Cell-specific expression of transfected brain identifier repetitive DNAs

Synthia H.Mellon⁺, John D.Baxter and Arthur Gutierrez-Hartmann^{§*}

Metabolic Research Unit and the Departments of Medicine and of Biochemistry and Biophysics,
University of California, San Francisco, CA 94143, USA

Received December 9, 1987; Revised and Accepted April 8, 1988

ABSTRACT

To define the neural-specific expression of rat repetitive identifier (ID) DNA, we co-transfected an intron B subclone of the rat growth hormone (rGH) gene, containing a tandem array of two type 2 repeats and a single ID monomer, and a plasmid conferring neomycin resistance into human SK-N-MC neuroblastoma, HeLa epidermal carcinoma, 293 kidney and 251 MG glioblastoma cells. Transcript analysis from both individual and pools of G418-resistant cells revealed that rGH intron B repeats were expressed only in SK-N-MC neuroblastoma cells as small, cytoplasmic RNAs of 85, 110, 155 and 180 bases. Primer-extension studies show these repetitive RNAs to contain a common 5' end that maps precisely to the beginning of the ID element and that type 2 transcripts are not stably expressed. However, ID DNA expression from two other transfected plasmids, each containing only the ID core sequence, was not restricted to the SK-N-MC cell line. These data show that the transfected rGH ID sequence is selectively expressed in a neural-specific manner resulting in BC-like RNAs, and furthermore, suggest that flanking DNA may play a role in cell-specific expression of certain repetitive DNA elements.

INTRODUCTION

Short length, middle-repetitive DNA consists of about 3×10^5 related sequences which are dispersed throughout the mammalian genome (1-3). Individual elements can be grouped into families based on sequence homology and length (4,5). Two of the most abundant rodent repetitive DNA families include the type 2 repeat and a related, but truncated, version containing 66% homology to the first 73 bp of the type 2 DNA, which we have termed the truncated repeat (TR) (4-9). The type 2 and TR sequences may be evolutionarily derived from tRNA genes, and many individual elements are transcribed *in vitro* by RNA polymerase III (pol III) (8-13). Additionally, flanking sequences, promoter organization and factor competition may all influence repetitive DNA expression *in vitro* (8,10).

Mechanisms which regulate repetitive DNA expression *in vivo*, however, remain essentially unknown. Type 2 transcripts are typically found in heterogenous nuclear (hn) RNA and are present in only small amounts as stable, cytoplasmic products (1,5,6). By contrast, TR transcripts are expressed in a tissue-specific and developmentally regulated manner in neural tissues as small, stable cytoplasmic RNAs (14-19; for review see ref. 20). Because these RNAs were initially identified in brain cytoplasm, they were termed BC RNAs, and the DNA

sequence to which they hybridized was termed brain identifier (ID) element (14). Although there are about 10^5 copies of TR DNA, recent data indicate that BC1 RNA is transcribed from only one or a few of these elements (18). We use the term "ID" to denote the small subset of TR sequences that are brain-specific.

BC1 RNA is a tripartite structure of about 160 bases containing the first 75 bp of the ID sequence as the 5' domain, an A-rich tract of heterogeneous length as the middle domain, and a unique 23 bp sequence as the 3' BC1-specific domain (18). BC1 RNA expression is developmentally regulated and restricted to neural tissue and the anterior pituitary (15-17). BC2 is about 100-110 bases long, abundant in brain but is also present in very low amounts in peripheral tissues (15,17). BC3 is about 75 bases long, probably represents the ID "core" sequence, and is also found in abundant quantities in the testes (15-17). Brain-specific transcripts analogous to the rat BC RNAs have been found in the mouse, hamster and monkey (17,19,20). However, in many cultured rodent cell lines, BC RNA expression is not restricted in a neural-specific manner (17,20). Also, it is not apparent in human cell lines, even at low hybridization stringency (17).

In the current study, we co-transfected an intron B subclone of the rat growth hormone (rGH) gene containing two type 2 repeats and a single TR arranged in tandem (8), or alternatively, individual ID clones, together with a plasmid conferring neomycin resistance into a variety of human neural and non-neural cell lines in order to further define the cell-specific expression of TR/ID DNA. By using human cell lines, we were able to analyze the expression of transfected rodent repeats, since these sequences do not cross-hybridize with human repetitive DNA sequences (1,2). The rGH intron B repeats were chosen for this study since: 1) their expression *in vitro* by HeLa cell extracts and in *Xenopus* oocytes has been characterized (Fig. 1) (8,9); 2) the type 2 repeats serve as an internal control; and 3) the rGH TR, but not the ID subclones, contain 3' flanking DNA which may include a unique domain conferring cell-specificity. Our data show that the rGH TR element is selectively expressed in transfected SK-N-MC neuroblastoma cells, but not in HeLa epidermal carcinoma, 293 kidney or 251 MG glioblastoma cells, resulting in significant quantities of stable, cytoplasmic BC-like RNAs. Primer extension studies mapped the 5' end of these RNAs precisely to the beginning of the rGH TR element and failed to demonstrate any transcripts containing type 2 sequences. By contrast, expression of two other ID subclones (p2A120 and p1B224), containing only the ID core repeat and a short A tract (14), was detected in transfected HeLa, 293 and 251 MG cells. These data indicate that certain members of the TR family can function as neural-specific identifier elements, and that flanking DNA structures possibly influence this behavior.

MATERIALS AND METHODS

Plasmid DNAs. The plasmid prGH-1250 (7,8) contains a 1250 bp PstI fragment of the Sprague-Dawley rGH gene, including the last five bases of exon 1 through the middle of intron

C, inserted into the PstI site of pBR322 (Figs. 1a and b). Plasmids p2A120 and p1B224, each containing a "core" ID sequence of 75 bp followed by about 7 bp of oligo A, and preceded by intron-like sequences, were kindly provided by Dr. G. Sutcliffe (14). The plasmid *pneodX* used in these studies was derived from the vector *pK_{Neo}* (kindly provided by Dr. D. Hanahan, Cold Spring Harbor), which contains the neomycin resistance gene and SV40 mRNA processing signals downstream from the SV40 enhancer and early promoter, the BamHI-EcoRI fragment of pBR322, and the lac UV5 promoter. The plasmid *pneodX* was obtained by deleting the BamHI-NdeI fragment of pBR322 and adding XhoI linkers at the EcoRI site; this shortens the vector length but maintains the origin of replication and ampicillin resistance sequences of pBR322 (21). Plasmid DNAs were prepared by polyethylene glycol precipitation of overnight bacterial cultures followed by banding in CsCl gradients (22,23).

Cell culture and DNA transfection. HeLa, SK-N-MC (24), 251 MG (25) and 293 (26) cell lines were obtained from the Tissue Culture Facility at the University of California, San Francisco. Cells were grown in Dulbecco's modified Eagle's medium (DMEM), 10% fetal bovine serum, 50 U/ml penicillin and 50 µg/ml streptomycin. Individual cell lines (1×10^6 cells per 10cm plate) were co-transfected with 1 µg of *pneodX* and 10 µg of either prGH-1250, p2A120 or p1B224 DNA by the calcium phosphate procedure (27). Cells were selected in medium containing 1 mg/ml G418 (Geneticin, Gibco) and pools of G418-resistant colonies were grown to mass culture. To obtain individual G418-resistant colonies, isolated clones were overlaid with a 3 mm paper disc (Whatman 3MM) impregnated with 0.2% trypsin. After 3 min., each 3 mm disc was transferred to a single well of a 24-well tissue culture plate and the individual colonies were grown to mass culture as described above.

RNA isolation and analysis. Nuclear and cytoplasmic RNA was isolated from control and transfected cells as described (28,29). RNA was either electrophoresed through 1.5% agarose gels containing formaldehyde or glyoxylated and electrophoresed through nondenaturing gels and then transferred to nitrocellulose paper (30,31). RNA size was estimated on 5% polyacrylamide-8M urea gels and electrophoretic transfer to nylon membranes (ZetaProbe, BioRad) at 30V overnight in 5mM Tris-acetate, pH 7.8, 2.5 mM sodium acetate, and 0.25 mM EDTA. Blots were then hybridized to the 1250 bp PstI insert of prGH-1250, which had been purified by gel electrophoresis and labelled by nick-translation to 10^8 cpm/µg. Primer extension studies were performed essentially as described by Singh et al. (32), using either a synthetic 21-base oligonucleotide specific for transcripts containing the rGH TR/ID sequence (see Fig.4) or a 16-base oligonucleotide specific for transcripts containing the rGH type 2 sequence. These oligonucleotides (kindly provided by Dr. J.A. Iwasa, California Biotechnology, Inc.) were end-labelled with [γ - 32 P]ATP (>7000 Ci/mmol, ICN) and T4 polynucleotide kinase (P-L Biochemicals). About 5×10^4 cpm of either the [32 P]-labelled 21-mer or 16-mer and 50 µg of cytoplasmic RNA were boiled for 3 minutes and then annealed for 3 hours at 50°C in 300mM NaCl, 10mM Tris-Cl pH 7.8, and 1mM EDTA. Following the hybridization, samples

were incubated with 100 U/ml of avian myeloblastosis virus reverse transcriptase (Life Sciences) for 1 hour at 42°C in a buffer adjusted to 60mM NaCl, 10mM Tris-Cl pH 7.8, 10mM dithiothreitol, 1mM dNTPs, 8mM MgCl₂, 50 mg/ml actinomycin D and 1000 U/ml RNasin (Promega Biotech). The reaction was stopped by phenol/chloroform extraction and ethanol precipitation. Extended products were dissolved in 80% formamide dye mix, heated at 90°C for 5 minutes and then loaded onto 8% polyacrylamide - 8M urea sequencing gels. Primer extended products were visualized by autoradiography at -70°C using an intensifying screen.

Figure 1: Structural organization and RNA products of the repetitive DNA sequences associated with the rGH gene. (a) Schematic of the 5.8 kb EcoRI-HindIII genomic fragment containing the Sprague-Dawley rGH gene (prGHeh-5.8) (7). Repetitive DNA elements (arrows, labeled A-J) are oriented in the direction of their transcription. Exons are designated E₁-E₅ and introns I_A-I_D. The PstI sites which define the 1250 bp fragment used in constructing the intron B subclone prGH-1250 (8) are shown. Wavy lines represent flanking plasmid DNA. (b) Enlargement of the rGH gene structures present in the prGH-1250 subclone. The symbols are as in (a). The 18 bp direct repeats that flank the three tandem, middle-repetitive DNA elements within intron B are shown as small, dark arrowheads. (c) The HeLa cell-free pol III transcription map of rGH-intron B. The RNA products of prGH-1250 plasmid DNA transcribed in HeLa cell-free extracts are aligned below the rGH region from which they are expressed. The RNAs initiating with the first type 2 repeat are labeled A1-3 and those initiating with the second type 2 repeat are labeled B1-2. The transcript lengths in nucleotides are also shown. Two initiation sites for the A1 and A2 transcripts result in the two lengths indicated (see ref 8). (d) The transcription map of intron B DNA obtained in *Xenopus* oocytes. The single, 73-base RNA is expressed in *Xenopus* oocytes using either prGHeh-5.8 or prGH-1250 DNA as a template or using injected HeLa cell-free transcripts; it contains sequences matching the first 73-76 bases of the TR element in intron B (see ref. 9). A dashed line at the ends of the arrow representing the TR transcript in "d" indicates that neither the 5' nor 3' ends of this transcript were mapped precisely. The TR transcript is aligned below the cell-free transcripts A1 and B2 from which it may have been processed.

RESULTS

Transcription map of rGH intron B repetitive DNA derived from cell-free and *Xenopus* oocyte transcription studies. Previous HeLa cell-free transcription studies revealed that the rGH-intron B type 2 structures are efficiently transcribed by pol III and result in two sets of overlapping primary transcripts of 175 to 730 nucleotides which map to pol III initiation and termination signals and utilize tRNA gene transcription factors (Fig. 1) (8). However, the rGH-TR elements exhibited little, if any, transcription activity *in vitro* and were found only as parts of larger RNAs directed by upstream type 2 elements. In *Xenopus* oocytes, however, transcripts derived from type 2 repeats do not accumulate as stable species (9). Instead, a single, 73-base RNA containing the TR/ID "core" structure and coinciding with the tRNA homology within the TR region accumulates preferentially by posttranscriptional processing of larger, pol III repeat transcripts (Fig.1d) (9).

Figure 2: Northern blot analysis of rodent repetitive DNA transcripts expressed in transfected human SK-N-MC neuroblastoma and HeLa cells. Cytoplasmic and nuclear RNA was prepared from pools of G418-resistant colonies transfected and control SK-N-MC and HeLa cells as described by Savouret et al. (28), and cytoplasmic RNA was prepared from the GC rat pituitary tumor cell line (lane labeled GH) as described by White and Bancroft (29). The GH lane contains 2.5 μ g of RNA, all other lanes contain 20 μ g of RNA. The RNAs were electrophoresed on 1.5% agarose-2.2M formaldehyde gels, transferred to nitrocellulose and hybridized as detailed in Methods. The location of the 28S RNA (about 5kb), 18S RNA (about 2kb), rGH mRNA (about 1 kb), and the bromophenol blue dye (BPB; about 0.11 kb) are included as size references.

Analysis of RNAs expressed from rGH-intron B repetitive DNA in transfected SK-N-MC and HeLa cells. Since the size of rGH intron B repetitive DNA transcripts can vary from 75 to 730 bases (Fig. 1), RNA from transfected cells was initially analyzed on 1.5% agarose-2.2M formaldehyde gels under conditions to retain small (75 bases) transcripts. Figure 2 shows a Northern blot of cytoplasmic and nuclear RNA isolated from pools of transfected SK-N-MC and HeLa cells resistant to G418 ("Pst 1250" lanes), and from control cells ("control" lanes), probed with the ^{32}P -labeled 1250 bp PstI fragment of the rGH gene. Abundant cytoplasmic RNA hybridized to this probe in the transfected SK-N-MC cells, but was not detected in the transfected HeLa cells. However, cytoplasmic RNA from untransfected SK-N-MC and HeLa cells, or tRNA alone, did not hybridize with this probe. The complete absence of stable rGH-intron B transcripts in transfected HeLa cells was particularly surprising, since whole cell extracts prepared from these cells actively transcribe the rGH-intron B type 2 repeats (Fig. 1) (8). The lane marked "GH" in Fig. 2 contains cytoplasmic RNA prepared from the GC rat pituitary tumor cell line; mature rGH mRNA is detected (noted on the left of the figure) since the 1250 bp probe contains exons 2 and 3 of the rGH gene (7,8). The GH lane also displays the smear pattern typically observed when rodent RNA is probed with rodent middle-repetitive DNA (4,14-17,19). This smear of radioactivity, which is most evident above the 18S marker, is presumed to represent leakage of nuclear RNA polymerase II transcripts containing interspersed, middle-repetitive DNA sequences (1,19).

To determine directly whether any rGH-intron B transcripts occur in large hnRNA structures and if the differential expression of intron B DNA is maintained at the nuclear level, nuclear RNA from control and transfected HeLa and SK-N-MC cells was prepared and analyzed. As shown in lanes marked "Pst 1250" in Fig. 2, transcripts slightly larger than the cytoplasmic repeat RNAs were found in nuclear RNA from transfected SK-N-MC cells and, to a much lesser degree, in HeLa cells. Large hnRNAs (>18S) containing intron B sequences were not found in either HeLa or SK-N-MC cells, whereas the intron B repeat transcripts tend to be restricted to the nucleus in transfected HeLa cells.

To determine the size of this RNA more precisely, we electrophoresed it on a 5% polyacrylamide-8M urea gel, electrophoretically transferred it to a nylon membrane, and probed it with the ^{32}P -labeled, 1250 bp Pst I fragment (Fig. 3). Three separate clonal lines (SK #1-2, SK #2-1, and SK #2-2) of G418 resistant SK-N-MC cells from two separate transfections with prGH-1250 DNA expressed RNAs of 85 and 110 nt. A fourth clonal line (SK #2-4) expressed additional RNAs of 180 and 155 nt (Fig. 3). The major RNA species observed in pools of transfected SK-N-MC cells grown to mass culture are 180 and 155 nt long. RNA from control SK-N-MC cells is devoid of sequences homologous to the rGH-intron B DNA. The minor signal observed at about 110 nt is probably a blotting artifact, since control lanes in several other studies do not contain this signal. The GC line of cultured rat pituitary tumor cells not only contains RNAs of 180 and 155 nt, but also expresses RNAs of 120 and 100 nt. As a

Figure 3: Electrophoretic size determination of rat repetitive RNA. Cytoplasmic RNA from individual clones of transfected SK-N-MC (SK#1-2, #2-1, #2-2 and #2-4) cells, pools of clones from transfected SK-N-MC cells, untransfected SK-N-MC cells, GC cells, and rat brain was prepared as described (29), denatured in 90% formamide sample buffer, and electrophoresed on a 5% polyacrylamide-8M urea gel. The RNAs were electrophoretically transferred onto a nylon membrane and probed with nick-translated PstI 1250 DNA, as described in Methods. The rat brain and GC pituitary RNA lanes contained 2.5 μ g of cytoplasmic RNA, all other lanes contain 20 μ g of RNA. The marker (M) lane contains 32 P-labeled, HpaII digested-pBR322 size standards. The approximate sizes of the various repeat transcripts are indicated.

comparison, the putative brain-specific transcripts of 155 nt (BC1) and 100 nt (BC2) found in rat brain RNA are also shown (Fig. 3) (17). The BC3 RNA of about 75 nt is barely visible in the brain tract in Fig. 3. Of particular note, a transcript equivalent in size to the BC1 RNA (155 nt) is also found in the SK #2-4, SK pool and cultured GC cells (Fig. 3). Additionally, the RNA from rat brain clone #2-4, GC pituitary cells and, to a much lesser extent, from the pool results in a hybridization signal which migrates as a smear at the top of the gel. This smear is typically observed in Northern blots probed with repetitive DNA and has been ascribed to mRNAs which contain repeat sequences in their 3' untranslated regions and to nuclear leak of hnRNAs (1,4,14-19). The lack of such a signal in the individual clones #1-2, #2-1 and #2-2 suggests that prGH-1250 DNA did not integrate downstream from a pol II promoter in these clones. Whereas the intense smear in clone #2-4 indicates that in this case prGH-1250 DNA may have integrated into an active pol II transcription unit. Finally, none of the rGH-intron B transcripts expressed in transfected human neuroblastoma cells are of a size similar to either the

Figure 4: Primer-extension analysis of rat repetitive DNA transcripts. (a) Structure of prGH-1250, as previously described in Fig. 1 b. (b) Location of the 21 base TR (or ID)-specific primer. (c) The sizes of the expected extended products and the corresponding 5' ends to which

they would map are indicated. (d) 8% sequencing gel of extended products specifically primed by the 21-base TR oligonucleotide. Primer extension was performed as described in Methods. The sources of cytoplasmic RNA are from pools of G418-resistant cells (50 μ g HeLa and SK-N-MC) brain tissue and GC pituitary cells (2.5 μ g brain and GC cell RNA) as indicated. The marker (M) lane displays DNA size standards. The size of the TR-extended products of 53 nt is indicated.

rGH-intron B transcripts generated *in vitro* in HeLa cell extracts or *Xenopus* oocytes (Figs. 1 and 3) (8,9).

Primer extension analysis of rGH-intron B RNAs expressed in transfected SK-N-MC cells. In Figures 2 and 3, it was not possible to distinguish if the intron B RNAs we identified in transfected SK-N-MC cells were derived from the type 2 and/or TR repetitive DNA because the probe used contains the entire intron B region of the rGH gene. Furthermore, we wished to determine the 5' end(s) for these transcripts. If pol III was responsible for these transcripts, then these RNAs should have a similar 5' end that should be located about 15 nt upstream from the A box of the pol III promoter (8,10,32,33). To address these points specifically, primer extension studies were performed using synthetic oligonucleotides specific for either the TR/ID or for the type 2 repetitive DNA sequence. The 21-base, TR/ID primer (5'-GGCCTTGCGCTTCCTAGGTAA-3') corresponds to the mid-portion of the rGH-TR (or ID) sequence (nucleotides 995-1015 in ref. 7). It contains 17 mismatches with the type 2 sequence in this region (see below) and, therefore, is specific for TR/ID transcripts (Fig. 4). The 5' end of the TR/ID primer is 53 nt downstream from the beginning of the TR/ID sequence, 226 nt downstream from the second type 2 sequence, and 422 nt downstream from the beginning of the first type 2 element (Fig. 4c). As shown in Fig. 4d, cytoplasmic RNA from control and transfected HeLa cells, control SK-N-MC cells and a mock tRNA reaction did not result in any products specifically extended by the TR/ID primer with reverse transcriptase. The DNA extension products migrating at about 42 bases are common to all the lanes, including the tRNA control, and are, therefore, not specific for rGH-intron B transcripts. However, a major extension product of 53 nt results when RNA from transfected SK-N-MC cells is primed with the TR/ID oligonucleotide (Fig. 4d). This indicates that the 5' end of at least some of the rGH-intron B transcripts corresponds precisely to the beginning of the TR/ID repeat. Furthermore, annealing the TR/ID primer with rat brain or GC cell cytoplasmic RNA results in an extended product of a similar size (Fig. 4d).

These data demonstrate that rat brain, cultured rat pituitary tumor cells and transfected human neuroblastoma cells all express small cytoplasmic RNAs that are homologous to the brain identifier sequence and that share a common 5' end. Although the 5' end of these transcripts corresponds precisely to the putative pol III transcription initiation site of the ID sequence, these studies cannot unambiguously exclude the possibility of rapid posttranscriptional processing of transcripts initiated at either of the upstream type 2 sequences resulting in the relative accumulation of an RNA species containing only the TR/ID repeat sequence (19).

Figure 5: Cell-specific expression of different ID elements. HeLa, 293, 251MG and SK-N-MC cells were co-transfected with *pneodX* (1 μ g) and 10 μ g of the plasmid indicated. Pools of G418 resistant colonies were grown to mass culture and cytoplasmic (50 μ g) RNA was primed with the ID primer. Brain RNA was similarly analyzed, and extended products were sized as in Fig. 4. The 53 nt extended products are indicated by arrows.

However, larger extension products mapping initiation events at the type 2 sequences (422 or 226 nt) were not detected, even after prolonged exposures of the gel in Fig. 4D (data not shown).

To determine directly if any cytoplasmic transcripts arose from either of the type 2 repeats but were terminated upstream from the TR element, primer extension studies were performed analogous to those described for the TR primer. In this case, a synthetic 16-base oligonucleotide specific for the type 2 repeat sequence, 5'-TGGAAGAGCAGTCGGG-3, corresponding to nucleotides 604-619 in the first type 2 repeat and 799-814 in the second type 2 repeat as numbered by Barta et al. (7), was used. The type 2-specific primer did not result in any extended DNA products, indicating that cytoplasmic RNA from control and transfected HeLa and SK-N-MC cells, and rat brain, do not contain type 2 sequences (data not shown).

Cell-specific expression of different ID sequences. To determine if the expression of the rGH intron B ID sequence is restricted to SK-N-MC neuroblastoma cells, we co-transfected human 251 MG glioblastoma cells and 293 embryonal kidney cells with prGH-1250 and p α neodX and isolated total RNA from pools of G418-resistant cells, as described previously. Primer extension analysis with the ID primer revealed that the rGH intron B ID sequence is expressed only in SK-N-MC cells (Fig. 5; Pst 1250 lanes). By contrast, two other ID sequences (p2A120 and p1B224) containing variable lengths of intron-like sequences, a 75 bp ID "core" structure in the middle and a short A tract at the 3' end (14), are both expressed in 293 embryonal kidney and 251 MG glioblastoma cells (Fig. 5, lanes p2A120 and p1B224). In HeLa cells, p2A120 ID DNA is expressed to a lesser degree, and expression of p1B224 is inconclusive due to some loss of sample prior to gel loading. Of particular importance is that all ID transcripts, regardless of cell-type in which they were expressed or vector from which they originated, once transcribed they contain the same 5' end. Thus, upstream type 2 sequences are not required to promote ID expression, since p2A120 and p1B224 are expressed. However, the type 2 sequences may somehow influence the cell-specific expression of the rGH ID element.

DISCUSSION

We have compared the expression of several rodent ID elements in different human cell lines and our data show that the expression of the rGH intron B ID element is restricted to SK-N-MC neuroblastoma cells, whereas the expression of other ID elements occurred in HeLa, 251 MG and 293 cells. The type 2 repetitive elements in rGH intron B served as an internal control, and stable cytoplasmic transcripts from this family of repeats were not detected in transfected HeLa or SK-N-MC cells. The cell-specificity of rGH intron B ID expression is apparently highly selective, since its expression was not found in either HeLa epidermal carcinoma cells; 251 MG glioblastoma cells, representing neural supportive cells (25); or in 293 embryonal kidney cells, which express the EIA product of adenovirus (26). The inability of EIA to influence the expression of rGH intron B ID DNA in 293 cells further suggests that rGH

ID expression in SK-N-MC cells is not due to effects of a nonspecific activator of pol III transcription (34).

The differential cell expression of ID elements from prGH-1250, p2A120 and p1B224 may be due to sequence differences. Recent data suggest that BC RNAs are transcribed from only one or a few of the 10^5 TR elements, and that this single TR element represents the neural-specific ID sequence, or BC1 gene(s), and is marked by a unique 23 bp sequence 3' to the A-rich domain (18). It is conceivable that p2A120 and p1B224 have lost neural-specific expression due to a lack of the BC-specific structure, since they were cloned by priming RNA with oligo-dT (14). By contrast, prGH-1250 contains 237 bp of genomic DNA 3' to the A-rich tract of its ID sequence (8). However, significant homology to the 23 bp BC-specific sequence is not found in the 1250 bp Pst I fragment in prGH-1250 (7,8). These data suggest that other sequences present in prGH-1250 are responsible for tissue-specific ID DNA expression. In this respect, the transcriptional regulatory element that promotes neural-specific expression of the JC virus contains significant homology with the ID repeat sequence and only minimal homologies with the 23-bp BC1-specific sequence, and 80% of transgenic mice containing early regions of JC virus developed adrenal neuroblastomas (35,36). Clearly, certain DNA structures, either alone or in combination, are required to result in neural-specific expression. The combination of a 75 bp ID element and a nonrepetitive 23 bp 3' element may be one mechanism, as suggested by DeChiaria and Brosius (18). However, prGH-1250 DNA transfected into human SK-N-MC neuroblastoma cells or JC virus regulatory regions resulting in neuroblastomas in transgenic mice provide model systems that may allow the determination of the precise sequences required. However, whether these sequences function positively to activate neural-specific transcription or function negatively to inhibit ID DNA expression in non-neuronal cells remains unanswered.

The structure of the rGH intron B ID transcripts in transfected SK-N-MC cells greatly mimics the BC RNAs in neural tissues (17,20). Specifically, they share the same 5' end (Figs. 4d and 5), several are of the same size (Fig. 3), and many contain A-rich tracts (data not shown). Furthermore, they both represent stable, cytoplasmic RNAs derived from repetitive DNA; a relatively uncommon finding (1). Although BC RNAs have not been shown to be derived from a single gene, current data indicates that the size differences of BC1 (160 nt), BC2 (110 nt) and BC3 (75 nt) are due to differences in the 3' end (15,18). Since the rGH intron B ID transcripts contain a homogenous 5' end, it is likely that size differences are also due to different 3' ends. The first pol III termination signal occurs 330 nt downstream from the beginning of the rGH ID element (8). Furthermore, there are neither signals for pol II polyadenylation nor weak pol III termination that coincide with the lengths (180, 155, 110 and 85 bases) of the ID RNAs in SK-N-MC cells (7,8). Most BC1 RNAs end with a stretch of thymidines, suggesting that *in vivo* BC1 is an unprocessed, pol III transcript (18). The mechanism whereby alternative 3' ends are generated to result in BC2 and BC3 RNAs remains

unknown. One speculation is that cell-specific, posttranscriptional processing may play a role in cytoplasmic stabilization of discrete classes of repetitive DNA transcripts (9).

Our data suggest that pol III transcribes the rat ID elements in transfected human cells. The common 5' end of the transcripts which maps precisely to the beginning of the ID monomer, whether expressed by p2A120, p1B224 or prGH-1250, is located 15 bases upstream from the pol III box A promoter element, a typical pol III initiation site (8,10,32,33). Additionally prGH-1250 DNA lacks the pol II promoter of the rGH gene (Fig. 1 and ref. 8). The lack of large, heterogeneous nuclear RNAs containing repeated sequences (Fig. 2) also suggests that pol II does not contribute significantly to ID DNA expression in transfected cells. However, the role of transcriptional versus posttranscriptional influences on neural-specific expression of the rGH intron B ID sequence remains unclear.

The biological significance of middle-repetitive DNA and its transcripts remains elusive. This DNA has been implicated in DNA recombination, DNA replication, gene regulation, transposition, or serving no function at all (selfish DNA) (1-3). The recent observations that transcripts from discrete repeated families are expressed and regulated in a tissue-, transformation- and stage-specific manner suggests that at least certain subsets of these repeated elements may have biological significance. The current findings lend support to this hypothesis, and the gene-transfer assay described here should prove very useful in testing directly the precise TR sequence requirements to result in a neural-specific ID element.

ACKNOWLEDGEMENTS

We thank Lynda Freedman for excellent technical assistance, Randy McKinnon, William Wood, Greg Owens and Bill Hahn for helpful discussions, and N.L. Eberhardt and W.L. Miller for encouragement and support. This work was supported by grants from the National Institutes of Health (HD 22013) and the Academic Senate of UCSF (2-506108-0954463) to S.H.M., by grants from the National Institutes of Health (AM 19997) and California Biotechnology, Inc. to J.D.B. and by a New Investigator Research Award (HD 16878) and a UCSF Faculty Development Award to A.G.-H.

Present addresses: ⁺Department of Obstetrics, Gynecology and Reproductive Sciences, Box 0546, University of California, San Francisco, CA 94143 and [§]Departments of Medicine and of Biochemistry, Biophysics and Genetics, B-151, University of Colorado Health Sciences Center, Denver, CO 80262, USA

*To whom correspondence should be addressed

REFERENCES

1. Jelinek, W.R. and Schmid, C.W. (1982) *Ann. Rev. Bioch.* 51, 813-844.
2. Singer, M.F. (1982) *Cell* 28, 433-434.
3. Sharp, P.A. (1983) *Nature* 301, 471-472.
4. Bennett, K.L., Hill, R.E., Pietras, D.F., Woodworth-Gutai, M., Kane-Haas, C., Houston, J.M., Heath, J.K. and Hastie, N.D. (1984) *Mol. Cell. Biol.* 4, 1561-1571.
5. Witney, F.R. and Furano, A.V. (1984) *J. Biol. Chem.* 259, 10481-10492.
6. Haynes, S.R. and Jelinek, W.R. (1981) *Proc. Natl. Acad. Sci. USA* 78, 6130-6134.
7. Barta, A., Richards, R., Baxter, J.D. and Shine, J. (1981) *Proc. Natl. Acad. Sci. USA* 78:4867-4871.

8. Gutierrez-Hartmann, A., Lieberberg, I., Gardner, D., Baxter, J.D. and Cathala, G.G. (1984) *Nucleic Acids Res.* 12, 7153-7173.
9. Gutierrez-Hartmann, A. and Baxter, J.D. (1986) *Proc. Natl. Acad. Sci. USA* 83, 3106-3110.
10. Paolella, G., Lucero, M.A., Murphy, M.H. and Baralle, F.E. (1983) *EMBO J.* 2, 691-696.
11. Daniels, G.R. and Deininger, P.L. (1985) *Nature* 317, 819-822.
12. Lawrence, C., McConnell, D. and Ramsay, W. (1985) *Nucleic Acids Res.* 13, 4239-4252.
13. Sakamoto, K. and Okada, N. (1985) *Nucleic Acids Res.* 13, 7195-7206.
14. Sutcliffe, J.G., Milner, R.J., Bloom, F.E. and Lerner, R.A. (1982) *Proc. Natl. Acad. Sci. USA* 79, 4942-4946.
15. Sutcliffe, J.G., Milner, R.J., Gottesfeld, J.M. and Reynolds, W. (1984) *Science* 225, 1308-1315.
16. Sutcliffe, J.G., Milner, R.J., Gottesfeld, J.M. and Lerner, R.A. (1984) *Nature* 308, 237-241.
17. McKinnon, R.D., Danielson, P., Brow, M.A.D., Bloom, F.E. and Sutcliffe, J.G. (1987) *Mol. Cell. Biol.* 7, 2148-2154.
18. DeChiara, T.M. and Brosius, J. (1987) *Proc. Natl. Acad. Sci. USA* 84, 2624-2628.
19. Sapienza, C. and St-Jacques, B. (1986) *Nature* 319, 418-420.
20. Chikaraishi, D.M. (1986) *Trends in Neurosc.* 9, 543-546.
21. Cattini, P.A., Anderson, T.R., Mellon, P., Baxter, J.D. and Eberhardt, N.L. (1986) *J. Biol. Chem.* 261, 13367-13372.
22. Humphreys, G.O., Willshaw, G.A. and Anderson, E.S. (1975) *Biochem. Biophys. Acta* 383, 457-463.
23. Clewell, D.B. and Helinski, D.R. (1969) *Proc. Natl. Acad. Sci. USA*, 62, 1159-1166.
24. Biedler, J.L., Helson, L. and Spengler, B.A. (1973) *Cancer Res.* 33, 2643-2652.
25. Ponten, J. and MacIntyre, E.H. (1968) *Acta Path. Microbiol. Scandinav.* 74, 465-486.
26. Graham, F.L., Smiley, J., Russell, W.C. and Nairn, R. (1977) 36, 59-72.
27. Graham, F. and Van der Erb, A. (1973) *Virology* 52, 456-467.
28. Savouret, J.-F., Cathala, G., Eberhardt, N.L., Miller, W.L. and Baxter, J.D. (1984) *DNA* 5, 365-376.
29. White, B.A. and Bancroft, F.C. (1982) *J. Biol. Chem.* 257, 8569-8572.
30. Silverman, S.J., Rose, M., Botstein, D. and Fink, G. (1982) *Mol. Cell. Biol.* 2, 1212-1219.
31. Thomas, P. (1980) *Proc. Natl. Acad. Sci. USA* 77, 5201-5205.
32. Singh, K., Carey, M., Saragosti, S. and Botchan, M. (1985) *Nature* 314, 553-556.
33. Hofstetter, H., Kressman, A. and Birnstiel, M.L. (1981) *Cell* 24, 573-585.
34. Gaynor, R.B., Feldman, L.T. and Beck, A.J. (1985) *Science* 230, 477-450.
35. Kenney, S., Natarajan, V., Strike, D., Khoury, G. and Salzman, N.P. (1984) *Science* 226, 1337-1339.
36. Small, J.A., Khoury, G., Jay, G., Howley, P. and Scangos, G.A. (1986) *Proc. Natl. Acad. Sci. USA* 83, 8288-8292.