Genomic structure and promoter analysis of putative mouse acetyl-CoA transporter gene[1]

Roop Singh Bora, Shinichi Ichikawa, Akiko Kanamori[2], Yoshio Hirabayashi*

Laboratory for Cellular Glycobiology and Sphingolipid Expression, Frontier Research Program, The Institute of Physical and Chemical Research (RIKEN), 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan

Received 17 March 2000

Edited by Guido Tettamanti

Abstract The acetyl-CoA transporter gene (*Acatn*) encodes a hydrophobic, multitransmembrane protein that is involved in the process of *O*-acetylation of sialic acid residues on gangliosides. *O*-Acetylated gangliosides have been found to play important roles in tissue development and organization during early embryonic stages. We have cloned the gene for mouse acetyl-CoA transporter. The gene spans approximately 20 kb and is composed of seven exons and six introns. A single transcription initiation site, 371 bp upstream of the ATG start codon, was identified. The promoter region was found to lack a TATA box. However, several potential transcription factor binding motifs such as AP1, AP2, C/EBPα, C/EBPβ, HSF, GATA2 and MZF1 were identified in the promoter region.

Key words: Acetyl-CoA transporter; Transcription factor; Exon; Intron

1. Introduction

Gangliosides comprise a diverse group of sialic acid-containing glycosphingolipids present on the plasma membrane of most vertebrate cells. Sialic acid residues on gangliosides are sometimes modified by *O*-acetylation at the 4-, 7-, 8- or 9-hydroxyl position generating further structural diversity of this sugar molecule [1,2]. The occurrence of the *O*-acetylated sialic acids appears to be regulated in a molecule-specific, tissue-specific and developmentally regulated manner, suggesting highly specific roles for this modification system in tissue development and organization [3–6]. The expression of 9-*O*-acetylated gangliosides is associated with neural cell differentiation and migration [7,8]. In addition expression of *O*-acetylated sialic acids on cell surfaces can alter the action of bacterial sialidases [9,10], the binding of pathogenic viruses [11,12], recognition by lectins [13] and activation of the alternative complement pathway [14]. Furthermore, aberrant expression of these ester groups accompanies malignant transformation, sometimes resulting in the production of oncofetal antigens [15]. In spite of its importance, the *O*-acetylation mechanism is poorly understood at the molecular and genetic levels. Earlier studies done by Varki's group [2] indicated that production of *O*-acetylated gangliosides is a complex process, requiring the co-localization of the acceptor gangliosides, acetyl-CoA and acetyltransferase, in the same compartment of the Golgi apparatus. The acetyl-CoA transporter is necessary for translocation of acetyl-CoA from the cytosol to the lumen of the Golgi apparatus. For further understanding of the *O*-acetylation mechanism, the isolation and characterization of each factor involved in the *O*-acetylation of gangliosides is necessary. Previously, we isolated a human cDNA, encoding a novel protein involved in the expression of *O*-acetylated gangliosides. The predicted protein structure and in vitro assay study indicated that this protein is an acetyl-CoA transporter [16]. Recently we have cloned the mouse acetyl-CoA transporter cDNA [17] and assigned the gene to 3E1–E3 [18]. However, the gene encoding the acetyl-CoA transporter (*Acatn*) has not been cloned until now. Our earlier studies showed that *Acatn* is developmentally regulated with a high expression level during early embryonic days and then there is a subsequent decrease in the expression level in the later stages [17]. The analysis of the gene will help in the understanding of the regulatory mechanism of *Acatn* expression and will provide the basis for the creation of mutant mouse models with a disrupted *Acatn* gene. In this paper we report the cloning and structural analysis of the mouse *Acatn* gene including its promoter region.

2. Materials and methods

2.1. Materials

Expand High Fidelity PCR system and digoxigenin (DIG)-DNA labeling kit were obtained from Boehringer Mannheim (Tokyo, Japan). Cycle sequencing kit was from Amersham Life Technologies (USA). All restriction endonucleases were from Toyobo (Tokyo, Japan). pZErO-1 vector and pCRII-TOPO TA cloning kit were obtained from Invitrogen (CA, USA). pGL3-Basic vector, luciferase assay kit and Poly A Tract mRNA isolation kit were purchased from Promega (WI, USA). Marathon cDNA Amplification Kit and Advantage cDNA polymerase mix were obtained from Clontech (CA, USA). SuperFect transfection reagent was purchased from Qiagen (Germany). All other reagents were of analytical grade.

2.2. DNA manipulation and nucleotide sequence determination

DNA manipulation, colony hybridization and Southern blot analysis were performed according to Sambrook et al. [19]. The nucleotide sequence was determined in both directions using a cycle sequencing

*Corresponding author. Present address: Laboratory for Memory and Learning, Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan. Fax: (81)-48-467 6317.
E-mail: hirabaya@postman.riken.go.jp

[1] The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases with accession numbers AB037362, AB037363, AB037364, AB037365, AB037366, AB037367 and AB037368.

[2] Present address: Laboratory of Experimental Pathology, Research Institute Aichi Cancer Center, 1-1 Kanokodan, Chikusa-ku, Nagoya 464-8681, Japan.

Abbreviations: PCR, polymerase chain reaction; Acatn, acetyl-CoA transporter; *Acatn*, gene encoding Acatn; DIG, digoxigenin; UTR, untranslated region; RACE, rapid amplification of cDNA ends

PII: S0014-5793(00)01524-6

kit, based on the dideoxy-chain termination method [20]. Li-COR 4000L sequencer (Li-COR, NE, USA) was used for analysis.

2.3. Isolation and characterization of the mouse Acatn gene

One clone containing the entire *Acatn* gene in pBeloBAC11 [21] vector was isolated by bacterial artificial chromosome (BAC) high-density filter screening (Genome System Inc., USA), using the ^{32}P-labeled 1.65 kb *Bam*HI/*Xho*I fragment (corresponding to nucleotide positions 1–1653 of the coding region) of mouse *Acatn* cDNA [17] as a probe. The isolated genomic clone was designated pBAC17269. This clone was digested with *Eco*RI, *Xba*I, *Bam*HI or *Pst*I and subcloned into the pZErO-1 vector. Clones harboring exons were isolated by colony hybridization using various regions of mouse *Acatn* cDNA as probes. The DIG-DNA labeling kit was used for probe preparation according to the manufacturer's instruction. The DNA fragment containing exon sequences was also isolated by PCR using BAC clone as the template directly.

2.4. PCR

The Expand High Fidelity PCR system was used according to the manufacturer's instruction. PCR was performed in 50 μl of reaction mixture containing 100 ng of template DNA, 1×Expand HF buffer, 2.5 mM $MgCl_2$, 200 μM each of dNTPs, a forward primer, a reverse primer (40 pmol each) and enzyme mix (0.75 U). Forward primer: 5′-GGAACATACATGACCCTTTT-3′ (corresponding to nucleotide positions 1723–1742) and reverse primer: 5′-CAAATCTTTAA-TGCTTACAAAAGCAGG-3′ (corresponding to nucleotide positions 2666–2640 of the human *Acatn* cDNA sequence) were used for DNA amplification. Thirty PCR cycles were carried out and each PCR cycle consisted of denaturation at 94°C for 1 min, annealing at 55°C for 1 min and extension at 72°C for 3 min. The amplified fragment was subcloned into the pZErO1 vector and the nucleotide sequence was determined.

2.5. Determination of the transcriptional initiation site

5′ Rapid amplification of DNA ends (RACE) was conducted to determine the transcriptional initiation site of the *Acatn* gene. Briefly, mRNA was isolated from the brain of 8–10 week old male BALB/c mice using the Poly A Tract mRNA isolation kit following the manufacturer's protocol. cDNA was synthesized from the above mRNA using the Marathon cDNA Amplification kit following the protocol of the manufacturer. Marathon cDNA adapter (5′-CTAATAC-GACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT-3′) was ligated to the double stranded cDNA. 5′ RACE PCR was performed using adapter primer 1 (AP1: 5′-CCATCCTAATACGACT-CACTATAGGGC-3′) and *Acatn*-specific antisense primer (5′-ACA-TTGCCCAAAAAGTAGCC-3′), corresponding to nucleotide positions 689–670. PCR was performed in 50 μl of reaction mixture containing 5 μl of diluted adapter-ligated double stranded cDNA, 1×cDNA PCR reaction buffer, 200 μM of each dNTP, AP1 primer, *Acatn*-specific antisense primer and 1 μl of Advantage cDNA polymerase mix (50×). Thirty PCR cycles were carried out and each cycle consisted of denaturation at 94°C for 1 min, annealing at 56°C for 1 min and extension at 72°C for 2 min. The resulting PCR products were analyzed by Southern blot hybridization using DIG-labeled *Acatn* cDNA fragment (corresponding to nucleotide positions 1–689) as probe. The PCR fragment that hybridized to the probe was cloned into the pCRII-TOPO vector using a TOPO-TA cloning kit following the protocol of the manufacturer. The *Acatn* transcriptional start site was identified by sequencing the cloned 5′ RACE fragment.

2.6. Reporter gene assay

A series of *Acatn* promoter and luciferase gene chimeras with progressively smaller promoter regions were constructed by site-directed PCR amplification. The sequences of the primers used were as follows:

- F1: 5′-AAGTTAAGGCAGGAGTATGGCCA-3′ (−1202 to −1180)
- F2: 5′-ACCTCCTTCACTCTGTGTATCTAG-3′ (−599 to −576)
- F3: 5′-GGAGTGTTTCCATTTAGTCGGGCA-3′ (−507 to −484)
- F4: 5′-TGGAAGTTATGTACTGAATTAAATA-3′ (−407 to −383)
- F5: 5′-ACGAGGGCCCAATGGCCGAGG-3′ (−302 to −282)
- F6: 5′-TGGCGTTATTCCTCCCTCGGGG-3′ (−204 to −183)
- F7: 5′-CTCCGTCGGGAGGCACGGAACC-3′ (−104 to −83)
- R1: 5′-GAGCTTGAAATGTTGAGATACTAG-3′ (−744 to −767)
- R2: 5′-GCGGGCACCCTCCTCAGTAG-3′ (−1 to −20).

PCR was performed as described above using clone 1 containing exon I and the 5′ flanking region of the *Acatn* gene, as template. Thirty cycles were carried out and each cycle consisted of denaturation at 94°C for 1 min, annealing at 60°C for 1 min and extension at 72°C for 2 min. The amplified fragments were purified and cloned into the *Sma*I site of pGL3 Basic vector. For transfection, HeLa cells were grown in 6-well plates (1×10^5 cells) in DMEM (Dulbecco's modified Eagle's medium, Gibco BRL) containing 10% fetal bovine serum (FBS) and grown in a humidified 5% CO_2 incubator at 37°C. Cells were transfected with 10 μg of plasmid DNA using SuperFect transfection reagent following the manufacturer's instruction. After 36 h of incubation cells were harvested, lysed and assayed for luciferase activity using the Luciferase Enzyme Assay kit following the manufacturer's protocol.

3. Results and discussion

3.1. Genomic organization of Acatn

As described in Section 2, one clone, designated pBAC17269, was isolated from a mouse 129Sv/J ES genomic library maintained in BAC vector. The genomic clone pBAC17269 was found to contain the entire *Acatn* gene by Southern blot analysis using various regions of mouse *Acatn* cDNA as probes (data not shown). This clone was digested with either *Eco*RI, *Bam*HI, *Xba*I or *Pst*I and cloned into the pZErO1 vector. Various subclones harboring the exon sequences were identified by colony hybridization using various regions of the *Acatn* cDNA as probes. The exon sequences and exon–intron boundaries of *Acatn* were determined by sequencing the subcloned fragments. In some cases, PCR-amplified fragments were used for analysis, after cloning into the pZErO1 vector. The structure of the *Acatn* gene is shown in Fig. 1. The gene spans approximately 20 kb and is composed of seven exons interrupted by six introns. Exon 1 contains 371 bp of 5′ untranslated region (UTR) and 775 bp of the coding sequence of the *Acatn* gene. Exon 7 contains 168 bp of coding sequence of *Acatn* and approximately 600 bp of 3′ UTR as confirmed by human cDNA sequence and PCR analysis. The sizes of the exons range from 54 to 1146 bp. Exon 1 is exceptionally large and its coding sequence corresponds to a highly conserved region among human, mouse, *Drosophila*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae* and *Escherichia*

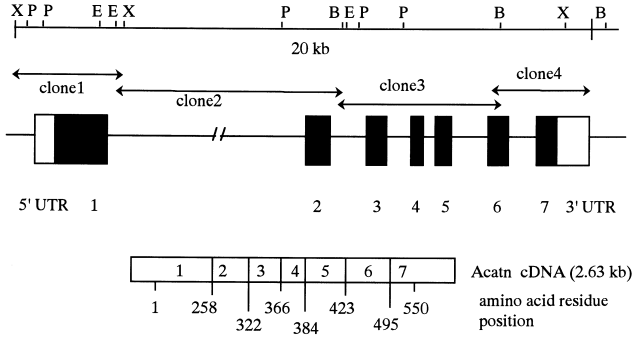

Fig. 1. Genomic organization of the *Acatn* gene. Exons are numbered and depicted as boxes. Open boxes indicate untranslated regions. The solid lines represent introns. Shown at the top is a 20 kb fragment of pBAC17269 that contains the entire *Acatn* gene, and displayed below are the subclones of pBAC17269 used for analysis. Clone 4 was obtained by cloning the PCR-amplified fragment into the pZErO1 vector. At the bottom, the exon numbers and their positions in *Acatn* cDNA are indicated. X, B, E and P stand for *Xba*I, *Bam*HI, *Eco*RI and *Pst*I, respectively.

Table 1
Exon–intron boundaries of the mouse *Acatn* gene

Exon	Exon size (bp) (coding region)	5′ splice site	3′ splice site	Intron no. (approximate size, kb)
1	1146 (775)	CTTCAG**gt**agtgtactt	tctaattc**ag**ATTTTC	I (9.9)
2	191	TCAAAG**gt**aagagatg	atgttgc**ag**ATTGGC	II (2.4)
3	131	CAGCAA**gt**atactgca	atcagca**ag**GTATAC	III (0.8)
4	54	CTACAG**gt**aactacag	ctctttgc**ag**ATTATT	IV (0.48)
5	118	CATCAG**gt**aagcgga	cttcttgc**ag**GTCACT	V (2.2)
6	216	ATCGAA**gt**aagtgagt	tgtttttat**ag**CTTTGC	VI (0.9)
7	769 (168)	–	–	–

Exon and intron sequences are shown in upper and lower case letters, respectively. The intron sequences adjoining the splice junctions are shown in bold.

coli homologues, on the basis of amino acid sequence [17], suggesting that it may be a functionally important domain of mouse *Acatn* and its homologues in various organisms. The nucleotide sequences of all the exons were identical to the corresponding sequences of mouse *Acatn* cDNA. The sizes of the introns range from 0.48 to 9.9 kb. The intron sizes were determined by Southern blot analysis, PCR and DNA sequencing. The exon–intron boundary sequences at the 5′ and 3′ ends of all the introns are GT and AG, respectively (Table 1), which are consensus sequences for pre-mRNA splicing recognition donor and acceptor sites [22].

The transcriptional initiation site of the *Acatn* gene was determined by 5′ RACE PCR. Southern blot analysis of the 5′ RACE products, using an *Acatn*-specific probe, detected a single band of approximately 1 kb, indicating that there is a single transcriptional start site. Sequence analysis identified the transcriptional initiation site, 371 nucleotide upstream of the ATG start codon, marked as position +1 in Fig. 2 and found to be 'A', a purine, the most common transcription initiating nucleotide.

3.2. The promoter region of Acatn

Clone 1, harboring a 2.7 kb *Xba*I-digested fragment from pBAC17269, was identified by colony hybridization. It contained 1.2 kb 5′ flanking region, 371 bp of 5′ UTR, all of exon 1 (775 bp) and part of intron I. Analysis of the 1.2 kb promoter region, immediately upstream of the transcription start site, failed to find any canonical TATA box or TATA-like sequence. However, several potential transcription factor binding motifs were identified in the promoter region between −904 and −1, using a computer-based sequence analysis, as shown in Fig. 2. These include AP1 (−554 to −548 and −467 to −461), AP2 (−295 to −288, −134 to −127 and −104 to −97), C/EBPα (−427 to −414), C/EBPβ (−363 to −352), MZF1 (−549 to −542) and GATA2 (−351 to −343 and −41 to −33) binding sites. In addition a few binding sites for HSF and cdxA transcription factors and CCAAT box (−831 to −827 and −819 to −815) were identified.

A series of *Acatn* promoter and luciferase gene chimeras with progressively smaller promoter regions were constructed by site-directed PCR amplification. The promoter activity of each construct was determined by transfection of recipient HeLa cells and measurement of resultant luciferase activity. The highest level of promoter activity was observed with the construct containing the sequence from −599 to −1 (Fig. 3, luc 3). Deletion of the sequence from −599 to −508 dramatically decreased the promoter activity (Fig. 3, luc 4), indicating that this particular region has a positive regulatory effect. This region contains AP1 and MZF1 binding motifs. Although *Acatn* appears to be a housekeeping gene because of its ubiquitous mRNA expression [16,17] and has a TATA-less promoter region, we could not identify any SP1 motif in the *Acatn* promoter, which is commonly required for efficient promoter activity, in many other housekeeping genes that lack a TATA box. As shown in Fig. 3, substantial promoter ac-

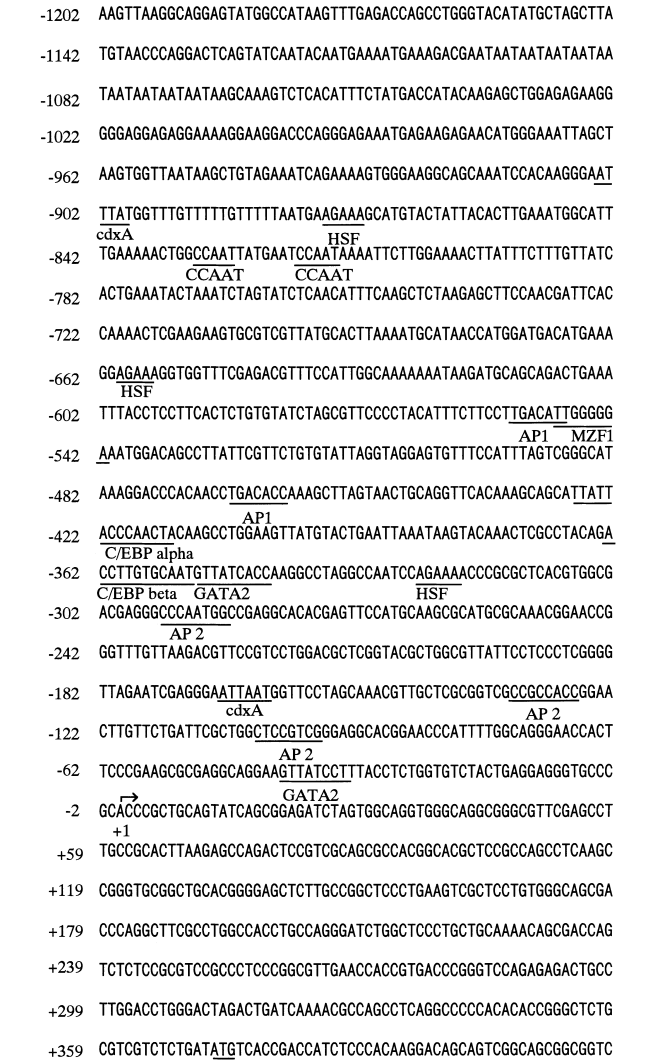

Fig. 2. The promoter region of *Acatn* and the consensus sequences of transcription factor binding elements. The sequence shows the 371 bp of 5′ UTR and 47 bp of coding region of exon 1 and 1202 bp of promoter region. The transcriptional initiation site is indicated by the arrow (nucleotide +1). The translation initiation site is underlined. Consensus sequences for AP1, AP2, C/EBPα, C/EBPβ, HSF, cdxA, GATA 2 and MZF1 binding are indicated.

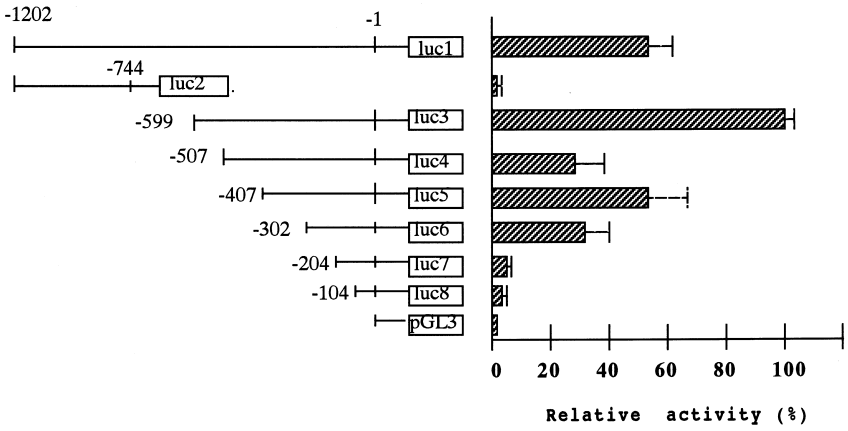

Fig. 3. Deletion analysis of the *Acatn* promoter. Luciferase (Luc) expression clones containing sequentially deleted fragments of the *Acatn* chimeric promoter constructs were transfected into HeLa cells. Following transfection, the cells were incubated for 36 h in DMEM containing 10% FBS. Luciferase activities were measured and the means (±S.D.) of triplicate transfection experiments were calculated. They are expressed as the percent activity relative to that observed in luc 3 (100%).

tivity was seen even after deletion of the promoter sequence from −1202 to −303 (Fig. 3, luc 6), but further deletion completely abolished the promoter activity (Fig. 3, luc 7 and 8), suggesting that the sequence from −302 to −205, which contains an AP2 binding motif (−295 to −288), is essential for promoter activity.

Our previous studies have indicated that the *Acatn* gene is developmentally regulated with a high expression level of *Acatn* mRNA during early embryonic stages as compared to later stages [17]. Interestingly, the *Acatn* promoter has been found to contain a binding motif for C/EBPα transcription factor, which is intimately linked to cellular differentiation and proliferation control in a variety of tissues [23]. C/EBPα was also found to be required for efficient insulin-sensitive glucose transport in murine fibroblast cell lines [24]. Similarly transcription factor HSF is important for the normal growth and development of cells. Deletion of the HSF gene in *Drosophila* resulted in defects in oogenesis and early larval development, with evidence indicating that these defects were not due to altered basal heat shock protein expression but rather to some other unknown target genes [25]. Hence, there is a possibility that these transcription factors including C/EBPα, HSF and AP1 might play important roles in the regulation of *Acatn* expression during development.

The data obtained from these studies will provide the basis for future studies on the regulation of the *Acatn* gene and for creation of mutant mouse models with altered *Acatn* expression.

Acknowledgements: We thank Dr. Yoshitaka Nagai for generous help and encouragement. This work was supported by a Frontier Research Program Grant of the Science and Technology Agency and a Grant-in-Aid for Scientific Research of Priority Area from the Ministry of Education, Science and culture to Y.H. (No. 05274106).

References

[1] Schauer, R. (1987) Methods Enzymol. 138, 132–161.
[2] Varki, A. (1992) Glycobiology 2, 25–40.
[3] Varki, A., Hooshmand, F., Diaz, S., Varki, N.M. and Hedrick, S.M. (1991) Cell 65, 65–74.
[4] Schlosshauer, B., Blum, A.S., Mendez-Otero, R., Barnstable, C.J. and Constantine-Paton, M. (1988) J. Neurosci. 8, 580–592.
[5] Zhang, G., Li, J., Kurono, S., Fujita, S.C., Furuya, S. and Hirabayashi, Y. (1997) Glycoconjug. J. 14, 847–857.
[6] Muchmore, E., Varki, N., Fukuda, M. and Varki, A. (1987) FASEB J. 1, 229–235.
[7] Mendez-Otero, R. and Cavalcante, L.A. (1996) Neurosci. Lett. 204, 97–100.
[8] Constantine Paton, M., Blum, A.S., Mendez-Otero, R. and Barnstable, C.J. (1986) Nature 324, 459–462.
[9] Corfield, A.P., Sander Wewer, M., Veh, R.W., Wember, M. and Schauer, R. (1986) Biol. Chem. Hoppe-Seyler 367, 433–439.
[10] Varki, A. and Diaz, S. (1983) J. Biol. Chem. 258, 12465–12471.
[11] Higa, H.H., Rogers, G.N. and Paulson, J.C. (1985) Virology 144, 279–282.
[12] Rogers, G.N., Herrler, G., Paulson, J.C. and Klenk, H.D. (1986) J. Biol. Chem. 261, 5947–5951.
[13] Ravindranath, M.H. and Paulson, J.C. (1987) Methods Enzymol. 138, 520–527.
[14] Varki, A. and Kornfeld, S. (1980) J. Exp. Med. 152, 532–544.
[15] Cheresh, D.A., Reisfeld, R.A. and Varki, A. (1984) Science 225, 844–846.
[16] Kanamori, A., Nakayama, J., Fukuda, M.N., Stallcup, W.B., Sasaki, K., Fukuda, M. and Hirabayashi, Y. (1997) Proc. Natl. Acad. Sci. USA 94, 2897–2902.
[17] Bora, R.S., Kanamori, A. and Hirabayashi, Y. (1999) Gene 238, 455–462.
[18] Bora, R.S., Kanamori, A. and Hirabayashi, Y. (1998) Cytogenet. Cell Genet. 83, 78–79.
[19] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Plainview, NY.
[20] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[21] Shizuya, H., Birren, B., Kim, U.-J., Mancino, V., Slepak, T., Tachiri, Y. and Simon, M. (1992) Proc. Natl. Acad. Sci. USA 89, 8794–8797.
[22] La Thangue, N.B. and Rigby, P.W.J. (1988) in: Transcription and Splicing (Hames, B.D. and Glover, D.M., Eds.), pp. 1–42, IRL Press, Washington, DC.
[23] Muller, C., Alunni-Fabbroni, M., Kowenz-Leutz, E., Mo, X., Tommasino, M. and Leutz, A. (1999) Proc. Natl. Acad. Sci. USA 96, 7276–7281.
[24] El-Jack, A.K., Hamm, J.K., Pilch, P.F. and Farmer, S.R. (1999) J. Biol. Chem. 274, 7946–7951.
[25] Jedlicka, P., Mortin, M.A. and Wu, C. (1997) EMBO J. 16, 2452–2462.