ region of *trk*A was isolated from a 129 SvJ mouse library (Figure 1a). Nucleotide sequence analysis allowed us to identify the first *trk*A coding exon as well as the first exon of the mouse IRR gene. This gene encodes a receptor tyrosine kinase highly related to the insulin receptor that is transcribed in the opposite

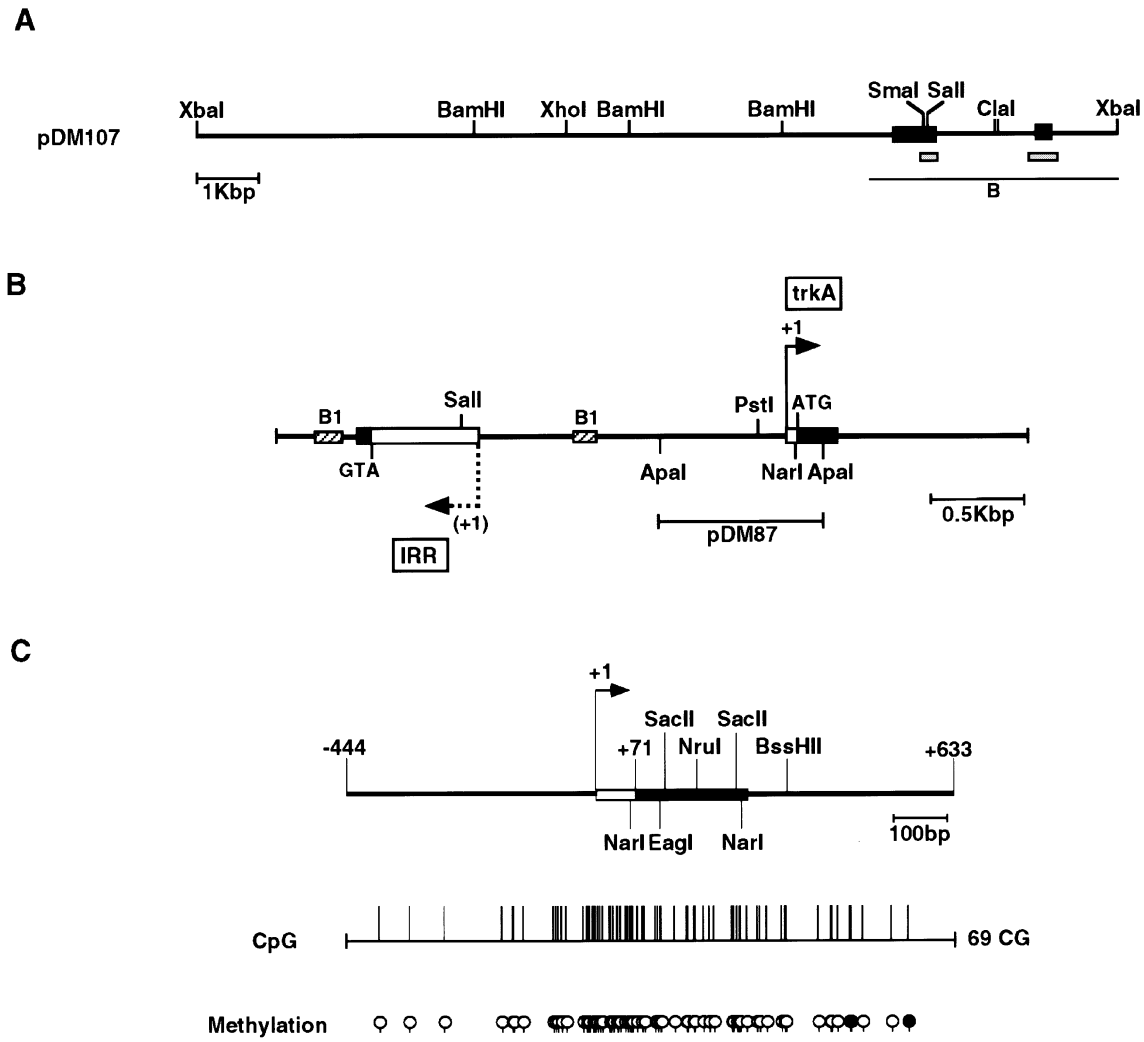

Figure 1 Genomic organization of the 5′ region of the mouse *trk*A gene. (**a**) Physical map of plasmid pDM107, encompassing the first exon of *trk*A as well as that of the mouse IRR gene (black boxes). pDM107 was derived from a genomic lambda clone isolated from a 129 SvJ mouse library, using a probe corresponding to the 5′ end of a previously described mouse *trk*A clone (Martin-Zanca *et al*., 1990). Shadowed boxes indicate the location of putative CpG islands identified by GRAIL: Gene Recognition and Assembly Internet Link (http://compbio.ornl.gov/Grail-1.3). (**b**) Detailed organization of the 3′ region of plasmid pDM107 (B-labeled line in **a**). Bent arrows indicate the location of the transcription start site of mouse *trk*A and the putative start site of the IRR gene. Open and filled boxes represent non-coding and coding exon sequences, respectively. B1 repetitive sequences are shown as crosshatched boxes. pDM87 is a plasmid used for S1 nuclease mapping of the *trk*A 5′ end. (**c**) Structure of the CpG island encompassing the first exon of mouse *trk*A. Nucleotides are numbered with respect to the *trk*A start of transcription (bent arrow, +1). High definition mapping of cytosine methylation of a 1078 bp DNA fragment encompassing the first exon (−444 to +633) was carried out as described by Frommer and colleagues (Clark *et al*., 1994). The distribution of the CG dinucleotides as well as their methylation status is shown. Open circles represent CpGs that are not methylated in mouse liver DNA; filled circles indicate the two CpGs that are methylated

direction to that of *trk*A, from the complementary DNA strand (Figure 1b). The transcription start site of the mouse IRR gene has not been reported; however, a putative TATA box, located 29 bp upstream to the start site of the rat gene, is conserved in the mouse gene, suggesting that the position of the start of transcription may be also conserved (Shier and Watt, 1992). The nucleotide sequence around the first coding exon of *trk*A is very GC-rich, suggesting that it may correspond to a CpG island (Bird, 1986). Sequence analysis using GRAIL (Gene Recognition and Assembly Internet Link) revealed that a 443 bp island, with a GC content of 73.38% and a ratio of observed/expected CpGs of 0.81, covers the first exon of *trk*A and extends some 190 bp upstream to the AUG translation initiation codon. Methylation analyses carried out using methylation-sensitive restriction enzymes showed that the CpG pairs contained within the *Sac*II, *Nru*I, *Bss*HII and *Eag*I recognition sites (Figure 1c) were not methylated in DNA from mouse tissues such as liver, which does not express *trk*A (not shown). High definition mapping of cytosine methylation, carried out following the procedure described by Frommer and colleagues (Clark *et al*., 1994), showed that, of the 69 CG dinucleotides included in a 1078 bp region around the first *trk*A exon, only two CG pairs located at the 3′ end of this region are methylated in liver DNA (Figure 1c). These results confirmed that the first coding exon of mouse *trk*A is contained within a CpG island. GRAIL analysis of the 5′ end of human *trk*A showed that it is also included in a putative CpG island, 563 bp long, with a GC content of 66.53% and a CpG score of 0.68 (not shown). Interestingly, a short cluster of CG dinucleotides within the 5′ end of the mouse IRR gene was identified by GRAIL as a 211 bp CpG island, with a GC content of 69.55% and a CpG score of 0.72 (Figure 1a).

Previous *in situ* hybridization analyses had shown that *trk*A is expressed in neural crest-derived sensory and sympathetic ganglia during mouse embryonic development. In this study we have analysed the expression of *trk*A by Northern blot using total RNA from kidney, spinal cord, and isolated DRG and TGG from E15–17 mouse embryos. As shown in Figure 2a, high levels of a 3.0 kbp *trk*A-specific mRNA were detected in DRG and TGG; lower expression levels were observed in spinal cord, while no expression was detected in kidney. The transcription start site of the mouse *trk*A gene was identified by a combination of primer extension and S1 nuclease protection analyses, using RNA from isolated DRG of E15–17 mouse embryos. As shown in Figure 2b, the extended cDNA products, 152 and 153 nucleotides in length, comigrated with the S1 digestion-protected fragments, thus locating the start of transcription of the mouse *trk*A gene at a CG dinucleotide, 71/70 nucleotides upstream to the AUG translation initiation codon, and showing that the first exon of the gene includes the AUG. The start of transcription of mouse *trk*A is just 1.6 kbp away from the putative start site of the IRR gene (Figure 1b).

To identify *trk*A genomic sequences containing promoter activity, transcriptional fusions between *trk*A and the bacterial chloramphenicol acetyl transferase (CAT) gene were generated and used in transient transfection assays on two cell lines which express *trk*A, N2a and PC12, as well as in the human epithelial cell line 293 which does not express *trk*A. Reporter gene expression was monitored by RNase protection or by

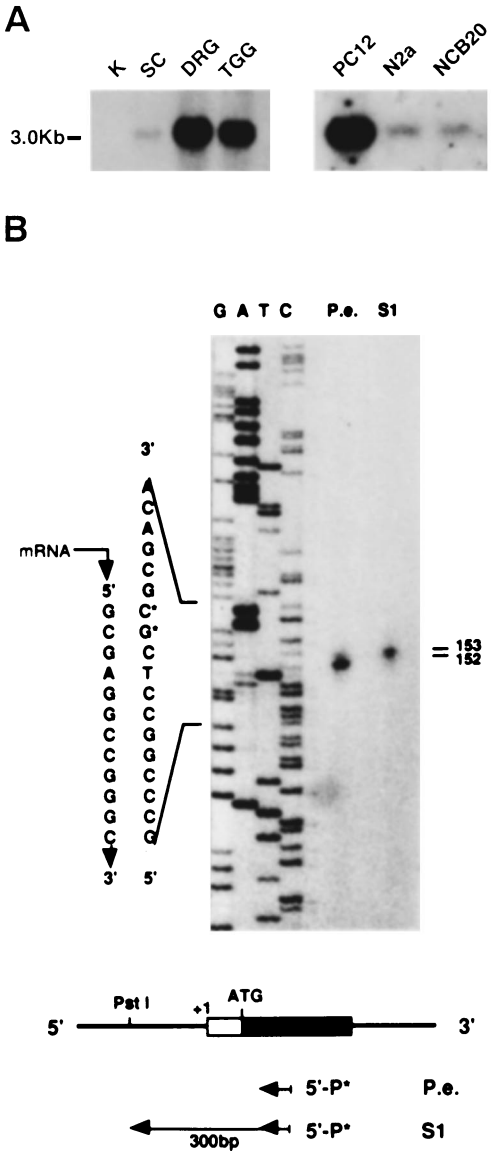

Figure 2 Identification of the transcription start site of the mouse *trk*A gene. (**a**) Northern blot analysis. 10 μg of total RNA were hybridized to a random primed ^{32}P-labeled 850 bp fragment corresponding to the 5′ region of a mouse *trk*A cDNA, as described (Martin-Zanca *et al*., 1990). DRG and TGG: dorsal root and trigeminal ganglia from E15–17 mouse embryos; K: kidney; SC: spinal cord; PC12: rat pheochromocytoma cell line; N2a; mouse neuroblastoma cell line; NCB20: mouse neuroblastoma-hamster neuron hybrid. (**b**) Primer extension and S1 nuclease digestion analyses of 20 μg of RNA from mouse embryonic DRG. For primer extension, a 5′-end-labeled oligonucleotide (N29: 5′-AGCATCAACGAAGTCATCAGAC-TGCCTAGC-3′), complementary to the mouse *trk*A sequence between nucleotides +52 and +82, relative to the AUG codon, was used. For S1 nuclease digestion, a 300 nt long 5′-end-labeled, single-stranded DNA probe was generated using 5′-end-labeled N29 as primer and plasmid pDM87 (Figure 1b) as template. Primer extension and S1 nuclease digestion analyses were performed according to published protocols (Sambrook *et al*., 1989). Reaction products (lanes P.e. and S1) were resolved on a 6% denaturing poliacrylamide gel, along a DNA sequencing ladder obtained from the pDM87 clone using the N29 oligonucleotide. 5′-P*: 5′-end-labeled N29 primer; S1: single-stranded 5′-end-labeled probe

CAT protein determination. Transient transfection analyses showed that a DNA fragment extending 1.8 kbp upstream of the trkA transcription start site exhibited promoter activity in N2a and PC12 cells but not in the 293 cell line. In order to define a smaller region capable of directing reporter expression in N2a and PC12 cells, a series of progressive 5′ deletions were derived from pMS8 (Figure 3a). As shown in Figure 3b, all deletion constructs, except pMS22, which extends only 43 nucleotides upstream to the transcription start site, directed accurate transcription initiation in these cells, as evidenced by the protection of a 224 nt fragment corresponding to the 5′ end of the *trk*A-CAT chimeric mRNAs (Figure 3b, lanes 3–15). No such protection was observed in untransfected cells or in cells transfected with control pUC-CAT plasmid

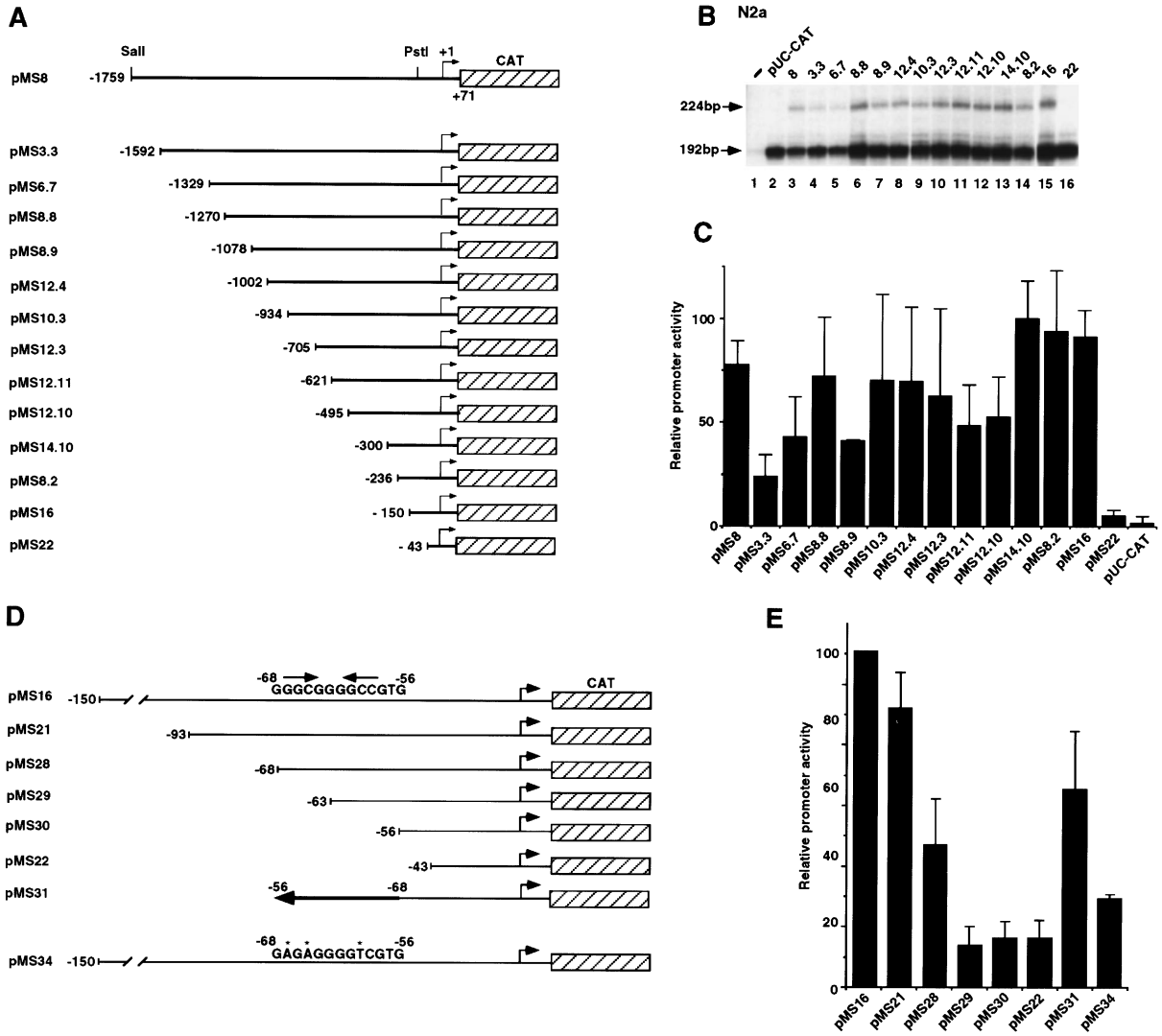

Figure 3 Functional characterization of the mouse *trk*A gene promoter. (**a** and **d**) *trk*A-CAT transcriptional fusions made on a promoter-less vector, pUC-CAT, containing four copies of the polyadenilation signal of SV40 upstream to the polylinker. pMS8 contains 1.8 kbp of *trk*A sequences extending from the *Sal*I site (at −1759) to the AUG codon (at +71). Progressive deletions from the 5′ end of this *trk*A fragment were generated by Exonuclease III or *Pst*I (pMS16) digestions. Shorter deletions (**d**) were generated by PCR. pMS31 contains the sequence −56 to −68 in the reverse orientation (underlined). Base substitutions in pMS34 (asterisks) were introduced according to the QuikChange Site-directed mutagenesis Kit protocol (Stratagene, La Jolla, CA, USA). Horizontal arrows indicate an inverted repeat. *trk*A-CAT plasmids (10 μg) and plasmid CMV-*lac*Z (1 μg), were co-transfected into N2a, PC12 or 293 cells by the calcium-phosphate method (Sambrook *et al*., 1989). After 40 h, reporter gene expression was determined by RNase protection assay or by measuring CAT protein levels using a CAT-ELISA kit (Boehringer Mannheim, Mannheim, Germany). In parallel, cell extracts were used to quantify β-galactosidase activity (Sambrook *et al*., 1989). (**b**) Representative RNase protection analysis of N2a transfected cells using a 397 nt riboprobe capable of identifying the 5′ end of the chimeric *trk*A-CAT mRNA. This fragment contains 153 nt complementary to the 5′ end of the CAT gene and 71 nt complementary to *trk*A transcript sequences, 5′ to the AUG codon. To identify the *lac*Z transcript, a 309 nt riboprobe containing 192 nt complementary to the 5′ end of the *lac*Z gene was used. 10 μg of total RNA were hybridized simultaneously to the two riboprobes and processed following standard protocols (Ausubel *et al*., 1995). Lanes 3–15 show a 224 nt protected fragment, specific for the 5′ end of the *trk*A-CAT fusion mRNA. The 192 nt protected fragment indicates expression of the CMV-*lac*Z internal control (lanes 2–16). (**c**) Relative CAT protein levels normalized with respect to β-galactosidase activity. CAT values (mean ± s.e., $n = 3–5$, duplicated experiments) are given as percentages relative to the promoter activity of plasmid pMS14.10 (100%). (**e**) Histogram showing normalized CAT protein levels. CAT values (mean ± s.e., $n = 2–8$, duplicated experiments) are given as percentages relative to the promoter activity of plasmid pMS16 (100%)

(Figure 3b, lanes 1 and 2). Similar results were obtained when the deletion constructs were transfected into the PC12 cell line; in contrast, none of the *trk*ACAT fusions exhibited promoter activity when transfected into the 293 cell line (not shown). Quantitative analyses of the promoter activity obtained by CAT protein determination, showed that all deletions, except for pMS22, exhibited significant levels of promoter activity in N2a cells (Figure 3c). These results indicated that a small region of 150 nt, 5′ to the *trk*A transcription start site, was capable of directing accurate, cell type-specific transcription, to levels comparable to those of fusions extending further upstream in *trk*A, indicating that *cis* elements responsible for cell type-specific transcription were located between nucleotides −150 and −43. To identify this essential *cis* element we dissected this region by generating four additional plasmids, pMS21, pMS28, pMS29 and pMS30 which extended 93, 68, 63 and 56 nucleotides upstream of the transcription start site of *trk*A, respectively (Figure 3d). Transient transfection assays showed that pMS21 and pMS28, extending to −93 and −98, exhibited 90 and 50% of the promoter activity of pMS16, respectively. However, removing five or 13 nucleotides at the 5′ end of pMS28, drastically reduced promoter activity, being the levels of CAT protein produced by pMS29 and pMS30 similar to those obtained with pMS22, a plasmid essentially incapable of directing accurate transcription as determined by RNase protection (Figure 3e). These results suggested that the 13 nucleotides between −68 and −56 (GGGCGGGGCCGTG) were important for promoter activity. This segment contains a 9 bp sequence with dyad symmetry (GGC GGG GCC) which seemed a good candidate for a *cis*-regulatory element that could function in either the direct or the reverse orientation. To test this hypothesis, the sequence complementary to nucleotides −68 to −56 was inserted at the 5′ end of plasmid pMS30, which extends to −56, generating plasmid pMS31 (Figure 3d). This resulted in the recovery of promoter activity to a level similar to that of pMS28 (Figure 3e). In order to analyse the sequence requirements of this *cis* element to drive transcription, three base pair substitutions, which destroyed the dyad symmetry, were introduced into the pMS16 backbone generating plasmid pMS34 (Figure 3d). As shown in Figure 3e, pMS34 exhibited a very low level of promoter activity in comparison to that of the parental pMS16 plasmid, suggesting that a *cis* element contained within this small GC-rich *trk*A sequence is required for cell-specific transcription of the *trk*A-CAT reporter fusions. In addition, this *cis*-regulatory element can function in both orientations, thus behaving as an enhancer that could bind *trans* acting nuclear factors. To test this hypothesis, electrophoretic gel mobility shift assays (EMSA) were performed using as a probe a 19-mer double-stranded oligonucleotide whose sequence (5′-GCGGGGCGGGGCCGTGGGT-3′) corresponds to that of the −71 to −53 *trk*A segment, and includes the 13 bp element identified as essential for transcription. As shown in Figure 4a, three DNA-protein complexes, I, II and III, were obtained using nuclear extracts from N2a cells (lane 2). These three complexes were competed specifically by the unlabeled oligonucleotide probe used at 50-fold molar excess (S, lane 3), whereas none of these complexes were competed by a 50-fold molar excess of an unrelated 19-mer oligonucleotide (NS, lane 4), nor by a 19-mer oligonucleotide containing three base pair substitutions that destroyed the inverted repeat sequence motif (M, lane 5). Moreover, no DNA-protein complexes were formed when the mutated oligonucleotide was used as a probe (Figure 4b). These results demonstrated the binding of specific proteins, present in N2a nuclear extracts, to the −71/−53 sequence. Nuclear extracts prepared from PC12 and NCB20, two cell lines which express *trk*A, yielded similar retardation patterns to those obtained with N2a extracts (not shown). In contrast, complexes I, II and III were essentially undetectable in the non-expressing 293 cell line; although a different non-specific complex was observed in these cells (Figure 4a, lanes 6–8; b, lanes 4–5). These data suggest that specific DNA-binding proteins interact with *cis* elements within the −71/−53 region, and that these proteins are probably involved in mediating cell-specific *trk*A expression. Interestingly, the use of excess amounts of non-specific oligonucleotides seems to favor the formation of the three specific complexes in N2a cells, and allows for their detection in 293 cells (Figure 4a, lanes 4, 5 and 8). Analysis of the 150 nucleotides upstream to the transcription initiation site using TESS, revealed the presence of multiple GC-rich

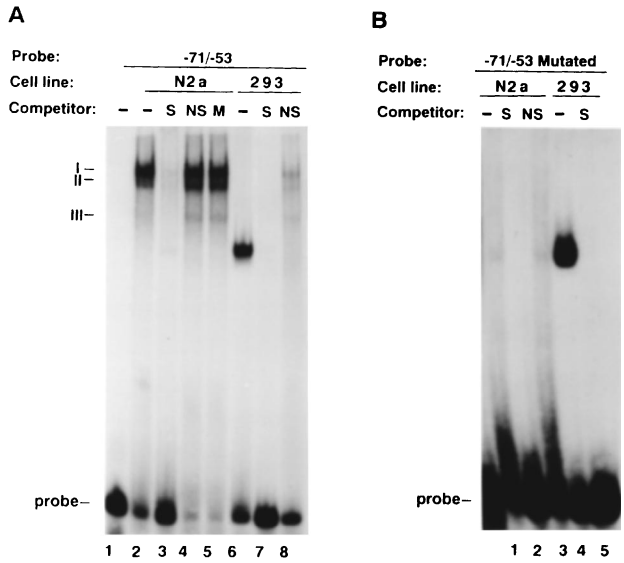

Figure 4 *Trans*-acting factors present in *trk*A-expressing cell lines bind to the −71/−53 *trk*A sequence. Electrophoresis gel mobility shift assay (EMSA) using the double-stranded 19-mer oligonucleotide −71/−53 (**a**) or a −71/−53 mutated oligonucleotide (**b**) as probes and nuclear extracts from *trk*A-expressing N2a cells and *trk*A-negative 293 cells. The three DNA-protein complexes I, II and III detected in N2a cells (lanes 2, 4 and 5 of panel **a**) are indicated. EMSA was performed according to described protocols (Ausubel *et al*., 1995), using 10 μg of nuclear protein extracts, prepared from subconfluent cell cultures, and 0.5–1 ng of ^{32}P-labeled double-stranded oligonucleotides. The sequences of the upper strands of each of the probes used were, −71/−53: 5′-GCGGGGCGGGGCCGTGGGT-3′; mutated −71/−53: 5′-GCGGATCGGGGTCGTGGGT-3′; and non-specific oligonucleotide: 5′-ATATTCTGAGACAGCCTTC-3′. −Indicates absence of competitor oligonucleotide. 50-fold molar excess of unlabeled competitors were added as follows, S: specific oligonucleotide probe; NS: non-specific, i.e. unrelated, oligonucleotide; M: mutated oligonucleotide

consensus sequences for transcription factors; in particular the segment between −93 and −56 contains three potential binding sites for Sp1, two for AP-2, and several for GCF (Figure 5a). Most of these sites are conserved in the promoter region of human *trk*A (Chang *et al*., 1998). The −71/−53 oligonucleotide contains consensus binding sites for Sp1, AP-2 and, with a single mismatch, for Egr-1. For this reason, antibodies specific for Sp1, Egr-1 and AP-2 proteins were used to test whether any of these factors were components of complexes I, II or III. As shown in Figure 5b, in the presence of a specific anti-Sp1 antibody the slower-migrating complex (I) changed its mobility and was further retarded, i.e. was supershifted, indicating that the Sp1 protein participates in this complex. The slight reduction in the intensity of complexes I, II and III observed in the presence of anti-Egr-1 and anti-AP2 antibodies may be due to non-specific interference of the antibodies with complex formation; however, neither anti-AP2 nor anti-Egr-1 antibodies supershifted any of these complexes, suggesting that neither factor is present in them.

The function of the NGF receptor, *trk*A, is controlled at different levels, an important one being the transcriptional regulation during development and in the adult life. Therefore, alterations of *trk*A transcription

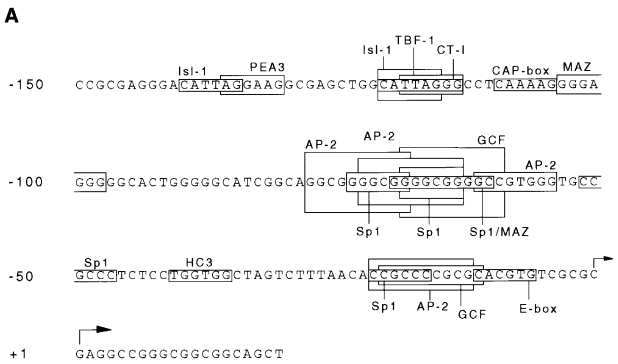

Figure 5 The Sp1 protein is present in complex I. (**a**) Location of potential transcription factor binding sites in the promoter region of mouse *trk*A, identified by TESS: Transcription Element Search Software on the WWW (http://agave.humgen.upenn.edu/tess/index.html). Nucleotides are numbered with respect to the main *trk*A transcription start site, indicated by the large bent arrow. (**b**) Gel mobility shift assays performed using N2a nuclear extracts and the −71/−53 double-stranded oligonucleotide as probe, in the presence or absence of 2 μg of anti-Sp1, anti-AP2 or anti-Egr1 antibodies (Santa Cruz Biotechnology, Santa Cruz, CA, USA). Antibodies were added to the reaction mixture and incubated for 1–2 h on ice before adding the labeled probe. I, II and III indicate DNA-protein complexes; SS indicates a supershifted complex formed in the presence of anti-Sp1 antibody

may underlie disorders such as neurodegeneration and tumor progression. In this report we have described the structure of the mouse *trk*A promoter, and the identification of *cis*-regulatory sequences that are essential for cell type-specific transcription of *trk*A-reporter gene fusions, and are capable of binding nuclear factors present in *trk*A-expressing cells. We have located the transcription start site just 70/71 nucleotides upstream to the AUG translation initiation codon, in a very GC-rich region, devoid of TATA, CAAT or initiator sequences. In spite of this, the start of transcription is unique, and lies 20 nucleotides downstream to a potential Sp1 binding site. In contrast, the human *trk*A gene seems to initiate transcription at several sites, all close to the AUG codon. The reason for this difference is not clear; however, it is possible that different transcription start sites are used in different cell types; thus, we have defined the mouse *trk*A start site in RNA from primary embryonic DRGs, while that of the human gene was determined in transformed cell lines (Chang *et al*., 1998).

A CpG island of unmethylated DNA encompasses the 5′ region of the mouse *trk*A gene, raising the possibility that methylation plays a role in regulating *trk*A expression. Methylation of the *trk*A island may suppress *trk*A expression, which in turn could result in the loss of NGF-dependent neuronal populations, leading to neurodegeneration. In fact, lack of *trk*A expression has been observed in specific neuronal populations of Alzheimer's disease patients (Bossiere *et al*., 1997), although the cause of this deficit is unknown. In the other hand, loss of *trk*A expression may be involved in neuroblastoma progression (Brodeur *et al*., 1997). Analysis of the 5′ region of mouse *trk*A revealed the close presence of the IRR gene, encoding a receptor tyrosine kinase, which is transcribed in the opposite direction, from the complementary DNA strand. The putative start of transcription of the mouse IRR gene is located just 1.6 kbp away from that of the *trk*A gene. It is unlikely that *trk*A and IRR are driven by a bi-directional promoter; however, they are co-expressed at the cellular level in both neuronal and non-neuronal tissues in rat (Reinhardt *et al*., 1994; Tsuji *et al*., 1996), suggesting that they may share regulatory elements acting both in *cis* and in *trans*. Similar genomic arrangements may exist in other species; thus, both the IRR and *trk*A genes are located in human chromosome 1 (Shier *et al*., 1990; Valent *et al*., 1997). We have shown that a small region of 150 nucleotides, 5′ to the mouse *trk*A gene, is required for promoter activity and for cell-type specific transcription. Detailed analyses of this region allowed us to identify a 13 bp *cis*-regulatory element that behaves as a tissue-specific enhancer. This element contains a 9 bp sequence with dyad symmetry composed of a three nucleotide inverted repeat separated by a stretch of three G residues (−67: GGC GGG GCC). Disruption of this sequence or mutations that destroyed the dyad symmetry drastically reduced transcription activity to levels similar to those of inactive fusions. A similar motif is conserved at an equivalent position in the human *trk*A promoter, although in this case there are four G residues separating the three-nucleotide inverted repeat (Chang *et al*., 1998). This mouse *trk*A *cis*-regulatory element was capable of binding nuclear

proteins present in *trk*A-expressing cell lines, giving rise to three specific DNA/protein complexes, which resembled those observed in human neuroblastoma cell lines, using a longer human DNA probe (Chang *et al*., 1998). These complexes were undetectable in *trk*A non-expressing cell lines, suggesting that the nuclear factors involved in their formation may contribute to the tissue-specific expression of *trk*A. However, these factors may be present, at low levels, in 293 cells being only apparent in the presence of large concentrations of non-specific competitors. An oligonucleotide containing a triple mutation that destroyed the dyad symmetry did not bind nuclear factors nor did it compete with the wild type oligonucleotide for complex formation, suggesting that the inverted repeat element is required both for transcription in transfection experiments in N2a cells, and for DNA-protein complex formation *in vitro*. This DNA segment contains potential binding sites for several factors, including Sp1, MAZ, GCF and AP-2. Sp1 and MAZ are general transcription activators that have been proposed as important regulators of TATA-less promoters (Parks and Shenk, 1996), whereas GCF acts as a repressor for GC-rich promoters (Kageyama and Pastan, 1989); in the other hand, AP-2 is a specific transcriptional activator present in neural crest-derived lineages (Mitchell *et al*., 1991). Thus, both positive and negative transcription factors as well as general and specific ones, can potentially interact with the *cis*-regulatory element at the *trk*A promoter in order to regulate its expression (Figure 5a). An anti-Sp1 antibody specifically recognized a component of complex I as evidenced by the slower migration of this complex in supershift experiments; in contrast, AP-2 and Egr-1 antibodies failed to supershift any of those three DNA-protein complexes. Sp1 by itself may not account for cell-type specific transcription but it may cooperate with other factors to regulate the *trk*A promoter. Different sets of factors are probably involved in the formation of complexes II and III and it is likely that precise regulation of *trk*A expression is dependent on the interaction of multiple transcription factors both at the *cis* element described in this report, as well as in other regions of the gene. Further studies are required to identify them.

Acknowledgements
We thank Dr F Antequera for valuable discussions and A Duwell for technical help. MP Sacristán and M Bonilla were recipients of postdoctoral and predoctoral fellowships from the Spanish Ministry for Education, respectively. This work was supported by grants from the Spanish Ministry for Education, DGICYT refs. PM91-0005 and PB94-1104, and by European Union Programs BMH2-CT96-0010 and BIO4-CT96-0285.

References

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA and Struhl K. (1995). *Curr. Proto. in Mol. Biol*. John Wiley and Sons, New York.
Barbacid M. (1995). *Curr. Opin. Cell Biol.,* **7,** 148–155.
Bird AP. (1986). *Nature,* **321,** 209–213.
Boissiere F, Faucheux B, Ruberg M, Agid Y and Hirsch EC. (1997). *Exp. Neurol.,* **145,** 245–252.
Brodeur GM, Nakagawara A, Yamashiro DJ, Ikegaki N, Liu XG, Azar CG, Lee CP and Evans AE. (1997). *J. Neurooncol.,* **31,** 49–55.
Clark SJ, Harrison J, Paul CL and Frommer M. (1994). *Nuc. Acids Res.,* **22,** 2990–2997.
Chang BB, Persengiev SP, de Diego JG, Sacristan MP, Martin-Zanca D and Kilpatrick DL. (1998). *J. Biol. Chem.,* **273,** 39–44.
Dalal R and Dajakiew D. (1997). *Mol. Cell Endocrinol.,* **134,** 15–22.
Davies AM. (1994). *J. Neurobiol.,* **25,** 1334–1348.
Ernfors P, Merlio JP and Persson H. Eur. (1992). *J. Neurosci.,* **4,** 1140–1158.
Fagan AM, Zhang H, Landis S, Smeyne RJ, Silos-Santiago I and Barbacid M. (1996). *J. Neurosci.,* **16,** 6208–6218.
Greco A, Miranda C, Pagliardini S, Fusetti L, Bongarzone I and Pierotti MA. (1997). *Genes Chrom. Cancer,* **19,** 112–123.
Holtzman DM, Kilbridge J, Li Y, Cunningham Jr ET, Lenn NJ, Clary DO, Reichardt LF and Mobley WC. (1995). *J. Neurosci.,* **15,** 1567–1576.
Indo Y, Tsuruta M, Hayashida Y, Karim MA, Ohta K, Kawano T, Mitsubuchi H, Tonoki H, Awaya Y and Matsuda Y. (1996). *Nat. Genet.,* **13,** 485–488.
Kageyama R and Pastan I. (1989). *Cell,* **59,** 815–825.
Lomen-Hoerth C and Shooter EM. (1995). *Neurochem.,* **64,** 1780–1789.
Martin-Zanca D, Oskam R, Mitra G, Copeland T and Barbacid M. (1989). *Mol. Cell. Biol.,* **9,** 24–33.
Martin-Zanca D, Hughes S and Barbacid M. (1986). *Nature,* **319,** 743–748.
Martin-Zanca D, Barbacid M and Parada LF. (1990). *Genes Dev.,* **4,** 683–694.
Mayerhofer A, Dissen GA, Parrott JA, Hill DF, Mayerhofer D, Garfield RE, Costa ME, Skinner MK and Ojeda SR. (1996). *Endocrinology,* **137,** 5662–5670.
Mitchell PJ, Timmons PM, Hébert JM, Rigby WJ and Tjian R. (1991). *Genes Dev.,* **5,** 105–119.
Miralles F, Philippe P, Czernichow P and Scharfmann R. (1998). *J. Endocrinol.,* **156,** 431–439.
Parks CL and Shenk T. (1996). *J. Biol. Chem.,* **271,** 4417–4430.
Reinhardt RR, Chin E, Zhang B, Roth RA and Bondy CA. (1994). *J. Neurosci.,* **14,** 4674–4683.
Sambrook J, Fritsch EF and Maniatis T. (1989). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
Shier P and Watt VM. (1992). *Mol. Endocrinol.,* **6,** 723–729.
Shier P, Willard HF and Watt VM. (1990). *Cytogenet. Cell Genet.,* **54,** 80–81.
Tam SY, Tsai M, Yamaguchi M, Yano K, Butterfield JH and Galli SJ. (1997). *Blood,* **90,** 1807–1820.
Torcia M, Bracci-Laudiero L, Lucibello M, Nencioni L, Labardi D, Rubartelli A, Cozzolino F, Aloe L and Garaci E. (1996). *Cell,* **85,** 345–356.
Tsuji N, Tsujimoto K, Takada N, Ozaki K, Ohta M and Itoh N. (1996). *Brain Res. Mol. Brain Res.,* **41,** 250–258.
Valent A, Danglot G and Bernheim A. (1997). *Eur. J. Hum. Genet.,* **5,** 102–104.
Wheeler EF, Gong H, Grimes R, Benoit D and Vazquez L. (1998). *J. Comp. Neurol.,* **391,** 407–428.
Yaar M, Eller MS, DiBenedetto P, Reenstra WR, Zhai S, McQuaid T, Archambault M and Gilchrest BA. (1994). *J. Clin. Invest.,* **94,** 1550–1562.