THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 274, No. 36, Issue of September 3, pp. 25623–25631, 1999

Printed in U.S.A.

Lysosomal and Cytosolic Sialic Acid 9-*O*-Acetylesterase Activities Can Be Encoded by One Gene via Differential Usage of a Signal Peptide-encoding Exon at the N Terminus*

(Received for publication, May 12, 1999, and in revised form, June 24, 1999)

Hiromu Takematsu‡, Sandra Diaz‡, Angela Stoddart§, Yu Zhang§, and Ajit Varki‡¶

From the ‡Glycobiology Research and Training Center, Divisions of Hematology-Oncology and Cellular and Molecular Medicine, University of California San Diego, La Jolla, California 92093 and the §Ontario Cancer Institute, University of Toronto, Ontario M5G 2M9, Canada

9-*O*-Acetylation is one of the most common modifications of sialic acids, and it can affect several sialic acid-mediated recognition phenomena. We previously reported a cDNA encoding a lysosomal sialic acid-specific 9-*O*-acetylesterase, which traverses the endoplasmic reticulum-Golgi pathway and localizes primarily to lysosomes and endosomes. In this study, we report a variant cDNA derived from the same gene that contains a different 5′ region. This cDNA has a putative open reading frame lacking a signal peptide-encoding sequence and is thus a candidate for the previously described cytosolic sialic acid 9-*O*-acetylesterase activity. Epitope-tagged constructs confirm that the new sequence causes the protein product to be targeted to the cytosol and has esterase activity. Using reverse transcription-polymerase chain reaction to distinguish the two forms of message, we show that although the lysosomal sialic acid-specific 9-*O*-acetylesterase message has a widespread pattern of expression in adult mouse tissues, this cytosolic sialic acid 9-*O*-acetylesterase form has a rather restricted distribution, with the strongest expression in the liver, ovary, and brain. Using a polyclonal antibody directed against the 69-amino acid region common to both proteins, we confirmed that the expression of glycosylated and nonglycosylated polypeptides occurred in appropriate subcellular fractions of normal mouse tissues. Rodent liver polypeptides reacting to the antibody also co-purify with previously described lysosomal sialic acid esterase activity and at least a portion of the cytosolic activity. Thus, two sialic acid 9-*O*-acetylesterases found in very different subcellular compartments can be encoded by a single gene by differential usage of a signal peptide-encoding exon at the N terminus. The 5′-rapid amplification of cDNA ends results and the differences in tissue-specific expression suggest that expression of these two products may be differentially regulated by independent promoters.

Sialic acids (Sias)[1] are a diverse family of acidic 9-carbon sugars typically found at the nonreducing end of sugar chains of animals throughout the deuterostome lineage (1, 2). Sias are now recognized not only as molecules responsible for negative charge and hydrophilicity of the cell surface but also as specific ligands playing important roles in intercellular and/or intermolecular recognition phenomena (3). There have been several vertebrate receptors reported for Sia-containing determinants, including the selectin family (4–6) and the Siglec family (7–10). Sias are also known to be targets for a variety of microbial proteins involved in host cell recognition, *e.g.* the influenza virus hemagglutinin (11) and one of the adhesins of *Helicobacter pylori* (12–14).

The Sia molecule can be modified by the addition of *O*-acetyl esters on the hydroxyl-groups of the 4, 7, 8, and 9 positions (2). This *O*-acetylation of Sia is a common modification found in mammalian cell surface sialoglycoconjugates and can also be present on free Sias in the cytosol (15, 16). One example of a well characterized *O*-acetylated sialoglycoconjugate is 9(7)-*O*-acetylated-G_{D3} and structurally related disialogangliosides, which show developmentally regulated expression and gradient patterns for its staining with monospecific antibodies in the developing brain (17, 18). Specific antibodies showed that the expression of 9-*O*-acetylated G_{D3} appears to be independently regulated from that of the parental molecule, G_{D3} (18), and occurs on the migrating neural cells (19).

Despite the widespread and regulated occurrence of this modification, the biological significance of *O*-acetylation is largely unknown. It has been shown that *O*-acetylation of Sia can mask the Sia-containing determinant from recognizing its counter receptor lectin Siglec-2 (CD22) (20). Direct or indirect evidence indicates that binding of other members of the Siglec family to their sialylated ligands can also be blocked by *O*-acetylation (21–25). In contrast, influenza virus C hemagglutinin specifically requires 9-*O*-acetylated Sia for binding to host cells (11, 26). These examples indicate that *O*-acetylation may be a key modification regulating Sia-dependent recognition events. In fact, cleavage of 9-*O*-acetyl groups from Sia molecules by transgenic expression of influenza C virus hemagglutinin 9-*O*-acetylesterase caused abnormalities in murine development (27).

The biosynthesis of *O*-acetylated sialoglycoconjugates is catalyzed by Sia-specific *O*-acetyltransferase(s) that use acetyl-CoA as an acetyl donor (28). This *O*-acetylation reaction appears to take place in the late Golgi apparatus, after the action of sialyltransferases (29). Various lines of evidence indicate the existence of a family of 9(7)-*O*-acetyltransferase(s) that are linkage- and molecule -specific in their action, likely explaining

* This work was supported by Grants HL57345 and GM32373 from the U. S. Public Health Service (to A. V.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF156856.

¶ To whom correspondence should be addressed: Cancer Center, 0687, UCSD School of Medicine, La Jolla, CA 92093-0687. Tel.: 619-534-3296; Fax: 619-534-5611.

[1] The abbreviations used are: Sia, sialic acid; Cse, cytosolic sialic acid 9-*O*-acetylesterase; Lse, lysosomal sialic acid-specific 9-*O*-acetylesterase; RT, reverse transcription; PCR, polymerase chain reaction; ConA, Concanavalin A; bp, base pair(s); PAGE, polyacrylamide gel electrophoresis; PNGase F, peptide *N*-glycosidase F.

the tissue- and stage-specific expression of this modification (30). These labile Golgi enzymes have so far proven refractory to purification or cloning. This modification can be removed by Sia-specific 9-*O*-acetylesterases. We and others have previously described two distinct forms of mammalian Sia *O*-acetylesterases, one in the cytosolic fraction and another in the lysosomal/endosomal compartment (31–34). This localization of the lysosomal Sia 9-*O*-acetylesterase (Lse) is somewhat unusual, considering its neutral pH optimum (34). Regardless, this enzyme is likely to be the one participating in the terminal lysosomal degradation of 9-*O*-acetylated sialoglycoconjugates. However, it would not have access to free 9-*O*-acetylated Sias that have been reported in the cytosolic fraction. These presumably result from the action of lysosomal sialidases on 9-*O*-acetylated Sias, followed by export of such molecules to the cytosol by the action of the lysosomal Sia exporter (15, 16). Such 9-*O*-acetylated Sias entering the cytosol are known to be poor substrates for reactivation by CMP-Sia synthase (35), and CMP-9-*O*-acetylated Sias in turn are poor substrates for some sialyltransferases (35). Thus, we have suggested that the function of the cytosolic Sia 9-*O*-acetylesterase (Cse) activity is to salvage any 9-*O*-acetylated molecules that escape the initial action of the Lse enzyme.

The Cse and Lse enzyme activities are both found in the ultracentrifugate supernatant of freeze-thawed rat liver (the soluble Lse enzyme is presumably released by the breakage of lysosomal membranes during freeze-thaw) or in detergent extracts. The two activities can be biochemically separated by ConA-Sepharose chromatography (the Lse binds and the Cse runs through the column) and/or by DEAE-cellulose chromatography (the Cse binds and the Lse runs through) (36). Furthermore, the reported properties of the Cse are somewhat different from that of the Lse (32, 37). Finally, some antibodies directed against the Lse did not seem to precipitate the Cse activity (33). Based on all of these criteria, it has always been assumed that the Lse and Cse enzymes are products of separate genes.

We earlier purified the Lse to homogeneity and obtained N-terminal amino acid sequences of the two subunits (33). These sequences were subsequently used to identify the Lse cDNA, which was initially isolated serendipitously during differential display analysis of different stages of murine hematopoietic development (38). The cDNA encoding the Lse was also independently cloned by means of differential display comparisons between different stages of the murine B lymphoid cell lineage (39). During the characterization of the latter system, several cDNAs were noted to have 5′ sequences that were different from the originally isolated Lse form. Here we report the characterization of one of these cDNAs and show that it can encode a cytosolic counterpart of the Lse.

EXPERIMENTAL PROCEDURES

Materials—All reagents used for this study were of appropriate grade for either biochemistry or molecular biology and unless otherwise stated were obtained from Sigma. Unless otherwise stated, typical experiments followed the *Current Protocols in Molecular Biology* (40), *Current Protocols in Protein Science* (41), or *Molecular Cloning* (42).

FLAG-tagged cDNA Constructs for the Transfection into COS7 Cells—7A3-C is a cDNA clone previously isolated from the 70Z/3 pre B-cell line cDNA library by cross-hybridization to the Lse (39). We subcloned the Lse into the expression vector pcDNAI/Amp (Invitrogen) after PCR-derived mutagenesis to add a FLAG epitope in its C terminus (Lse-FLAG) as reported previously (38). We then exchanged the 5′ region of Lse from the 7A3-C using the *Bal*I site within the coding sequence of both forms and *Hin*dIII site in the multiple cloning site of pcDNAI/Amp (giving Cse-C-FLAG). Plasmid DNA was prepared using a Qiagen maxi-column (Qiagen) and transiently transfected into COS cells using LipofectAMINE (Life Technologies, Inc.), according to the manufacturer's instructions. Briefly, 20 μg of plasmid DNA is premixed with 80 μl of LipofectAMINE reagent, and this mixture was used to transfect COS7 cells 18 h after the subculture into 10 each of 140-mm dishes. At the time of transfection, the cell density was ~25–30% confluent. After the transfection procedure in Opti-MEM for 5 h, α-minimum essential medium with 20% fetal calf serum was added to a final of 10%, followed by an overnight incubation in α-minimum essential medium.

Western Blotting Analysis of Cse-C Transiently Expressed in COS7 Cells—Transfected cells were washed with phosphate-buffered saline, scraped off from the dish, washed twice with phosphate-buffered saline, and homogenized in 25 mM Tris-HCl, pH 7.3, containing 1 mM dithiothreitol using gentle sonication. The homogenate was centrifuged 500 × *g* for 5 min, and the post-nuclear supernatant was then ultracentrifuged at 100,000 × *g* for 40 min. The resulting supernatant was then incubated with anti-FLAG M2-Sepharose (Eastman-Kodak) resin overnight at 4 °C with rotation. The M2 resin was packed into a minicolumn and eluted with 100 mM glycine-HCl, pH 3.0. The eluted fractions were neutralized immediately by adding drops of 1 M Tris-HCl, pH 8.5, and aliquots of each were subjected to 12.5% SDS-PAGE in either reduced or nonreduced conditions. Gels were then electrotransferred to nitrocellulose membranes (Bio-Rad) in a wet condition and blocked in 2% dry milk in Tris-buffered saline with 0.2% Tween. The blot was exposed to the primary anti-FLAG antibody M2 (1:2500 in Tris-buffered saline) for 1 h followed by washing and exposure to the secondary goat anti-mouse IgG horseradish peroxidase (Bio-Rad) for 1 h. Membranes were washed for more than 1 h. Horseradish peroxidase activity was then developed using the Supersignal substrate (Pierce) for 5 min. The signal was visualized by detection of chemiluminescence using x-ray film (Kodak) for periods ranging from a few seconds to 10 min.

Indirect Immunofluorescent Staining of Transiently Transfected COS7 Cells—COS7 cells were transfected with FLAG-tagged constructs in 8-well chamber slides as described above. After overnight culture, cells were fixed with ice-cold acetone:methanol 1:1 and subjected to staining (the signal following 2 days of culture was too strong for the analysis). Briefly, fixed cells were washed with phosphate-buffered saline, incubated with 10 μg/ml of anti-FLAG antibody M2 (Eastman-Kodak) for 30 min at room temperature. Unbound antibody was washed off, and the cells were incubated with 1 μg/ml of fluorescein isothiocyanate-conjugated goat anti-mouse IgG for 30 min at room temperature. Unbound antibody was again washed off, and the slides were coverslipped using 9:1 glycerol/phosphate-buffered saline for viewing with a Zeiss Axiophot epifluoresence microscope. Images were captured using Adobe Photoshop software with a Sony CCD camera.

RT-PCR Analysis of Two Forms of cDNA—Total cellular RNA was purified from various tissues of C57Bl/6 strain mice using Trizol reagent (Life Technologies, Inc.). Total RNA (20 μg) was reverse transcribed by superscript II (Life Technologies, Inc.) in a 20-μl reaction. 2.5 μl of first strand cDNA was used for PCR with 30 cycles for Lse amplification and 40 cycles for Cse-C amplification (preincubation at 94 °C for 2 min, 94 °C for 0.5 min, 60 °C for 1 min, and 72 °C for 2 min) using a Gene Amp 2400 (Perkin-Elmer). Primer sequences for the amplification were carried out using Lse-1 (anneals specific to Lse cDNA) *versus* Lse-J (should anneal to both forms) for Lse amplification and Cse-11 (anneals specific to Cse-C cDNA) *versus* Lse-J for Cse-C amplification. Sequence and position of these primers are as follows, Lse-1 (sense primer for Lse corresponding to positions cDNA 2–23) 5′-atcaggatcttcacaaacatggt-3′, Lse-J (antisense primer for Lse/Cse-C corresponding to Lse cDNA positions 816–795) and 5′-gaaacctctcttctggacaag-3′, and Cse-11 (sense primer for Cse-C corresponding to cDNA positions 124–144) 5′-aaaggacatgaggactcctcac-3′. Predicted amplified fragments were 810 bp for Lse and 890 bp for Cse-C (more than one band including 890 bp was actually seen for the latter). RT-PCR products were then directly sequenced or subcloned into a TA-cloning vector pCR II (Invitrogen), and multiple clones were sequenced by the cycle dideoxy termination sequencing method with an ABI model 310 autosequencer using M13–20 forward or reverse primers as primers. Not all the bands shown in agarose gel were recovered after gel purification, possibly because some are actually heterogenously annealed DNAs of different size.

Production of a Chicken Anti-serum Directed against the Common Esterase C Terminus—We had previously found that rabbits produced very low titer antibodies against intact Lse (33). We therefore decided to raise antibodies against a defined peptide region in chickens. The 69-amino acid portion shared by Lse and Cse-C cDNAs (see Fig. 1*B*) was amplified using Pwo polymerase (Boehringer Mannheim) and the product subcloned into pMAL-p2 vector (New England Biolabs) as an in-frame maltose-binding protein fusion protein. The fusion protein was

induced in LB containing 0.3 mM isopropyl 1-thio-β-D-galactoside at 37 °C for 3 h. Secreted fusion protein was directly purified from the medium by passage through a maltose column. The purity of the purified protein was confirmed by Coomassie Brilliant Blue staining after SDS-PAGE. The purified protein was used to immunize Rhode Island Red female chickens four times over a period of 2 months. After 2 months from the initial immunization, nonfertile eggs were collected, and the IgY was purified from egg yolk by using the Promega Eggstract kit.

Preparation and Fractionation of a High Speed Supernatant from Mouse Liver—Mouse (C57Bl/6) liver was subjected to three cycles of freeze and thaw in hypotonic buffer to release the soluble lysosomal enzymes. The mixtures were then homogenized on ice with a polytron for a total of 2 min with intervals of 30 s on ice. The homogenates were ultracentrifuged at 100,000 × *g* for 40 min, and the supernatant fractions were then subjected to column chromatography on ConA-Sepharose (Amersham Pharmacia Biotech). The bound fractions were eluted with 100 mM α-methylmannopyranoside in 20 mM KPO_4 pH 8, 0.1 M NaCl. ConA run-through fractions positive for Sia 9-*O*-acetylesterase activity (see assay described below) were pooled, dialyzed against 20 mM KPO_4 pH 8, and loaded onto a DE52 (DEAE-cellulose) column (Whatman). After washing, the bound proteins were eluted with 15 ml of 150–400 mM NaCl gradient, followed by 5 ml of 400 mM NaCl (36).

Western Blotting Analysis of Chromatography Fractions—25 μl each of the fractions from the ConA and DE52 columns were subjected to 12.5% SDS-PAGE under reduced conditions as described above. Transferred protein blots were incubated with anti-esterase C12 IgY antibody as the primary antibody, followed by development with a peroxidase-conjugated donkey anti-chicken IgY secondary antibody (Bio-Rad) and Supersignal substrate (Pierce).

Assay for the Sialic Acid 9-O-Acetylesterase—As described previously (43), [9-*O*-acetyl-^{3}H]Neu5,9Ac_2 is prepared by labeling purified rat liver Golgi with [acetyl-^{3}H]acetyl coenzyme A. The enzyme source (50–100 μl) is incubated with the substrate (10,000 cpm) at 37 °C for 1 h, and the reaction is quenched by the addition of equal volume of "stopping mixture," followed by mixing with a toluene-based scintillation mixture. The uncleaved substrate cannot enter the toluene phase, but free [^{3}H]Acetate (the cleavage product) can, allowing determination of the activity.

PNGase F Digestion of the Lse/Cse-C Fractions—Aliquots of fractions positive for the activity/Western blot signal for the chicken antibody were incubated with PNGase F for the cleavage of *N*-linked oligosaccharides. The digested material was compared with sham-treated material and analyzed by Western blotting.

RESULTS

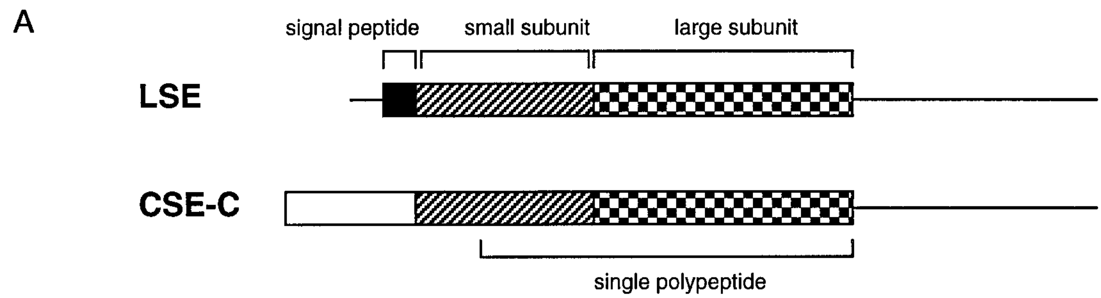
Two Different Types of Messages Are Encoded by the Lse Gene—During characterization of the cDNA for the lysosomal Sia 9-*O*-acetylesterase (Lse) (39), we cloned a novel cDNA from a pre B cell line library (70Z/3), in which the signal peptide encoding region of Lse in its 5′ region was substituted with a novel sequence (Fig. 1). Because the substituted region completely matches exon 1 of the Lse gene,[2] this new message seemed likely to be derived from an alternate promoter usage. The new cDNA is missing the original ATG codon of the Lse but has multiple in-frame ATG codons throughout the substituted 5′ region and the common sequence. Because the consensus sequence for endoplasmic reticulum targeting (the signal peptide) would be missing from the putative polypeptide encoded by this cDNA, we considered it a candidate for the previously described cytosolic Cse activity; this cDNA is hereafter called the Cse candidate (Cse-C).

Cse-C Codes for a Distinct Protein When Transiently Expressed in COS7 Cells—To determine the functional open reading frame of Cse-C, we created constructs for transient expression of Cse-C and the original Lse, each with a FLAG epitope tag incorporated into its C terminus. Anti-FLAG affinity chromatography was used to isolate epitope-tagged proteins from COS7 cells transiently transfected with either Lse-FLAG or Cse-C-FLAG. Production of polypeptides were studied by Western blotting using the anti-FLAG M2 antibody. Most of the Lse protein was recovered as a secreted form in the medium. This is in keeping with previous reports that a significant fraction of lysosomal enzymes are secreted into the medium when hyperexpressed in COS cells (38). In contrast, the Cse-C-FLAG product was mainly recovered in the cell lysate. Western blotting analysis of the anti-FLAG-Sepharose eluate showed that the secreted Lse was the expected uncleaved single band in the medium (the apparent molecular mass of 82 kDa is higher than that of the polypeptide backbone, presumably because of *N*-glycosylation and possibly *O*-glycosylation; see below). The Cse-C-FLAG product purified from cell lysates showed two bands of 51.5 and 28 kDa under reducing conditions (Fig. 2*A*). The smaller 28-kDa band showed a slower mobility in nonreducing conditions, indicating that the Cse-C-FLAG product can be cleaved by an endogenous proteinase to form heteromeric subunits (data not shown). Assuming that the 28-kDa fragment originated from the 51.5-kDa pro-polypeptide and accounting for the Kozak sequences (44) required for a functional starting codon (Fig. 2*B*), it is likely that the translation machinery in COS7 cells is using the Met-75 ATG downstream from the N terminus of the Lse open reading frame to create a protein missing the signal sequence and some of the N-terminal region (a total of 75 amino acids is truncated, compared with the mature Lse). The comparison between the Lse and putative Cse-C sequences is shown in Fig. 1*B*. The FLAG epitope-tagged forms of both molecules showed Sia 9-*O*-acetylesterase activity when affinity purified with anti-FLAG-Sepharose and assayed (data not shown). However, the Cse-C activity was rather unstable to routine elution with low pH. We therefore eluted the beads with a more gentle method, using the FLAG peptide itself (Fig. 2*C*). The activity recovered by this method was also not completely stable, falling by about half during 1 week of storage at 4 °C. In contrast, the Lse-FLAG product showed activity that was stable to acid elution, as well as prolonged storage at 4 °C (data not shown). These properties are in keeping with those previously reported for lysosomal and cytosolic Sia 9-*O*-acetylesterase activities from biological sources (see "Discussion"). Because of the small amount of the Cse-C-FLAG product and activity available from the cell lysate and the instability to storage, more detailed studies of its enzymatic activity were not pursued.

The Cse-C Form Localizes to the Cytosol of Transiently Transfected COS7 Cells—To confirm that the Cse-C encodes a cytosolic protein, we again transfected COS7 cells with the Lse-FLAG and Cse-C-FLAG constructs and detected the subcellular localization of the products by indirect immunofluorescent staining using the anti-FLAG M2 antibody. Staining of the Lse-FLAG product showed that the portion retained in the cells indeed gave a punctate pattern consistent with the lysosome/endosome/Golgi localization expected because of the signal sequence at the N terminus. In contrast, the Cse-C-FLAG product was detected diffusely throughout the cytosol with the subcellular organelles appearing as unstained regions (Fig. 3).

RT-PCR Analysis Indicates that the Two Forms of cDNA Have a Different Profile of Message Expression in Normal Mouse Tissues—Previously reported Northern blot analyses would not have distinguished between the expression of the two forms of esterase message in normal tissues. To examine the relative expression of the two forms in different tissues, we used sets of primers to selectively amplify the distinct 5′ sequences of Lse and Cse-C. This RT-PCR analysis shows that the expression of the two forms are quite different in various tissues of the mouse. Lse showed a widespread expression (Fig. 4*A*), suggesting that this message is derived from a general expression type of promoter. In contrast, the Cse-C showed rather tissue-specific expression with the strongest signal

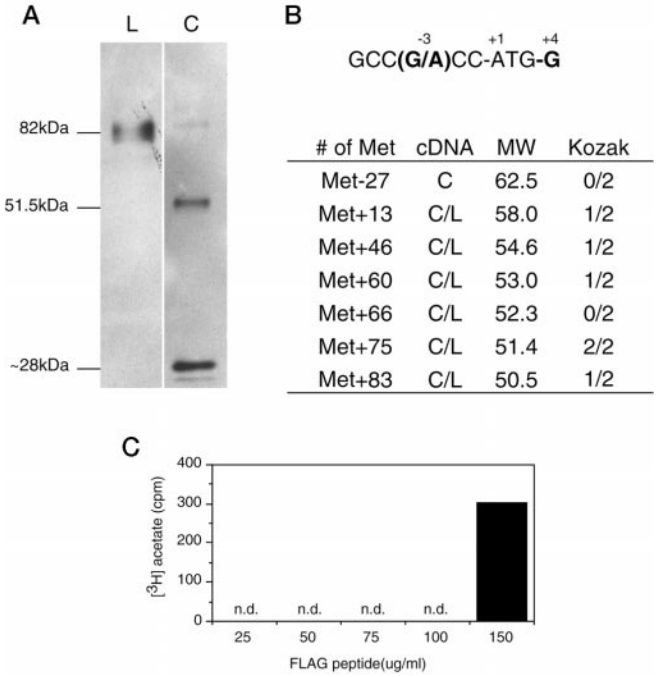
[2] H. Takematsu and A. Varki, unpublished data.



B

```
CSE cDNA   1                                                                                    cacacaacactggctgctcatatatcttcaggaacaccct
CSE cDNA  41   gagcaggctgtcacactgtgacatccaacacttcttcctgttggattccatcactgcctcggagcacagcttaaggctagtgggaaacctctcttctggacaaggtgcctggactttgtc

LSE AA   -23                                                        M  V  S  P  G  P  V  F  G  I  V  L  L  I  I  A  R  V  S  R  S  A  G
LSE cDNA   1                                   aatcaggatcttcacaaacatggtttccccggggcctgtgtttgggatagtgctgctcataatcgcgcgagtcagcagaagtgcag
CSE cDNA 161   caagaaaataaaagcaaggaaccatacatactgaatgaatgacacagacagtcactattacgggagctgagaatttggatgagcacaaagactttctgaagaaagctggcctgaagcaag

LSE AA     1    I  G  F  R  F  A  S  Y  I  D  N  Y  M  V  L  Q  K  E  P  S  G  A  V  I  W  G  F  G  T  P  G  A  T  V  T  V  T  L  C  Q
LSE cDNA  86   GTATTGGTTTTCGCTTTGCTTCATACATCGATAATTACATGGTGCTGCAGAAGGAGCCTTCAGGGGCTGTGATCTGGGGCTTTGGTACACCTGGAGCCACAGTGACAGTGACCTTGTGCC
CSE cDNA 281   GTATTGGTTTTCGCTTTGCTTCATACATCGATAATTACATGGTGCTGCAGAAGGAGCCTTCAGGGGCTGTGATCTGGGGCTTTGGTACACCTGGAGCCACAGTGACAGTGACCTTGTGCC

LSE AA    41    G  Q  E  T  I  M  K  K  V  T  S  V  K  E  P  S  N  T  W  M  V  V  L  D  P  M  K  P  G  G  P  F  E  V  M  A  Q  Q  T  L
LSE cDNA 165   AAGGTCAGGAAACCATCATGAAGAAAGTGACCAGTGTGAAAGAACCCTCCAACACCTGGATGGTGGTACTGGACCCTATGAAGCCTGGGGGACCTTTTGAAGTGATGGCGCAACAGACTT
CSE cDNA 401   AAGGTCAGGAAACCATCATGAAGAAAGTGACCAGTGTGAAAGAACCCTCCAACACCTGGATGGTGGTACTGGACCCTATGAAGCCTGGGGGACCTTTTGAAGTGATGGCGCAACAGACTT
CSE AA    75                                                                                                       M  A  Q  Q  T  L

LSE AA    81    G  T  M  N  F  T  L  R  V  H  D  V  L  F  G  D  V  W  L  C  S  G  Q  S  N  M  Q  M  T  V  S  Q  I  F  N  A  S  K  E  L
LSE cDNA 285   TGGGGACAATGAATTTCACCCTGAGAGTCCATGATGTCTTATTTGGAGATGTCTGGCTTTGCAGTGGGCAGAGTAACATGCAGATGACTGTTTCACAGATCTTTAACGCCTCAAAGGAGT
CSE cDNA 521   TGGGGACAATGAATTTCACCCTGAGAGTCCATGATGTCTTATTTGGAGATGTCTGGCTTTGCAGTGGGCAGAGTAACATGCAGATGACTGTTTCACAGATCTTTAACGCCTCAAAGGAGT
CSE AA    81    G  T  M  N  F  T  L  R  V  H  D  V  L  F  G  D  V  W  L  C  S  G  Q  S  N  M  Q  M  T  V  S  Q  I  F  N  A  S  K  E  L

LSE AA   121    S  D  T  A  A  Y  Q  S  V  R  I  F  S  V  S  L  I  Q  S  E  E  E  L  D  D  L  T  E  V  D  L  S  W  S  K  P  T  A  G  N
LSE cDNA 405   TGTCCGACACTGCTGCCTACCAGTCTGTGCGCATCTTCTCTGTCTCCCTCATCCAATCAGAGGAGGAGTTGGATGACCTTACTGAGGTTGACCTGTCATGGTCCAAGCCCACCGCAGGAA
CSE cDNA 641   TGTCCGACACTGCTGCCTACCAGTCTGTGCGCATCTTCTCTGTCTCCCTCATCCAATCAGAGGAGGAGTTGGATGACCTTACTGAGGTTGACCTGTCATGGTCCAAGCCCACCGCAGGAA
CSE AA   121    S  D  T  A  A  Y  Q  S  V  R  I  F  S  V  S  L  I  Q  S  E  E  E  L  D  D  L  T  E  V  D  L  S  W  S  K  P  T  A  G  N

LSE AA   161    L  G  H  G  N  F  T  Y  M  S  A  V  C  W  L  F  G  R  Y  L  Y  D  T  L  Q  Y  P  I  G  L  V  S  S  S  W  G  G  T  Y  I
LSE cDNA 525   ACTTGGGCCATGGAAATTTCACTTACATGTCTGCGGTGTGCTGGCTCTTTGGGCGTTACCTGTATGACACTCTGCAATATCCCATTGGGTTGGTCTCCTCCAGTTGGGGTGGGACGTACA
CSE cDNA 761   ACTTGGGCCATGGAAATTTCACTTACATGTCTGCGGTGTGCTGGCTCTTTGGGCGTTACCTGTATGACACTCTGCAATATCCCATTGGGTTGGTCTCCTCCAGTTGGGGTGGGACGTACA
CSE AA   161    L  G  H  G  N  F  T  Y  M  S  A  V  C  W  L  F  G  R  Y  L  Y  D  T  L  Q  Y  P  I  G  L  V  S  S  S  W  G  G  T  Y  I

LSE AA   201    E  V  W  S  S  R  R  T  L  K  A  C  G  V  P  N  T  R  D  E  R  V  G  Q  P  E  I  K  P  M  R  N  E  C  N  S  E  E  S  S
LSE cDNA 645   TTGAAGTCTGGTCATCTAGAAGGACACTGAAAGCCTGTGGAGTCCCTAATACAAGGGATGAAAGAGTTGGTCAGCCCGAGATAAAGCCCATGAGGAATGAATGTAACAGTGAGGAGTCCT
CSE cDNA 881   TTGAAGTCTGGTCATCTAGAAGGACACTGAAAGCCTGTGGAGTCCCTAATACAAGGGATGAAAGAGTTGGTCAGCCCGAGATAAAGCCCATGAGGAATGAATGTAACAGTGAGGAGTCCT
CSE AA   201    E  V  W  S  S  R  R  T  L  K  A  C  G  V  P  N  T  R  D  E  R  V  G  Q  P  E  I  K  P  M  R  N  E  C  N  S  E  E  S  S

LSE AA   241    C  P  F  R  V  V  P  S  V  R  V  T  G  P  T  R  H  S  V  L  W  N  A  M  I  H  P  L  Q  N  M  T  L  K  G  V  V  W  Y  Q
LSE cDNA 806   CATGTCCTTTTAGGGTTGTCCCATCTGTTCGTGTGACTGGACCGACAAGACACTCCGTCCTTTGGAACGCCATGATCCATCCACTGCAAAACATGACTTTGAAAGGGGTGGTGTGGTACC
CSE cDNA1001   CATGTCCTTTTAGGGTTGTCCCATCTGTTCGTGTGACTGGACCGACAAGACACTCCGTCCTTTGGAACGCCATGATCCATCCACTGCAAAACATGACTTTGAAAGGGGTGGTGTGGTACC
CSE AA   241    C  P  F  R  V  V  P  S  V  R  V  T  G  P  T  R  H  S  V  L  W  N  A  M  I  H  P  L  Q  N  M  T  L  K  G  V  V  W  Y  Q

LSE AA   281    G  E  S  N  A  D  Y  N  R  D  L  Y  T  C  M  F  P  E  L  I  E  D  W  R  Q  T  F  H  Y  G  S  Q  G  Q  T  D  R  F  F  P
LSE cDNA 885   AGGGAGAGTCCAATGCGGATTATAACAGGGACCTGTACACCTGCATGTTTCCTGAACTCATTGAAGACTGGCGCCAGACCTTCCACTACGGCTCCCAGGGTCAGACAGATCGTTTCTTTC
CSE cDNA1121   AGGGAGAGTCCAATGCGGATTATAACAGGGACCTGTACACCTGCATGTTTCCTGAACTCATTGAAGACTGGCGCCAGACCTTCCACTACGGCTCCCAGGGTCAGACAGATCGTTTCTTTC
CSE AA   281    G  E  S  N  A  D  Y  N  R  D  L  Y  T  C  M  F  P  E  L  I  E  D  W  R  Q  T  F  H  Y  G  S  Q  G  Q  T  D  R  F  F  P

LSE AA   321    F  G  F  V  Q  L  S  S  Y  M  L  K  N  S  S  D  Y  G  F  P  E  I  R  W  H  Q  T  A  D  F  G  H  V  P  N  P  K  M  P  N
LSE cDNA1005   CATTTGGGTTTGTCCAGCTGTCTTCATACATGTTAAAGAACTCATCGGATTATGGATTTCCAGAGATCCGTTGGCATCAAACTGCAGACTTTGGCCATGTCCCCAACCCGAAGATGCCTA
CSE cDNA1241   CATTTGGGTTTGTCCAGCTGTCTTCATACATGTTAAAGAACTCATCGGATTATGGATTTCCAGAGATCCGTTGGCATCAAACTGCAGACTTTGGCCATGTCCCCAACCCGAAGATGCCTA
CSE AA   321    F  G  F  V  Q  L  S  S  Y  M  L  K  N  S  S  D  Y  G  F  P  E  I  R  W  H  Q  T  A  D  F  G  H  V  P  N  P  K  M  P  N

LSE AA   361    T  F  M  A  V  A  I  D  L  C  D  R  D  S  P  F  G  S  I  H  P  R  D  K  Q  T  V  A  Y  R  L  H  L  G  A  R  A  V  A  Y
LSE cDNA1125   ATACTTTTATGGCTGTCGCTATTGACCTCTGTGATAGAGACTCACCCTTTGGAAGTATCCACCCTCGAGACAAACAGACTGTGGCCTATCGTCTGCACTTGGGTGCTCGAGCTGTGGCAT
CSE cDNA1361   ATACTTTTATGGCTGTCGCTATTGACCTCTGTGATAGAGACTCACCCTTTGGAAGTATCCACCCTCGAGACAAACAGACTGTGGCCTATCGTCTGCACTTGGGTGCTCGAGCTGTGGCAT
CSE AA   361    T  F  M  A  V  A  I  D  L  C  D  R  D  S  P  F  G  S  I  H  P  R  D  K  Q  T  V  A  Y  R  L  H  L  G  A  R  A  V  A  Y

LSE AA   401    G  E  K  N  L  T  F  Q  G  P  L  P  K  K  I  E  L  L  A  S  N  G  L  L  N  L  T  Y  D  Q  E  I  Q  V  Q  M  Q  D  N  K
LSE cDNA1245   ATGGTGAGAAAAATTTGACCTTTCAAGGACCACTACCTAAGAAGATAGAACTCTTGGCTTCCAATGGCCTGCTCAACCTCACATATGACCAGGAAATCCAGGTGCAAATGCAAGATAACA
CSE cDNA1481   ATGGTGAGAAAAATTTGACCTTTCAAGGACCACTACCTAAGAAGATAGAACTCTTGGCTTCCAATGGCCTGCTCAACCTCACATATGACCAGGAAATCCAGGTGCAAATGCAAGATAACA
CSE AA   401    G  E  K  N  L  T  F  Q  G  P  L  P  K  K  I  E  L  L  A  S  N  G  L  L  N  L  T  Y  D  Q  E  I  Q  V  Q  M  Q  D  N  K

LSE AA   441    T  F  E  I  S  C  C  S  D  R  H  C  K  W  L  P  A  P  V  N  T  F  S  T  Q  T  L  I  L  D  L  N  A  C  L  S  T  V  V  A
LSE cDNA1365   AAACATTTGAGATCTCCTGTTGTAGTGACCGTCATTGCAAATGGCTTCCAGCTCCTGTTAACACTTTCTCCACCCAGACTCTGATTTTGGATCTCAATGCTTGTCTTAGCACAGTGGTTG
CSE cDNA1601   AAACATTTGAGATCTCCTGTTGTAGTGACCGTCATTGCAAATGGCTTCCAGCTCCTGTTAACACTTTCTCCACCCAGACTCTGATTTTGGATCTCAATGCTTGTCTTGGCACAGTGGTTG
CSE AA   441    T  F  E  I  S  C  C  S  D  R  H  C  K  W  L  P  A  P  V  N  T  F  S  T  Q  T  L  I  L  D  L  N  A  C  L  S  T  V  V  A

LSE AA   481    V  R  Y  A  W  T  T  W  P  C  E  Y  K  Q  C  A  V  Y  H  T  S  S  M  L  P  A  P  P  F  I  A  Q  I  S  H  R  G  I  *
LSE cDNA1485   CTGTTCGCTATGCCTGGACCACATGGCCTTGTGAATATAAGCAATGTGCTGTTTACCACACCAGCAGTATGTTGCCAGCTCCTCCCTTCATTGCTCAAATTTCACACAGGGGTATCTAAG
CSE cDNA1721   CTGTTCGCTATGCCTGGACCACATGGCCTTGTGAATATAAGCAATGTGCTGTTTACCACACCAGCAGTATGTTGCCAGCTCCTCCCTTCATTGCTCAAATTTCACACAGGGGTATCTAAG
CSE AA   481    V  R  Y  A  W  T  T  W  P  C  E  Y  K  Q  C  A  V  Y  H  T  S  S  M  L  P  A  P  P  F  I  A  Q  I  S  H  R  G  I  *
```

FIG. 1. **Differing sequences at the 5′ ends of cDNAs derived from the Lse gene.** *A*, schematic of the cDNA structures of the Lse and the putative Cse clones. Noncoding regions of Lse cDNA are indicated by *black lines*. The previously reported Lse enzyme purified from rat liver (36) was a heterodimer of two polypeptides as depicted by the differently patterned *boxes*. Proteolytic cleavage takes place to a variable extent between the differently patterned *boxes* during Lse maturation in the endosomes/lysosomes. The *white box* in Cse-C indicates the sequence that replaces



# of Met	cDNA	MW	Kozak
Met-27	C	62.5	0/2
Met+13	C/L	58.0	1/2
Met+46	C/L	54.6	1/2
Met+60	C/L	53.0	1/2
Met+66	C/L	52.3	0/2
Met+75	C/L	51.4	2/2
Met+83	C/L	50.5	1/2

FIG. 2. **Epitope-tagged products of the Lse and Cse-C cDNAs expressed in COS cells.** *A*, Western blot of affinity purified Lse and Cse-C from transfected COS7 cells. Lse-FLAG (*L*) was purified from the culture medium and Cse-C-FLAG (*C*) from cell lysates of transfected COS7 cells as described under "Experimental Procedures." The Cse-C-FLAG construct contains the exact sequences shown in Fig. 1*B*. Purified proteins are subjected to SDS-PAGE under reducing conditions and detected by immunoblot with an anti-FLAG antibody after transfer to a nitrocellulose membrane. The molecular masses of bands are indicated by the *arrows*. The relative amounts of the 51.5- and 28-kDa bands varied in different preparations. *B*, Kozak concensus sequences of putative start sites in the Cse-C construct, matched together with expected molecular masses. The consensus sequence for translational initiation is shown above. The *chart* lists the positions of potential initiation codon (Met residue), the type of the cDNA where the potential ATG codons exist (cDNA), putative molecular mass of each translation product, and matching numbers to the Kozak sequence in position −3 and +4 residues of each ATG codon (44). The most likely start codon is at Met^{75}, which is the only one that shares the two most conserved residues of the Kozak sequence and fits the molecular mass of the product found in the COS7 cells. *C*, COS7 cells were transfected with Cse-C-FLAG vector and cell lysates incubated with M2 anti-FLAG affinity gel as described under "Experimental Procedures." The gel was then packed into a column and eluted stepwise with increasing concentration of the FLAG peptide in Tris-buffered saline. 50-μl aliquots of each fraction was assayed for 9-*O*-acetylesterase activity with [9-*O*-acetyl-^{3}H]Neu5,9Ac_2 as described under "Experimental Procedures." *n.d.* indicates that the activity was not detected.

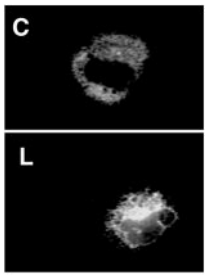


FIG. 3. **Indirect immunofluorescence staining of transfected COS cells.** COS7 cells were transiently transfected with Lse-FLAG (*L*) or Cse-C-FLAG (*C*) constructs. Subcellular localization was identified by immunofluorescent staining using a monoclonal antibody against the FLAG epitope tag as described under "Experimental Procedures." The Lse construct gives mainly a perinuclear Golgi staining pattern, along with punctate staining typical of lysosomes and endosomes. In contrast, the Cse-C construct gave a widespread diffuse staining pattern localized to the cytosol.

found in the brain and ovary and some detectable signal in the liver and thymus (Fig. 4*B*). Interestingly, multiple bands were sometimes observed from the Cse-C amplification, which were especially strong in the brain. Some of the major bands appearing in Cse-C amplification were directly sequenced. This analysis showed that only liver has the exact sequence expected from 7A3-C clone and that there can be multiple populations of 5′ noncoding sequences attached to the beginning of universal Lse exon 2 (Fig. 4*B*). In other tissues, the dominant signal amplified has truncated 7A3-C sequence and corresponds to the 7A3-D clone shown in 5′-rapid amplification of cDNA ends rapid amplification of cDNA ends studies from the 70Z/3 cell line (39). This indicates that alternative splicing events are occurring in the processing of this message. However, none of the other form sequences have an alternative open reading frame to code any new type of translational product. Thus, despite the multiple 5′ variations found in the new cDNAs, the protein coded by all these messages is expected to be the same. It is of course possible that these different 5′ noncoding regions may affect the efficiency of translation and/or alter the stability of the message. Also, because we found variable types of message in the brain, there remains a possibility of other as yet unidentified type of messages derived from independent promoters and/or splicing that might not be detected by this RT-PCR experiment.

A Chicken IgY Directed against the Esterase C Terminus—We had previously reported that a rabbit polyclonal antibody that reacted with rat liver Lse did not cross-react with Cse activity in the same tissue. However, this antibody was produced by a single rabbit out of many that failed to respond. We assume that this difficulty in raising a mammalian antibody response against the intact Lse protein is because of a high degree of conservation of the protein. We therefore raised polyclonal antibodies in chickens, directed against a 69-amino acid region that is common to the Lse and the Cse protein (Fig. 1*B*, amino acids 381–450). This C-terminal peptide was expressed in a bacterial system as a maltose-binding fusion protein, purified as described, and used to immunize chickens. The IgY was purified from egg yolk and studied for reaction by Western blot analysis with the recombinant intact Lse or with freeze-thaw extracts of tissues. One of the positive chicken antibodies (C-12) was found to specifically detect bands in a similar pattern in both mouse and rat tissues using the freeze-thaw supernatant fraction that is known to have both the Lse and Cse activities (data not shown; no bands were seen with control chicken IgY preparations under the same conditions of staining).

Cse-C Polypeptides Correspond to Major Activity for Cytosolic Sialic Acid 9-O-Acetylesterase in Buffer-extracted Mouse Liver—Lse and Cse from the rat liver were previously reported as 9-*O*-acetylesterase activities that bound to ConA or DEAE, respectively (36). To correlate these activities with the Lse gene-derived polypeptides reactive with the chicken antibody, we studied mouse liver. The tissue was homogenized in the

the Lse exon 1. The Cse-C message is expected to encode a single polypeptide. *B*, primary sequences and proposed open reading frames of the 5′ ends of the Lse and the putative Cse (7A3-C clone). The translational initiation of Cse-C sequence is depicted here as starting at Met^{75} (see Fig. 2 for other possibilities). The *underlined* nucleotide sequence shows the Cse-C specific 5′ sequence. *Underlined* peptide sequences indicate the polypeptide residues used for immunization of chickens for IgY preparation. *Arrows* in the Lse amino acid sequence indicate the N termini of the small and large subunit, respectively.

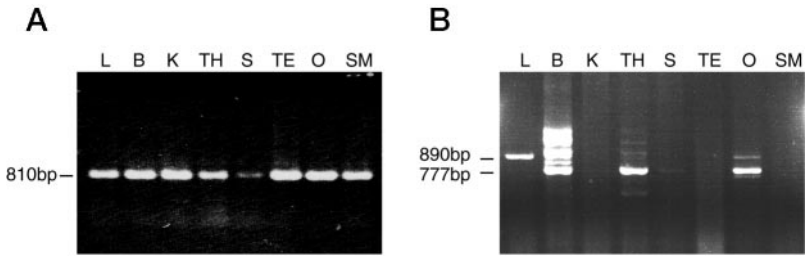


FIG. 4. **RT-PCR analysis of expression of Lse and Cse-C messages in mouse tissues.** As described under "Experimental Procedures," the RT-PCR method was applied using primers specific for the 5′ ends of the Lse (*A*) and Cse-C (*B*) messages. Total cellular RNA was purified from various mouse tissues including liver (*L*), brain (*B*), kidney (*K*), thymus (*TH*), spleen (*S*), testis (*TE*), ovary (*O*), and skeletal muscle (*SM*). First strand cDNA derived from each source was subjected to PCR by the primer set, which only amplifies each form of message. PCR products were detected by ethidium bromide staining after 1.2% agarose gel electrophoresis. The sizes of Lse bands were 810 bp as expected in contrast to the Cse-C that the expected size is 890 bp, which appeared only in the liver. In other tissue, Cse-C gave 777 bp amplification, which is the size of alternate Cse-C type clone originally identified as clone 7A3-D (39) in 5′-rapid amplification of cDNA ends experiment in 70Z/3 cell line. For the ovary sample shown in *panel B*, one-fifth of the template amount was used.

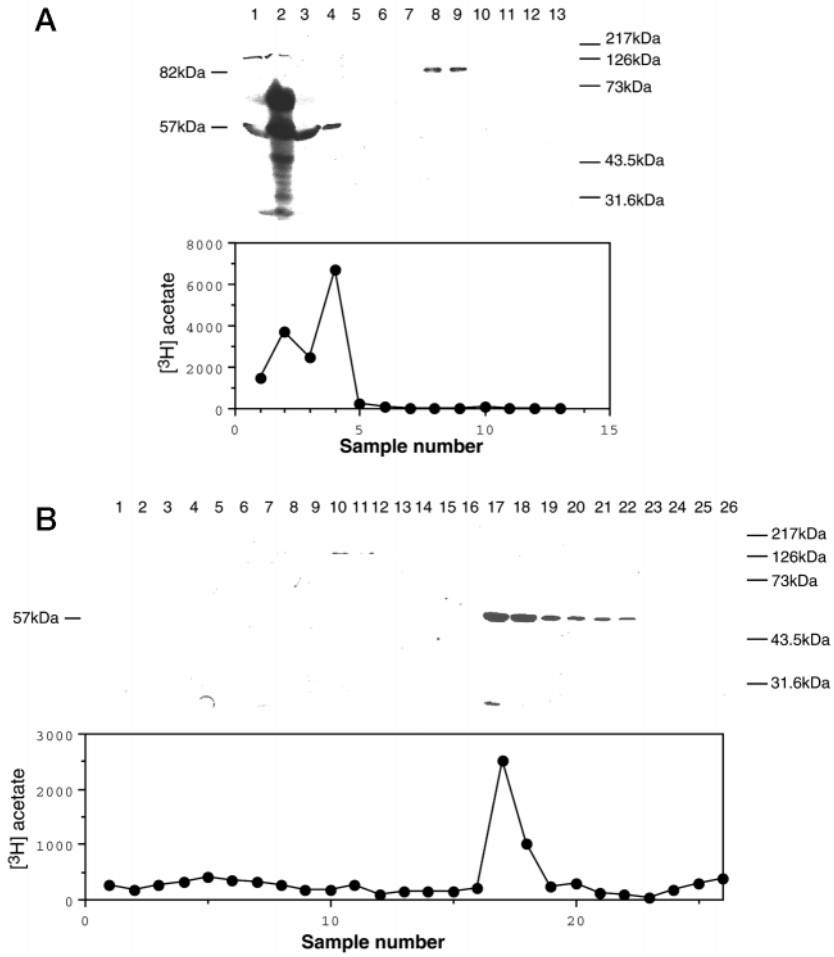


FIG. 5. **Western blot and activity analysis of chromatography fractions from mouse liver extract.** Freshly obtained mouse liver was extracted as described under under "Experimental Procedures" and subjected first to ConA-Sepharose affinity chromatography. The run-through fraction from the ConA column was then applied to a DEAE-cellulose (DE52) column that was washed and eluted with a linear gradient of NaCl, as described under "Experimental Procedures." Aliquots from the fractions were assayed for Sia 9-O-acetylesterase activity ([^{3}H]acetate release) or by SDS-PAGE under reducing conditions, followed by Western blotting with the chicken IgY antibody directed against a 69-amino acid sequence shared by the Lse and Cse-C cDNAs. Molecular mass standard markers are shown to the *right*, and major signal size are shown to the *left* of the blot.

buffer with a Polytron after two freeze-thaw cycles, and the homogenates were then ultracentrifuged to collect a fraction that should contain both the Cse and the Lse (the latter being released from disrupted lysosomes). The extracts were first applied to the ConA-Sepharose column and eluted with α-methyl-mannoside (Lse activity should bind to the column and be eluted). The flow-through fraction from ConA-Sepharose was then applied to the DEAE-cellulose column and then eluted with a NaCl gradient (Cse should bind and be eluted). Column fractions were assayed for Sia 9-*O*-acetylesterase activity and also studied by Western blotting using the C-12 Chicken IgY directed against the shared 69-amino acid region of both enzymes.

As shown in Fig. 5*A*, Western blotting with this antibody indeed gave the expected signal for the Lse polypeptide in the ConA-eluted fraction of the liver extract (~82 kDa). However, Sia 9-*O*-acetylesterase activity was not easily detectable in these fractions, probably because of the much lesser extent of Lse expression in mouse liver when compared with rat liver (from where we originally purified the enzyme). This is also consistent with the observation in transfected COS7 cells that Western blotting gave a much higher sensitivity than 9-*O*-acetylesterase activity assay. We therefore carried out a similar study on rat liver extracts, where the Lse forms a much higher proportion of total 9-*O*-acetylesterase activity (36). Indeed, as expected we found that the ConA eluted fractions from rat liver extracts contained a similar ~82-kDa band that coincided with the esterase activity assay (data not shown).

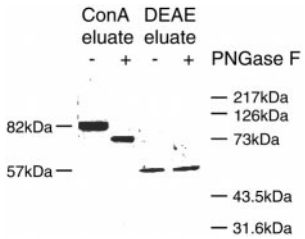


FIG. 6. **Western blotting analysis of PNGase F-digested Lse and Cse-C fractions.** Aliquots of fractions obtained in the experiment shown in Fig. 5 (sample tube 8 from the ConA elution and 17 from the DEAE elution) were treated or sham-treated with PNGase F to remove *N*-glycans, then subjected to SDS-PAGE under reducing conditions, and Western blotted with the chicken IgY antibody directed against a 69-amino acid sequence shared by the Lse and Cse-C cDNAs. Molecular mass standard markers are shown to the *right* of the blot.

The ConA run-through fractions from the mouse liver extract contained esterase activity which was maximal in sample tube 4 rather than in the peak sample tube 2. This is possibly because of nonspecific inhibition by the large amount of extract protein in fraction 2 (this may also account for the large number of nonspecific bands in the Western blot of this fraction, which did not persist in the following step). A pool of the ConA run-through (sample tubes 1–4) was applied to a DEAE column, which was eluted with increasing concentrations of NaCl. The major peak of esterase activity coincided with the elution of single 57-kDa band on Western blotting (Fig. 5*B*). Because the FLAG epitope-tagged protein studied above showed weak but detectable activity and encoded a protein expressed in the cytosol, it is likely that the 57-kDa polypeptide is responsible for the buffer-extracted Cse activity in mouse liver. The minor difference in the size may be caused by an unknown protein modification of the cytosolic protein. Alternatively, a different ATG start codon might be used in mouse liver than in COS cells.

Polypeptides derived from the Lse message (with a signal peptide encoding sequence) are expected to carry *N*-glycans, whereas those derived from the Cse-C message (or others without a signal peptide encoding sequence) should not. To examine this issue, we studied the fractions obtained above by Western blotting following PNGase F digestion to release *N*-glycans (Fig. 6). As predicted, PNGase F treatment of the ConA-eluted Lse band caused a shift of the ConA eluted 82-kDa band to 72 kDa (because the expected size of the Core polypeptide is 61 kDa, this suggests the presence of other modifications, such as *O*-glycosylation). In contrast to the result with the ConA bound fraction, there was no effect of PNGase F on the putative Cse-C products (DE52 eluted fraction, Fig. 6). This further confirms that Cse-C-related message products are indeed cytosolic proteins without *N*-glycosylation.

DISCUSSION

Based on the data in this study, we propose that a single Sia 9-*O*-acetylesterase gene can encode two differently localized proteins by differential usage of the signal peptide encoding exon found at the N terminus. To date, there are relatively few such examples wherein a single gene can produce polypeptides targetted to different cell compartments (45). This was first demonstrated in the yeast invertase (SUC2) gene (46, 47). A mammalian gene for the actin-modulating protein Gelsolin uses an alternative promoter to derive plasma and cytosolic forms message with and without signal peptide sequence, respectively (48). Recently, interleukin-15 was also shown to target to different compartments. In this case, the interleukin-15 gene encodes two different transcripts with or without the full-length signal sequence peptide by means of alternative splicing, and the products with the incomplete signal sequences were shown to be targetted to the cytosol instead of being secreted out from the cells (49). However, in all examples, the functional relevance of the "alternative form" are unknown (*i.e.* cytosolic invertase, plasma form of Gelsolin, cytosolic form of interleukin-15 in cytosol). In the case of the Sia 9-*O*-acetylesterase gene, both the Lse and Cse have potential substrates in their respective locations. Moreover, Lse has an additional localization signal for targetting to the lysosome, apparently using mannose 6-phosphorylation.[3]

The only difference between the two messages is in the 5′ terminus, and the 5′-rapid amplification of cDNA ends product of the Lse-type message did not show any additional sequences in its 5′ region. Taken together with the differences in tissue-specific expression, it seems likely that the Lse and Cse use different promoters for their expression. In this regard, we do have some information concerning the genomic organization of exon 1 of Lse and exon 2 of both forms. We know that in between these exons, there is no exon for Cse detected by Southern hybridization.[2] Thus, the Cse promoter and transcription initiation site must be somewhere in the more 5′ region of the esterase gene, upstream of the first Lse exon, which encodes its signal peptide. To definitively prove differential promoter usage and to completely rule out any unusual form of alternative splicing, we need to clone and characterize the entire genomic organization of the esterase gene, together with a complete promoter analysis.

Using Western blot analysis with an antibody against the common C terminus, we were able to differentiate the activity of the Lse and Cse forms, because the two proteins behaved in a distinct way by column chromatography. In retrospect, characterization of the two enzymes had already shown some shared features other than the fact that both have a common substrate, 9-*O*-acetyl Sia. Both activities are inhibited by diisopropyl fluorophosphate treatment, suggesting a serine-active site in each. The two activities also have a neutral pH optimum (32, 34, 36). In regard to the last point, it is notable that Lse was shown to co-localize with conventional "acid hydrolases" by an EM study (34). This suggests that the Lse-catalyzed reaction may takes place in the early endosome where the pH of the compartment is closer to neutral. Because lysosomal degradation of the oligosaccharides takes place from the outer terminii, it could be that the enzymes cleaving these act in advance of the other exoglycosidases, whereas the pH of the compartment is still not too acidic, making lysosomal hydrolysis more efficient. An alternate possibility we had suggested earlier is that the pH of lysosomes might not always be acidic but rather fluctuates between neutral and acidic states (34).

In contrast to the Lse, the Cse has been proposed as a "recycling" enzyme acting on the free cytosolic pool of 9-*O*-acetylated Sia. Because *O*-acetylated forms of Sia are poor substrates for CMP-NeuAc synthase, for some of the sialyltransferases, and for acylneuraminic acid pyruvate lyase, the presence of Cse would assure a higher efficiency in the recycling of Sias. In this regard it is interesting that the Cse activity is very high in the brain, where 9-*O*-acetylated Sias are commonly expressed. Of course, for this model to work, other Sia-recognizing proteins such as the lysosomal sialidases and the lysosomal Sia exporter should not be too sensitive to the presence of *O*-acetyl groups. This matter has not yet been explored.

The substrate specificities for the Lse and Cse might also be subtly different. Purified rat liver Lse was shown to cleave *O*-acetyl groups only from the 9-position of Sias (36). In contrast, Schauer and colleagues (32) reported that a bovine brain Cse showed activity toward the 4-*O*-acetylated Sia as well.

[3] S. Diaz and A. Varki, unpublished data.

Because their bovine Cse was not purified to homogeneity, it is not possible to be certain that the latter activity was derived from the same polypeptide. However, it is interesting to speculate that a Cse lacking the N terminus might have relaxed substrate specificity relative to the Lse by assuming that brain Cse is also coded by the Cse-C form of this gene.

The recombinant Cse-C needed to be exposed to low pH for a brief moment during acid elution from the anti-FLAG resin, and this may account for the lower recovery of activity compared with that obtained with the more expensive method, elution using the FLAG peptide. Even with the latter elution method, the purified material lost activity during further storage. This finding is actually in keeping with previous observations that naturally occurring Cse is not a stable enzyme (in contrast to Lse, which is very stable both *in vivo* and *in vitro*). It is noteworthy that Cse-C lacks a single cysteine residue in its primary amino acid sequence compared with Lse. This can be another explanation for the difference in its stability. Using COS7 cell transfection experiments, we could not obtain enough recombinant Cse protein to examine all the details of its activity. Ultimately, purification, characterization, and crystallization of those two enzymes in the presence of the various substrates will be required to fully compare their activities.

For Lse, it was previously suggested that a naturally occurring lysosomal cleavage of this polypeptide might expose the activity (33). However, we found the active nonclipped form both in the medium of transiently transfected COS7 and in a rat liver extract. Thus, it is fair to conclude that the cleavage of the Lse is not essential for exposing the activity (although we cannot rule out a further increase in activity upon cleavage). We found two bands of Cse signals in the COS7 transfection experiment. Because we did not detect 28-kDa signal in the mouse liver, it is likely that the 28-kDa fragment represents proteolytic cleaved material derived from the 51-kDa fragment. But it has not yet been possible to identify which form is active in the COS7 cells. Since the 28-kDa band is dramatically reduced in amount in the Western blot under nonreducing conditions, it is likely that it is associated with an N-terminal subunit through an S-S bond as also found in Lse, even though the precise site of cleavage appears to be different. This can be explained by the fact that the two proteins are processed in different cell compartments.

In the mouse liver, we were not able to recover activity of the Lse, even though the Lse message is strongly expressed in the RT-PCR experiment, and an immunoreactive polypeptide was detectable. This suggests the possibility that, as in the COS cells, the bulk of mouse liver Lse might be secreted. This may also explain why our initial attempt to produce antibodies against purified rat Lse in rabbits was not very successful (33). The only one among many immunized rabbit sera that showed anti-Lse activity had an inhibitory effect for the Lse activity but not for Cse. This is probably because that particular polyclonal antibody was directed against a very limited epitope that was unique for Lse. We used a chicken for immunization to raise an antibody because genomic Southern blot cross-hybridization of mouse full-length Lse cDNA was a lot weaker with chicken genomic DNA compared with DNA from mammalian species.[2] The availability of this antibody reactive to primary amino acid sequence made it possible for us to study the translated polypeptides of both messages after using ConA and DEAE columns to physically separate them. Of course, although it is very likely that the Cse-C is encoding the cytosolic activity found in the mouse liver, it remains possible that the antibody is cross-reacting with a product from a closely related gene. The other unlikely possibility is proteolytic cleavage of the Lse followed by de-*N*-glycosylation and export to the cytosol (such proteins are normally rapidly degraded by proteasomes).

PNGase F digestions confirmed that only the ConA binding Lse band not the Cse-C band carries *N*-glycans, further supporting the cytosolic location of the latter. In contrast to the extensive cleavage of the Cse (51–28 kDa) that we saw in COS cells, buffer-extracted Cse activity in mouse liver coincided primarily with the of 57-kDa polypeptide without cleaved material. The difference in size of the COS7-derived band (51 kDa) and that from the mouse liver (57 kDa) might indicate an unknown post-translational modification of this enzyme in different tissues/systems, because we used the "liver form" clone for the transient COS7 expression experiment. Alternatively, since there are multiple in-frame ATG codons throughout this cDNA (Fig. 2*B*), it is possible that these two cell types are using a different start codon for translational initiation.

From these data we conclude that at least some portion of the previously reported Cse activity can be explained by the alternate cytosolic form derived from the Lse gene. However, our ongoing studies (data not shown) suggest that in tissues other than liver (*e.g.* brain), there are some cytosolic 9-*O*-acetylesterase activity peaks that do not coincide with the immunoreactive bands. Thus, it is probable that there are one or more other genes that can generate Cse activity in other tissues. Furthermore, when mouse liver was extracted in the presence of detergent (0.1% Triton X-100), additional immunoreactive bands were detected that did not coincide with peaks of Sia 9-*O*-acetylesterase activity (data not shown). It is apparent that there are even greater complexities in the forms of message and polypeptide that can be derived from the Lse gene. Further studies are needed to sort out all these complexities.

With regard to the presence of a Cse-like activity previously reported in red blood cells (50), the significance is not really clear, because there should be no turnover of Sias in these cells. It is possible that this Cse has an alternative role in such cells or that it is left over from the earlier phase of red cell development. Even though the molecular mechanisms and topology are not easy to explain, it is noteworthy that expression of a cytosolic sialidase resulted in a reduced amount of intercellular G_{M3} expression and increased amounts of its hydrolyzed product lactosyl ceramide in stably transfected B16 melanoma cells (51) and epidermoid carcinoma cells (52). If a similar situation applies for *O*-acetylated gangliosides, it is possible that this cytosolic enzyme can still regulate *O*-acetylated ganglioside expression in certain cell types.

Expression of the *O*-acetylation on Sia is regulated in temporal and spatial patterns during the embryogenesis and organogenesis (2). This modification is also known to be selectively expressed on certain underlying sialoconjugate epitopes (30). These studies imply that this modification is playing a role in the specific cell-cell recognition events during the developmental process. To study the mechanism and significance of regulation of Sia *O*-acetylation in the biological systems, information regarding all the metabolizing steps is essential. In the present study, we report the cDNA cloning and the regulation of biosynthetic mechanism of lysosomal and cytosolic forms of *O*-acetylation hydrolyzing enzymes. Further studies should be pursued to examine how these factors are affecting a given cell systems through the hydrolysis of *O*-acetyl groups on Sias. Ultimately, to assign a definitive function, the *O*-acetylation profile and phenotype resulting from *in vivo* genetic disruption of the Cse and Lse activities needs to be studied.

Acknowledgments—We thank Dr. Nissi Varki for help with immunofluorescent staining of the COS7 cells and Dr. Lubor Borsig for critical discussions throughout this study.

REFERENCES

1. Schauer, R. (1982) *Sialic Acids: Chemistry, Metabolism and Function, Cell Biology Monographs,* Vol. 10, Springer-Verlag, New York
2. Varki, A. (1992) *Glycobiology* **2,** 25–40
3. Varki, A. (1997) *FASEB J.* **11,** 248–255
4. Kansas, G. S. (1996) *Blood* **88,** 3259–3287
5. McEver, R. P., and Cummings, R. D. (1997) *J. Clin. Invest.* **100,** 485–492
6. Varki, A. (1997) *J. Clin. Invest.* **99,** 158–162
7. Crocker, P. R., Clark, E. A., Filbin, M., Gordon, S., Jones, Y., Kehrl, J. H., Kelm, S., Le Douarin, N., Powell, L., Roder, J., Schnaar, R. L., Sgroi, D. C., Stamenkovic, K., Schauer, R., Schachner, M., Van den Berg, T. K., Van der Merwe, P. A., Watt, S. M., and Varki, A. (1998) *Glycobiology* **8**, v
8. Powell, L. D., and Varki, A. (1995) *J. Biol. Chem.* **270,** 14243–14246
9. Kelm, S., and Schauer, R. (1997) *Int. Rev. Cytol.* **175,** 137–240
10. Cornish, A. L., Freeman, S., Forbes, G., Ni, J., Zhang, M., Cepeda, M., Gentz, R., Augustus, M., Carter, K. C., and Crocker, P. R. (1998) *Blood* **92,** 2123–2132
11. Herrler, G., Rott, R., Klenk, H. D., Muller, H. P., Shukla, A. K., and Schauer, R. (1985) *EMBO J.* **4,** 1503–1506
12. Miller-Podraza, H., Abul Milh, M., Teneberg, S., and Karlsson, K. A. (1997) *Infect. Immun.* **65,** 2480–2482
13. Miller-Podraza, H., Bergstrom, J., Milh, M. A., and Karlsson, K. A. (1997) *Glycoconj. J.* **14,** 467–471
14. Johansson, L., and Karlsson, K. A. (1998) *Glycoconj. J.* **15,** 713–721
15. Schauer, R. (1970) *Hoppe. Seylers. Z. Physiol. Chem.* **351,** 749–758
16. Schauer, R., and Wember, M. (1971) *Hoppe-Seyler's Z. Physiol. Chem.* **352,** 1282–1290
17. Schlosshauer, B., Blum, A. S., Mendez Otero, R., Barnstable, C. J., and Constantine-Paton, M. (1988) *J. Neurosci.* **8,** 580–592
18. Sparrow, J. R., and Barnstable, C. J. (1988) *J. Neurosci.* **21,** 398–409
19. Blum, A. S., and Barnstable, C. J. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 8716–8720
20. Sjoberg, E. R., Powell, L. D., Klein, A., and Varki, A. (1994) *J. Cell Biol.* **126,** 549–562
21. Collins, B. E., Yang, L. J. S., Mukhopadhyay, G., Filbin, M. T., Kiso, M., Hasegawa, A., and Schnaar, R. L. (1997) *J. Biol. Chem.* **272,** 1248–1255
22. Collins, B. E., Kiso, M., Hasegawa, A., Tropak, M. B., Roder, J. C., Crocker, P. R., and Schnaar, R. L. (1997) *J. Biol. Chem.* **272,** 16889–16895
23. Kelm, S., Brossmer, R., Isecke, R., Gross, H. J., Strenge, K., and Schauer, R. (1998) *Eur. J. Biochem.* **255,** 663–672
24. Strenge, K., Schauer, R., Bovin, N., Hasegawa, A., Ishida, H., Kiso, M., and Kelm, S. (1998) *Eur. J. Biochem.* **258,** 677–685
25. Shi, W. X., Chammas, R., Varki, N. M., Powell, L., and Varki, A. (1996) *J. Biol. Chem.* **271,** 31526–31532
26. Muchmore, E., and Varki, A. (1987) *Science* **236,** 1293–1295
27. Varki, A., Hooshmand, F., Diaz, S., Varki, N. M., and Hedrick, S. M. (1991) *Cell* **65,** 65–74
28. Diaz, S., Higa, H. H., Hayes, B. K., and Varki, A. (1989) *J. Biol. Chem.* **264,** 19416–19426
29. Chammas, R., McCaffery, J. M., Klein, A., Ito, Y., Saucan, L., Palade, G., Farquhar, M. G., and Varki, A. (1996) *Mol. Biol. Cell* **7,** 1691–1707
30. Shi, W. X., Chammas, R., and Varki, A. (1996) *J. Biol. Chem.* **271,** 15130–15138
31. Higa, H. H., Butor, C., Diaz, S., and Varki, A. (1989) *J. Biol. Chem.* **264,** 19427–19434
32. Schauer, R., Reuter, G., Stoll, S., and Shukla, A. K. (1989) *J. Biochem.* (*Tokyo*) **106,** 143–150
33. Butor, C., Higa, H. H., and Varki, A. (1993) *J. Biol. Chem.* **268,** 10207–10213
34. Butor, C., Griffiths, G., Aronson, N. N., Jr., and Varki, A. (1995) *J. Cell Sci.* **108,** 2213–2219
35. Higa, H. H., and Paulson, J. C. (1985) *J. Biol. Chem.* **260,** 8838–8849
36. Higa, H. H., Manzi, A., and Varki, A. (1989) *J. Biol. Chem.* **264,** 19435–19442
37. Schauer, R. (1988) *Methods Enzymol.* **138,** 611–626
38. Guimaraes, M. J., Bazan, J. F., Castagnola, J., Diaz, S., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Varki, A., and Zlotnik, A. (1996) *J. Biol. Chem.* **271,** 13697–13705
39. Stoddart, A., Zhang, Y., and Paige, C. J. (1996) *Nucleic Acids Res.* **24,** 4003–4008
40. Authors (1993) *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience, New-York
41. Coligan, J. E. (1996) *Current Protocols in Protein Science*, John Wiley & Sons, Inc., Brooklyn, NY
42. Sambrook, J., Maniatis, T., and Fritsch, E. F. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
43. Higa, H. H., Manzi, A., Diaz, S., and Varki, A. (1989) *Methods Enzymol.* **179,** 409–415
44. Kozak, M. (1991) *J. Cell Biol.* **115,** 887–903
45. Danpure, C. J. (1995) *Trends Cell Biol.* **5,** 230–238
46. Carlson, M., and Botstein, D. (1982) *Cell* **28,** 145–154
47. Carlson, M., Taussig, R., Kustu, S., and Botstein, D. (1983) *Mol. Cell. Biol.* **3,** 439–447
48. Kwiatkowski, D. J., Mehl, R., and Yin, H. L. (1988) *J. Cell Biol.* **106,** 375–384
49. Tagaya, Y., Kurys, G., Thies, T. A., Losi, J. M., Azimi, N., Hanover, J. A., Bamford, R. N., and Waldmann, T. A. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 14444–14449
50. Varki, A., Muchmore, E., and Diaz, S. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 882–886
51. Tokuyama, S., Moriya, S., Taniguchi, S., Yasui, A., Miyazaki, J., Orikasa, S., and Miyagi, T. (1997) *Int. J. Cancer* **73,** 410–415
52. Meuillet, E. J., Kroes, R., Yamamoto, H., Warner, T. G., Ferrari, J., Mania-Farnell, B., George, D., Rebbaa, A., Moskal, J. R., and Bremer, E. G. (1999) *Cancer Res.* **59,** 234–240