Vol. 178, No. 1, 1991 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 15, 1991 Pages 291-296

A NOVEL BRAIN-SPECIFIC mRNA ENCODING NUCLEAR PROTEIN (NECDIN) EXPRESSED IN NEURALLY DIFFERENTIATED EMBRYONAL CARCINOMA CELLS

Kei Maruyama, Mihoko Usami, Takako Aizawa, and Kazuaki Yoshikawa*

Department of Molecular Biology
Psychiatric Research Institute of Tokyo
2-1-8, Kamikitazawa, Setagaya-ku
Tokyo 156, Japan

Received May 30, 1991

SUMMARY: A novel DNA sequence has been isolated from a subtraction cDNA library of P19 embryonal carcinoma cells treated with retinoic acid which induces neural differentiation of the stem cells. The cDNA insert (4B) hybridized with a single 1.7 kb mRNA, whose abundance was markedly increased in P19 cells after retinoic acid treatment. The 1.7 kb mRNA was also expressed in the brain, but not in other non-neuronal tissues. A 1.6 kb cDNA insert (4BFL), which was cloned by screening another cDNA library with the 4B probe, encodes a novel protein sequence of 325 amino acids (Mr 36,831). The protein expressed in 4BFL-transfected COS cells was translocated into the nuclei as detected with antibodies against subsequences of the predicted protein. The antibodies stained the nuclei of neurally differentiated P19 cells but not of the undifferentiated stem cells. This novel mRNA encoding the nuclear protein, termed necdin, may represent a useful marker for the differentiation and development of brain cells.

Differentiation and development of neurons and glial cells in the central nervous system (CNS) of vertebrates may be accomplished through complex interactions between series of gene expression and external stimuli (e.g., hormones and growth factors). In the analysis of the molecular mechanism underlying differentiation and development of the CNS cells, the use of in vivo systems may lead to the difficulties associated with their complexities. The P19 embryonal carcinoma (EC) cell line, a clonal line of murine EC cells, differentiates into neurons, astrocytes, fibroblast-like cells (1,2), and microglia-like phagocytes (3) by treating the stem cells with retinoic acid (RA). Since the cytogenesis observed in RA-treated P19 EC cells resembles that of the CNS cells in vivo, P19 EC cells have been used as an in vitro model for the differentiation and development of the CNS cells (2,3,4).

Recently many genes (cDNAs) expressed in mature CNS cells have been cloned and well characterized (e.g., genes encoding neurotransmitter-synthesizing enzymes, neuropeptide precursors, neurotransmitter receptors, and ion-channels). Little is known, however, about genes associated with the differentiation and develop-

*To whom correspondence should be addressed.

Abbreviations: CNS, central nervous system; EC, embryonal carcinoma; RA, retinoic acid; kb, kilobase(s); ABC, avidin-biotin-peroxidase complex.

0006-291X/91 $1.50

ment of the CNS cells. We report here a cDNA sequence encoding novel nuclear protein, which is expressed exclusively in the CNS and the differentiated P19 cells.

MATERIALS AND METHODS

Cell cultures: P19 EC cells were cultured and induced to differentiate as described (2,3): Cultures of P19 stem cells were aggregated in bacterial culture-grade dishes in the absence (RA^-) and presence (RA^+) of 0.5 μM RA (Sigma) for 4 days. When the cells were plated in tissue culture-grade dishes, only the RA^+ cells were differentiated into neurons, astrocytes, and fibroblast-like cells. To enrich the neuron population in the mixed cultures, the RA^+ P19 cells were treated with 5 μg/ml cytosine arabinoside (Sigma) for 4 and 6 days (2).
RNA extraction and Northern blot analysis: Total cellular RNA was extracted from cultured cells and various organs of male ddY mouse by the differential ethanol precipitation method (5) with the modifications (6). Amounts of RNA were quantified by ultraviolet absorption at 260 nm. Total RNA (5 μg per lane) was electrophoresed on a 6 % formaldehyde/1.2 % agarose gel containing 0.5 μg/ml ethidium bromide, and blotted onto a nylon membrane (Hybond-N, Amersham). RNA blots were fixed on the membrane by UV irradiation and hybridized with [^{32}P]-labeled 4B probe. Hybridization and washing were performed under the conditions reported previously (7,8).
Subtraction cloning and differential hybridization: Poly(A)$^+$ RNA was purified with oligo dT latex (Nihon Roche) from total RNA of RA^- and RA^+ P19 EC cell aggregates. A cDNA library of RA^+ mRNA pool was constructed in the lambda ZAP system (Stratagene), and single strand (anti-sense) cDNA in the pBluescript vector was rescued by helper phage infection. The common sequences between RA^+ pool and RA^- pool were removed from the RA^+ cDNA pool by hybridization with biotinylated RA^- mRNA using the Subtractor system (Invitrogen). RA^+-specific single-stranded DNA was used to transform competent Escherichia coli strain DH5α (BRL) to construct the subtraction library. About 200 colonies were randomly selected from the library. Plasmid DNA samples were blotted on a membrane, and screened by differential hybridization with [^{32}P]-labeled cDNA probes synthesized from the RA^- and RA^+ mRNA pools. A plasmid, which contained a 0.6 kilobase (kb) insert (4B), was selected from those showing higher signal intensities by hybridization with RA^+ cDNA.
Sequence analysis: A larger cDNA insert (4BFL) was isolated from a lambda gt10 cDNA library constructed from the RA^+ mRNA pool. EcoRI fragment of the 4BFL insert was subcloned into the EcoRI site of Bluescript II vector (Stratagene) for DNA sequence analysis. DNA sequence was determined in both directions by the dideoxy-chain termination method (9). Sequences were analyzed using the computer program Genetyx (SDC, Tokyo). Sequence homology was analyzed by the Genetyx CD system (SDC) using following databases: EMBL-GDB (release 24), LASL-GDB (release 65), NBRF-PDB (release 26), and SWISS-PROT (release 15).
Cell-free translation: 4BFL cDNA in Bluescript II vector was transcribed by T7 RNA polymerase, and the transcript was capped at its 5' end using a kit (Stratagene). Synthesized RNA was translated in the rabbit reticulocyte lysate system supplemented with [^{35}S]-methionine (Amersham) (7). [^{35}S]-labeled protein was then electrophoresed on SDS-polyacrylamide gel (15% acrylamide).
Expression of 4BFL-encoded protein in COS cells: A 1.6 kb EcoRI fragment of 4BFL cDNA was subcloned into the unique EcoRI site of the expression vector p91023(B) to construct p94BFL. COS-1 cells , donated by the Japanese Cancer Research Resources Bank, were transfected with p94BFL according to the DEAE-dextran method followed by chloroquine-shock treatment (10,11). After 72-hr incubation, p94BFL-transfected COS cells were lysed. Samples were then electrophoresed on SDS-polyacrylamide gels (15% acrylamide) and imunoblotted with antibodies N1 (1:500) and C2 (1:500) by avidin-biotin-peroxidase complex (ABC) technique using a kit (Vector Labs) (12); the antibodies N1 and C2 were raised against synthetic peptides corresponding to the residues 58-72 and 309-325, respectively, of the deduced protein sequence.
Immunocytochemistry: The p94BFL-transfected COS cells were stained with the antibodies N1 (1:500) and C2 (1:500) by detection with fluorescein isothiocyanate-conjugated anti-rabbit immunoglobulins (1:80) (Tago), and were photographed with a fluorescence microscope (Nikon Microphot FXA-FL)(12). Undifferentiated and differentiated P19 cells were stained with the antibody N1 by detection with 3,3'-diaminobenzidine, a substrate of peroxidase, using the ABC technique (12).

RESULTS

From the subtraction cDNA library of RA^+ P19 EC cells, we have isolated 6 cDNA clones showing intense hybridization signals with [^{32}P]-labeled cDNA prepared from RA^+ mRNA. One of these clones had a 0.6 kb insert (4B) which hybridized with a 1.7 kb mRNA species (4B mRNA) in P19 EC cells; the level of 4B mRNA was markedly increased after RA treatment, reached a plateau during days 4-14, and decreased there-

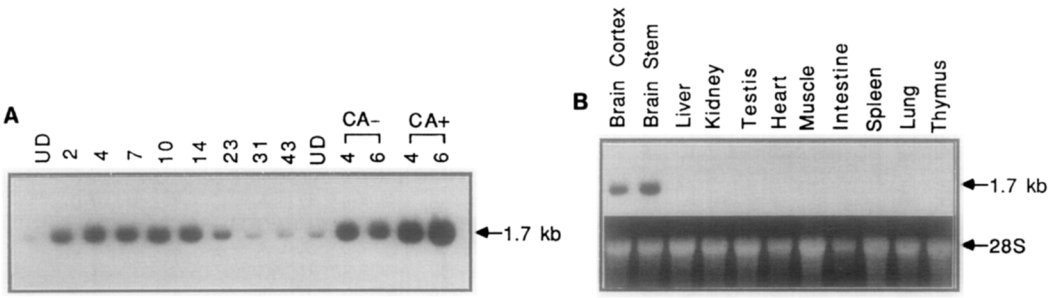

Fig.1.Northern blot analysis of 4B mRNA in neurally differentiated P19 cells and various mouse organs. A: abundances of 4B mRNA (1.7 kb) in neurally differentiated P19 cells. The numbers on top are the durations (days) in culture after RA treatment. UD; undifferentiated control. The differentiated P19 cells were also incubated for 4 and 6 days in the absence (CA-) and the presence (CA+) of cytosine arabinoside. **B:** abundances of 4B mRNA in various mouse organs. Each organ was dissected from a 65-day-old mouse. Fluorescent intensities of 28S ribosomal RNA (28S) reacting with ethidium bromide are shown for assessment of the integrity of RNA blotted on the membrane (lower panel).

after (Fig.1A). Cytosine arabinoside, which enriches the neuron population in mixed P19 cultures (2), increased the abundance of 4B mRNA, indicating that neurons contain a higher level of 4B mRNA than non-neuronal cells. Fig.1B shows the distribution of 4B mRNA in various organs of adult mouse; 4B mRNA was present in the brain (cerebral cortex and brain stem), but not in other non-neuronal tissues examined.

By screening a lambda gt10 cDNA library constructed from RA^+ P19 cells with the 4B probe, we have isolated a clone having a larger 1.6 kb cDNA (4BFL). Both 4B- and 4BFL-cDNA hybridized with the 1.7 kb mRNA in neurally differentiated P19 cells (not shown). This 1.6 kb insert appeared to be nearly full length in comparison with the size (1.7 kb) of 4B mRNA. Determination of total nucleotide sequence of 4BFL revealed that the cDNA contains an open reading frame from nucleotides 1-975 followed by the termination codon TAA; this region encodes a protein of 325 amino acids (calculated Mr 36,831)(Fig.2). Neither the nucleotide sequence of 4BFL nor the deduced amino acid sequence showed appreciable homologies to known sequences in the databases. There are no consensus sequences or motifs of asparagine-linked glycosylation sites, cysteine-rich domain, leucine zipper, or homeodomain. Hydrophobicity profile revealed that this protein sequence is devoid of a signal sequence and a membrane-spanning domain (data not shown).

In order to confirm the assignment of the open reading frame, we have synthesized the protein by cell-free translation of in vitro transcribed 4BFL cDNA (Fig.3A). A major band of Mr 42 kDa [^{35}S]-labeled protein was synthesized in this system. We then tried to characterize the 4BFL-encoded protein by transfecting the cDNA into cultured COS-1 cells (Fig. 3B); two antibodies against distinct parts of the deduced protein sequence reacted with a major 42 kDa band. These results indicate that the coding region of the 4BFL sequence is substantially translated into the protein, although the molecular size (42 kDa) of the synthesized peptides in two systems is higher than that of the predicted protein (37 kDa). Since the protein encoded by 4BFL cDNA

```
                                                          ACCGAGG  -61
GAGTGCCCGCTCCAAGAGCTCCAAGCCGCATCGGTCCTGCTCTGATCCGAAGGCGCAGAC   -1

ATGTCGGAACAAAGTAAGGACCTGAGCGACCCTAACTTTGCAGCCGAGGTCCCCGACTGT   60
M  S  E  Q  S  K  D  L  S  D  P  N  F  A  A  E  V  P  D  C     20
GAGATGCAGGACAGCGATGCCGTTCCGGTGGGGATCCCTCCTCCCGCTTCTCTGGCCGCT  120
E  M  Q  D  S  D  A  V  P  V  G  I  P  P  P  A  S  L  A  A     40
AACCTCGCAGGGCCACCGTGCGCTCCCGAAGGCCCTATGGCAGCCCAACAGGCCTCGCCA  180
N  L  A  G  P  P  C  A  P  E  G  P  M  A  A  Q  Q  A  S  P     60
CCGCCCGAAGAACGGATAGAAGATGTTGACCCTAAAATCCTGCAGCAGGCCGCAGAGGAG  240
P  P  E  E  R  I  E  D  V  D  P  K  I  L  Q  Q  A  A  E  E     80
GGCCGCGCCCACCAGCCCCAGAGTCCAGCCCGGCCGATCCCAGCACCGCCAGCCCCTGCC  300
G  R  A  H  Q  P  Q  S  P  A  R  P  I  P  A  P  P  A  P  A    100
CAGCTGGTGCAGAAGGCGCACGAGCTCATGTGGTACGTGTTGGTGAAGGACCAGAAGAGG  360
Q  L  V  Q  K  A  H  E  L  M  W  Y  V  L  V  K  D  Q  K  R    120
ATGGTCCTCTGGTTTCCAGACATGGTGAAAGAGGTCATGGGCAGCTACAAGAAATGGTGC  420
M  V  L  W  F  P  D  M  V  K  E  V  M  G  S  Y  K  K  W  C    140
AGAAGCATCCTCAGGCGCACCAGCGTCATCCTCGCCAGAGTGTTCGGGCTGCACCTGAGG  480
R  S  I  L  R  R  T  S  V  I  L  A  R  V  F  G  L  H  L  R    160
CTGACCAATCTCCACACCATGGAGTTTGCCCTGGTCAAAGCCCTCAGCCCAGAGGAGCTA  540
L  T  N  L  H  T  M  E  F  A  L  V  K  A  L  S  P  E  E  L    180
GACAGGGTGGCGCTCAACAACCGTATGCCCATGACAGGCCTCCTGCTCATGATCCTGAGC  600
D  R  V  A  L  N  N  R  M  P  M  T  G  L  L  L  M  I  L  S    200
CTCATCTATGTGAAGGGCCGCGGGGCCAGAGAGGGTGCGGTCTGGAATGTGCTGCGCATC  660
L  I  Y  V  K  G  R  G  A  R  E  G  A  V  W  N  V  L  R  I    220
CTGGGGCTGAGGCCCTGGAAGAAGCACTCCACCTTCGGAGACGTGAGGAAGATAATCACC  720
L  G  L  R  P  W  K  K  H  S  T  F  G  D  V  R  K  I  I  T    240
GAGGAGTTCGTCCAGCAGAATTACCTGAAGTACCAGCGTGTGCCCCACATCGAGCCTCCC  780
E  E  F  V  Q  Q  N  Y  L  K  Y  Q  R  V  P  H  I  E  P  P    260
GAGTACGAGTTCTTCTGGGGGTCCAGAGCTAACCGTGAAATCACCAAGATGCAGATCATG  840
E  Y  E  F  F  W  G  S  R  A  N  R  E  I  T  K  M  Q  I  M    280
GAGTTCCTGGCCAGAGTCTTCAAGAAAGATCCCCAGGCGTGGCCTTCCCGATACAGGGAG  900
E  F  L  A  R  V  F  K  K  D  P  Q  A  W  P  S  R  Y  R  E    300
GCTCTGGAGCAGGCCAGAGCTCTGCGGGAGGCTAATCTTGCTGCCCAGGCCCCCCGCAGC  960
A  L  E  Q  A  R  A  L  R  E  A  N  L  A  A  Q  A  P  R  S    320
AGTGTCTCTGAGGACTAAAAAGGTCCAGGGGCACACTGATAGTTTCTGACCCATACTAGG 1020
S  V  S  E  D  *                                              325

GCTGTGTAAGGGTGGGGTTGAGTCATTAGAGTATCCCAAATCCACAGTGCAGTATTTCAT 1080
GTATAATTTTTAAGTTTTCCATACAGTGCTTTTGTACCTTGTAATGCTATTCATTTGTGT 1140
ACTCGTGTAGTGTTTAAGATTGATGCATGTGTGATAAGTATTTGGTACTTTCACTTTTGT 1200
GCTTTCGTGCATTTTTGTACAAGAGATGTGCTGTGCTAAACTTGTGAAATACATTGAGGT 1260
GTTCTGTATCTTGTTCTTTGTATGGGACTGATGATCTGTATCGACAAAGAAGGCCCTGGA 1320
GAGTTAGCAGGACTTAACAGCAACGCAGACCTGAGCAAGAGAAAGGTCAAGGCCTTTCTC 1380
CATATGACTTCAACTGGCACAGGAAGCATCCATGTGGAATGGACTGATTTGAACTGGACT 1440
GTTCTCAGTGTAGGCACTTAGCACCCTTTACAAAACATGTATGCAACCCCACCATAAATA 1500
AACGTTAAAATGAGCATTAAAAAAAAAAAA                                 1530
```

Fig.2. Nucleotide sequence and the deduced amino acid sequence of 4BFL cDNA. Nucleotide residues are numbered in the 5'- to 3'-direction, beginning with the first residue of the ATG triplet encoding the initiating Met. The open reading frame has been tentatively assigned to yield the longest polypeptide sequence (325 amino acid residues). The polyadenylation signal (AATAAA) is underlined.

is unrelated to known proteins, we termed this novel protein necdin for neurally differentiated EC cell-derived factor.

Intracellular localization of necdin in the 4BFL-transfected COS-1 cells was immunocytochemically determined with the antibodies N1 and C2 (Fig.4); both antibodies reacted with immunoreactive materials localized in the nuclei, indicating that synthesized necdin is translocated into the nucleus (Fig.4A,B). The nuclei of undifferentiated P19 cells showed little reactivity toward the antibody N1 (Fig.4C), but those of the neurally

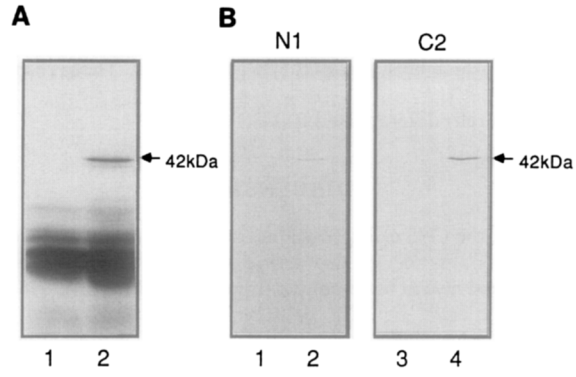

Fig.3. Molecular size determination of 4BFL cDNA-encoded protein (necdin). **A:** cell-free translation of in vitro transcribed 4BFL cDNA. 1: no addition of the transcript 2: addition of the transcript. The arrow indicates the size of the major band (42 KDa). **B:** immunoblotting of 4BFL-encoded protein expressed in COS-1 cells. COS cells were transfected with p94BFL by the DEAE-dextran method. After 72-hr incubation, the cells were lysed, and the samples were analyzed by immunoblotting with the antibodies N1 (1:500) and C2 (1:500). 1 and 3, mock-transfected COS cells; 2 and 4, p94BFL-transfected COS cells. The arrow indicates the size of immunoreactive protein (42 KDa).

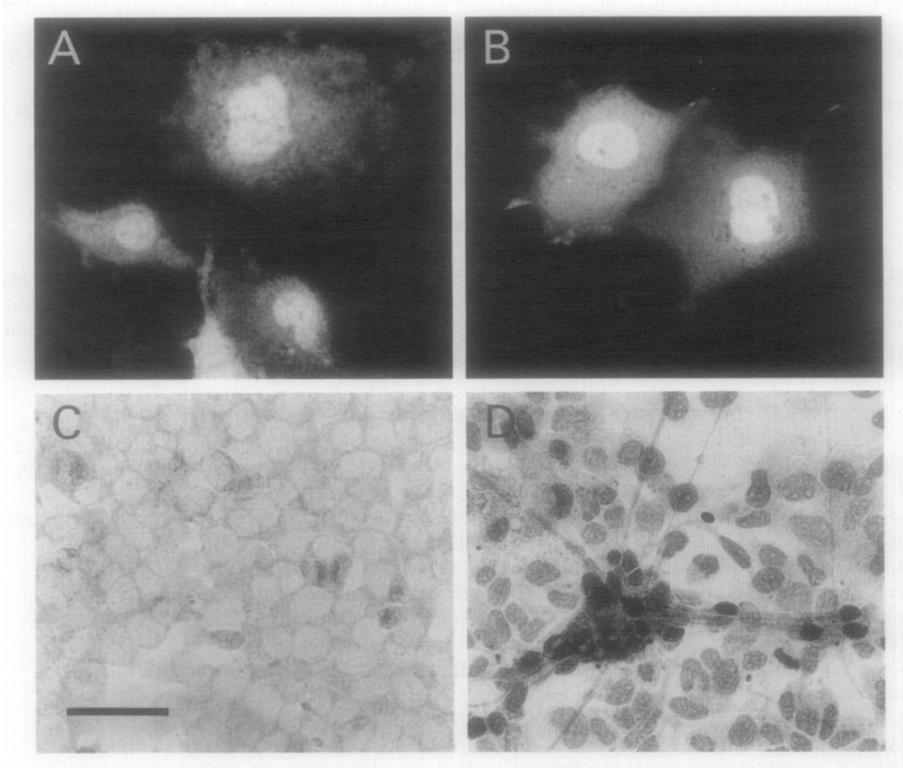

Fig.4. Immunocytochemical localization of necdin expressed in p94BFL-transfected COS cells and differentiated P19 EC cells. COS cells were immunocytochemically examined 48 hrs after transfection with p94BFL. RA-treated P19 EC cells were cultured for 3 days , and stained using the ABC technique. **A:** p94BFL-transfected COS cells, stained with N1 (1:500). **B:** p94BFL-transfected COS cells, stained with C2 (1:500). **C:** undifferentiated P19 stem cells, stained with N1 (1:500) **D:** neurally differentiated P19 cells, stained with N1 (1:500). Smaller and aggregated nuclei are of differentiated neurons. Scale bar = 50 μm for A-D.

differentiated cells were positively stained (Fig.4D); the nuclei of neurons were more immunoreactive than those of non-neuronal cells, of which most are astrocyte progenitors. These results indicate that necdin is a nuclear protein expressed in neurally differentiated P19 EC cells.

DISCUSSION

There are several markers for the CNS cell species; for example, neurofilaments (for neurons), glial fibrillary acidic protein (for astrocytes), and myelin basic protein (for oligodendrocytes). However few markers common to the neuroectodermal cell types have been reported to date. Necdin is present in both neurons and astrocyte progenitors derived from P19 cells (this study). Moreover we have found that necdin is present in the nuclei of neurons and glial cells in primary cultures prepared from mouse brain (unpublished observations). Thus, necdin may represent a useful tool for dissecting molecular bases of the differentiation and development of the CNS cells.

Ectopic expression of exogenous c-jun proto-oncogene sequences in undifferentiated P19 EC stem cells leads to differentiation into mixed populations of endoderm- and mesoderm-like cells (13). It is tempting to speculate that overexpression of necdin in undifferentiated P19 stem cells may induce (or modulate) differentiation along the neuroectodermal pathway. To test this possibility, we are planning to transfect necdin cDNA into P19 EC stem cells. This line of study may provide an insight into physiological roles of this novel nuclear protein in the differentiation and development of the CNS cells.

ACKNOWLEDGMENTS

We are grateful to Drs. M.McBurney (University of Ottawa) and H.Hamada (University of Tokyo) for generous provision of P19 EC cells, to Dr.R.J. Kaufman (Genetics Institute) for the gift of p91023(B) and to Mr. K.Kato for photographic assistance. This work was supported in part by the Ministry of Education, Science and Culture of Japan (Grant Nos. 022401031/02455028/02740376) and by the Nissan Science Foundation.

REFERENCES

1. McBurney, M.W., Jones-Villeneuve, E.M.V., Edwards, M.K.S., and Anderson, P.J. (1982) Nature 299, 165-167.
2. Rudnicki, M.A., and McBurney, M.W. (1987) in Teratocarcinomas and Embryonic Stem Cells: A Practical Approach (Robertson, E.J., Ed) pp. 19-49, IRL Press, Oxford.
3. Aizawa, T., Haga, S., and Yoshikawa, K. (1991) Devel. Brain Res. 59, 89-97.
4. Yoshikawa, K., Aizawa, T., and Maruyama, K. (1990) Biochem. Biophys. Res. Commun. 171, 204-209.
5. Chirgwin, J.M., Przbyla, A.E., MacDonald, R.J., and Rutter, W.J. (1979) Biochemistry 18, 5294-5299.
6. Yoshikawa, K., and Aizawa, T. (1988) Mol. Brain Res. 4, 87-96.
7. Yoshikawa, K., Williams, C., and Sabol, S.L. (1984) J. Biol. Chem. 259, 14301-14308.
8. Yoshikawa, K., Hong, J-S., and Sabol, S.L. (1985) Proc. Natl. Acad. Sci. USA 82, 589-593.
9. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
10. Sompayrac, L.M., and Danna, K.J. (1981) Proc. Natl. Acad. Sci. USA 78, 7575-7578.
11. Maruyama, K., and MacLennnan, D.H. (1988) Proc. Natl. Acad. Sci. USA 85, 3314-3318.
12. Maruyama, K., Terakado, K., Usami, M., and Yoshikawa, K. (1990) Nature, 347, 566-569.
13. de Groot, R.P., Kruyt, F.A.E., van der Saag, P.T., and Kruijer, W. (1990) EMBO J. 9, 1831-1837.