Glycoconjugate Journal 18, 205–213, 2001

Cloning and expression of human sialic acid pathway genes to generate CMP-sialic acids in insect cells

Shawn M. Lawrence[1], Kathleen A. Huddleston[2], Noboru Tomiya[3], Nam Nguyen[2], Yuan C. Lee[3], Willie F. Vann[4], Timothy A. Coleman[2] and Michael J. Betenbaugh[1]*

[1]*Department of Chemical Engineering, The Johns Hopkins University, Baltimore, MD 21218, USA*, [2]*Protein Development, Human Genome Sciences, Rockville, MD 20850, USA*, [3]*Department of Biology, The Johns Hopkins University, Baltimore, MD 21218, USA*, [4]*Laboratory of Bacterial Toxins and Laboratory of Bacterial Polysaccharides, Food and Drug Administration, Bethesda, MD 20892, USA*

The addition of sialic acid residues to glycoproteins can affect important protein properties including biological activity and in vivo circulatory half-life. For sialylation to occur, the donor sugar nucleotide cytidine monophospho-sialic acid (CMP-SA) must be generated and enzymatically transferred to an acceptor oligosaccharide. However, examination of insect cells grown in serum-free medium revealed negligible native levels of the most common sialic acid nucleotide, CMP-N-acetylneuraminic acid (CMP-Neu5Ac). To increase substrate levels, the enzymes of the metabolic pathway for CMP-SA synthesis have been engineered into insect cells using the baculovirus expression system. In this study, a human CMP-sialic acid synthase cDNA was identified and found to encode a protein with 94% identity to the murine homologue. The human CMP-sialic acid synthase (Cmp-Sas) is ubiquitously expressed in human cells from multiple tissues. When expressed in insect cells using the baculovirus vector, the encoded protein is functional and localizes to the nucleus as in mammalian cells. In addition, co-expression of Cmp-Sas with the recently cloned sialic acid phosphate synthase with N-acetylmannosamine feeding yields intracellular CMP-Neu5Ac levels 30 times higher than those observed in unsupplemented CHO cells. The absence of any one of these three components abolishes CMP-Neu5Ac production in vivo. However, when N-acetylmannosamine feeding is omitted, the sugar nucleotide form of deaminated Neu5Ac, CMP-2-keto-3-deoxy-D-glycero-D-galacto-nononic acid (CMP-KDN), is produced instead, indicating that alternative sialic acid glycoforms may eventually be possible in insect cells. The human CMP-SAS enzyme is also capable of CMP-N-glycolylneuraminic acid (CMP-Neu5Gc) synthesis when provided with the proper substrate. Engineering the CMP-SA metabolic pathway may be beneficial in various cell lines in which CMP-Neu5Ac production limits sialylation of glycoproteins or other glycans.

Keywords: baculovirus, CMP-Neu5Ac, insect, KDN, sialic acid

Abbreviations: CMP-SA, cytidine monophospho-sialic acid; CMP-Neu5Ac, CMP-N-acetylneuraminic acid; Cmp-Sas, human CMP-sialic acid synthase gene; CMP-KDN, CMP-2-keto-3-deoxy-D-glycero-D-galacto-nononic acid; CMP-Neu5Gc, CMP-N-glycolylneuraminic acid; CHO, Chinese hamster ovary; Neu5Ac, N-acetylneuraminic acid; ManNAc, N-acetylmannosamine; neuA, E. coli CMP-sialic acid synthase gene; AcCMP-SAS, recombinant baculovirus containing Cmp-Sas; Sf9, Spodoptera frugiperda cell line; AcSAS, recombinant baculovirus containing human sialic acid phosphate synthase gene; EST, expressed sequence tag; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; TBA, thiobarbituric acid; CTP, cytidine triphosphate; Neu5Gc, N-glycolylneuraminic acid; KDN, 2-keto-3-deoxy-D-glycero-D-galacto-nononic acid; MWCO, molecular weight cut-off; HPAEC, high performance anion exchange chromatography; DMB, 1,2-diamino-4,5-methylene-dioxybenzene dihydrochloride; KDN, 2-keto-3-deoxy-D-glycero-D-galacto-nononic acid.

Introduction

N-glycosylation is a particularly important post-translational modification that can affect the solubility, biological activity, and *in vivo* circulatory half-life of glycoproteins [1]. Mammalian cell lines produce glycoproteins with complex

*To whom correspondence should be addressed: M. J. Betenbaugh, Department of Chemical Engineering, The Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218. Tel.: (410) 516-5461; Fax: (410) 516-5510; E-mail: beten@jhu.edu

glycan patterns typically terminating in sialic acid residues. Although insect cells *N*-glycosylate secretory and membrane proteins, the final *N*-glycosylation pattern includes mostly high mannose or paucimannosidic (low mannose) structures [2,3]. Some oligosaccharides terminating in *N*-acetylglucosamine or galactose have also been observed [2–4]. Expression of recombinant galactosyltransferases has increased the relative amounts of *N*-glycans terminating in galactose [4–6]. The differences between mammalian and insect glycosylation can have a significant impact on a protein's characteristics and are important as recombinant glycoproteins are increasingly being used as therapeutic agents. For example, the *in vitro* activity of recombinant human thyrotropin expressed in insect cells was five times higher than that produced by Chinese hamster ovary cells (CHO), yet the *in vivo* activity of CHO thyrotropin was higher in mice [7]. The reason for this difference in activity was thought to be a more rapid clearance of the insect produced thyrotropin due to the absence of *N*-glycans terminating in sialic acid residues [7]. It is not yet known whether or not insects have the basic metabolic machinery for glycoconjugate sialylation. While some studies have suggested that insects or insect cells can perform sialylation under certain conditions [8–11], other studies have indicated that the essential genes may not be present or at least not active [3,12–13].

Sialylation requires the metabolic generation of the sugar nucleotide cytidine monophospho-sialic acid (CMP-SA) followed by the transfer of the sialic acid to an acceptor oligosaccharide in the Golgi apparatus by sialyltransferases. Previous studies indicate that insect cells grown in serum-free medium have undetectable levels of the most common sialic acid nucleotide, CMP-*N*-acetylneuraminic acid (CMP-Neu5Ac) [13,14] as well as negligible levels of its metabolic precursor, *N*-acetylneuraminic acid (Neu5Ac) [15]. Since CMP-Neu5Ac and Neu5Ac are not easily incorporated into cells [16], alternative strategies should be considered to produce Neu5Ac and CMP-Neu5Ac if sialylation is desired in insect cells. One approach is to engineer the necessary enzymatic pathways for CMP-SA synthesis. Unlike Neu5Ac or CMP-Neu5Ac, the precursor to Neu5Ac, *N*-acetylmannosamine (ManNAc) [16], and its analogs [17,18] are readily incorporated into cells. In fact, feeding of ManNAc in concert with the expression of the recently cloned sialic acid phosphate synthase gene generated large intracellular pools of Neu5Ac in insect cells [15]. In the case of mammalian cells, Neu5Ac is subsequently activated to the nucleotide sugar CMP-Neu5Ac by CMP-sialic acid synthase [19]. The activity of a CMP-sialic acid synthase was described in pigs in 1962 [20], but the gene has only recently been cloned in *E. coli* [21] and mouse [22].

Using homology searches, we identified the human *Cmp-Sas* gene (GenBank AF397212) and cloned it into a transfer vector to produce the recombinant baculovirus (AcCMP-SAS). In order to determine if insect cells could be modified to generate CMP-SA pools, *Spodoptera frugiperda* (Sf9) cells were co-infected with AcCMP-SAS and the baculovirus containing recombinant human sialic acid phosphate synthase (AcSAS) [15] in the presence of ManNAc. The result was the production of large pools of intracellular CMP-Neu5Ac, overcoming a major limitation to the sialylation of glycoproteins in insect cells. Interestingly, another sugar nucleotide, CMP-2-keto-3-deoxy-D-*glycero*-D-*galacto*-nononic acid (CMP-KDN), was generated in insect cells following infection with AcSAS and AcCMP-SAS when the cells were not supplemented with ManNAc. The enzyme CMP-SAS could also synthesize CMP-*N*-glycolylneuraminic acid (CMP-Neu5Gc) when provided with the proper substrates *in vitro*. The production of alternative nucleotide sugars in separate cell cultures may potentially allow the generation of glycoproteins with different sialic acid termini in insect cells.

Materials and methods

Gene identification, characterization, cloning, and expression

The *E. coli neuA* coding sequence was used to query the Human Genome Sciences Inc. (Rockville, MD) human cDNA database with BLAST software. One EST clone from a human prostate cell line demonstrated significant homology to *neuA* and was further characterized. The procedures used for Northern blotting, *in vitro* transcription and translation, and baculovirus cloning were the same as those described for work with the human sialic acid phosphate synthase gene [15]. For PCR amplification, the forward primer, 5′-TGTAATACGACTCACTATAGGGC*GGATCC***GCCATC**ATGGACTCGGTGGAGAAGG, contained a synthetic T7 promoter sequence (underlined), a BamHI site (italics), a KOZAK sequence (bold), and a sequence corresponding to the first six codons of *Cmp-Sas* (GenBank AF397212). The reverse primer, 5′- GTAC*GGTACC***TTACTA**TTTTTGGCATGAATTATTAACTTTTTCC, contained an Asp 718 site (italics), two in-frame stop codons (bold), and sequence representing the last six codons of *Cmp-Sas*.

Enzyme localization

Protein localization was determined with the Pierce NE-PER™ kit and modifications of a previously published procedure [23]. T-175 flasks were plated with 16×10^6 Sf9 cells, and the medium was replaced. Two flasks of cells were infected with A35-, AcSAS-, or AcCMP-SAS-amplified viruses for approximately 60 h. Cells were washed twice with PBS and incubated in 3 ml of the described hypotonic buffer [23]. After centrifugation, 400 µl of stripping solution (10 mM Tris HCl pH 7.3, 1% NP-40, 0.5% deoxycholate, 10 mM NaCl, 1.5 mM $MgCl_2$) was added to the pellet. The supernatant was removed and the soluble nuclear fraction taken after incubation with 200 µl of the described lysis buffer [23]. The protein content of all fractions was determined using the Pierce (Rockford, IL) BCA assay kit with a Molecular Devices (Sunnyvale, CA) 96-well plate reader. Total protein from each fraction was analyzed by 12.5% SDS-PAGE. Samples were transferred to a PVDF membrane, stained with Ponceau S,

appropriate bands excised, and submitted for *N*-terminal protein sequencing using an ABI-494 sequencer (PE Biosystems, Foster City, CA).

CMP-sialic acid synthase assays

To assay CMP-sialic acid synthetic ability, CHO-K1 cells (ATCC, Manassas, VA) were grown to confluence in T-75 flasks at 37°C in a humidified atmosphere with 5% CO_2 in Dulbecco's Modified Eagle medium (Invitrogen, Carlsbad, CA) supplemented with 10% FBS, 100 U/ml penicillin, 100 μg/ml streptomycin, 100 μM MEM essential amino acids, and 4 mM L-glutamine (Invitrogen, Carlsbad, CA). After washing with PBS, 1 ml of suspension solution (0.2 M Tris, 0.15 M NaCl, 1% NP-40 pH = 9.0) was added [24]. Sf9 cells were plated in 6-well dishes with fresh Ex-Cell 405 medium (JRH Biosciences, Lanexa, KS) and infected with the A35 control virus or AcCMP-SAS virus or left uninfected for 60 h. The cells were washed with PBS, and the contents of two wells were harvested with 400 μl of the resuspension solution in each well. The resulting mixtures were centrifuged for 15 min at 14,000 RPM, and the supernatant was used for enzyme assays. The protein content of each sample was determined as previously described. The activity assays were modified from the TBA assay used to measure CMP-SA formation [25]. Twenty μl of lysates were incubated with 100 μl of substrate solution [25] at 37°C for 40 min. Forty μl of 1.6 M $NaBH_4$ was used to reduce free Neu5Ac, and the mixture was then treated with 40 μl of 20 N H_3PO_4. Released Neu5Ac was oxidized with 0.1 ml 25 mM periodic acid for 30 min at 37°C then treated with 0.1 ml of 62 mM sodium bisulfite. The use of bisulfite instead of arsenite gave similar responses. 1.0 ml of 0.1 M TBA was added, and the mixture was placed in a boiling water bath for 7.5 min. The tubes were placed on ice and 2 ml of DMSO with 5% concentrated HCl was added. The absorbance at 552 nm was measured using a Spectronic 20 spectrophotometer (Bausch and Lomb, Rochester, NY), and the amount of CMP-Neu5Ac formed calculated by Neu5Ac standards.

Nuclei for enzymatic assays were prepared by taking 30 ml of Sf9 cells at a density of $1–1.2 \times 10^6$ after infection with AcCMP-SAS or control virus (A35) for three days. The cells were spun down and washed with PBS. Cells were homogenized on ice in 3 ml 20 mM Tris HCl pH 7.2, 0.32 M sucrose, 3 mM $MgCl_2$ with 20 strokes of a tight-fitting Dounce homogenizer. 0.25 M sucrose (2.4 ml) was added, and the mixture centrifuged at 700g for 10 min. 2.4 M sucrose (3 ml) was added, and the mixture spun at 50,000g for 1 h. The precipitate was resuspended in 200 μl 0.25 M sucrose, 1 mM $MgCl_2$, 10 mM Tris HCl pH 7.6 and centrifuged at 7000g for 10 min. 20 μl of substrate solution [25] with 2.8 mM of sialic acid (Neu5Ac, Neu5Gc, or KDN) was incubated with 5 μl of the nuclear preparation for 1 h at 37°C. Adenosine monophosphate was added as an internal standard, and a recently developed procedure was used to determine nucleotide sugar content [14]. Intracellular Neu5Gc content was determined by DMB derivatization according to previously described methods [15].

CMP-sialic acid levels

Sf9 cells were grown in 6-well plates and infected with combinations of A35, AcSAS, and AcCMP-SAS viruses along with 10 mM ManNAc feeding for approximately 90 h. CHO cells grown according to the culture conditions described above were grown in T-75 flasks with and without 10 mM ManNAc supplementation for 48 h. All cells were washed with PBS and aliquots taken for protein quantification. Cells were resuspended in 75% ethanol and sonicated on ice for 30 s. Lysates were centrifuged, frozen in liquid nitrogen, and lyophilized. Samples were resuspended in 120 μl of 40 mM phosphate buffer, pH 9.2, to stabilize CMP-sialic acids. Samples were centrifuged again and the supernatant filtered through 10,000 molecular weight cut-off membranes. CMP-sialic acids were detected using HPAEC [14] and quantified versus the response curves of standard CMP-Neu5Ac (Sigma, St. Louis, MO), CMP-KDN (Dr. Yasuhiro Kajihara), and CMP-Neu5Gc (Dr. Akemi Suzuki, RIKEN Frontier Research Systems, Japan).

Results

Gene identification

The human CMP-sialic acid synthase gene (GenBank AF397212) was identified from the expressed sequence tag (EST) database of Human Genome Sciences Inc. based on homology searches using the known *E. coli* CMP-sialic acid synthase gene [21]. The human sialic acid phosphate synthase gene had previously been identified from the *E. coli* gene using a similar technique [15]. The murine sequence was subsequently published [22], and the homologies between the human, bacterial, and murine sequences are shown (Figure 1). Over the entire sequence, the putative human gene is 24% identical and 48% similar to the bacterial sequence and 94% identical and 96% similar to the murine sequence. Northern blots probing human mRNA samples using the gene sequence showed ubiquitous transcription of the gene in all human tissues tested (Figure 2A).

Using *in vitro* transcription and translation, the putative gene was shown to encode a protein of the expected molecular weight (Figure 2B). The same gene was used to generate a recombinant baculovirus, AcCMP-SAS, and this virus was used to infect Sf9 insect cells. Pulse labeling followed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) separation showed a protein of 49 kDa (Figure 2C, lane 1), consistent with the predicted weight of 48.5 kDa, that was not observed in the control infection (A35) (Figure 2C, lane 2).

Enzyme localization

CMP-sialic acid synthetic activity in mammals has been shown to localize to the nucleus [19], and fluorescent labeling

```
        1                                                   50
Murine  MDALEKGAVT SGPAPRGRPS RGRPPKLQR. SRGA.GRGLE KPPHLAALVL
 Human  MDSVEKGAAT SVSNPRGRPS RGRPPKLQRN SRGGQGRGVE KPPHLAALIL
E.coli  .......... .......... .......... .......... MRTKIIAIIP

        51                                                  100
Murine  ARGGSKGIPL KNIKRLAGVP LIGWVLRAAL DAGVFQSVWV STDHDEIENV
 Human  ARGGSKGIPL KNIKHLAGVP LIGWVLRAAL DSGAFQSVWV STDHDEIENV
E.coli  ARSGSKGLRN KNALMLIDKP LLAYTIEAAL QSEMFEKVIV TTDSEQYGAI

        101                                                 150
Murine  AKQFGAQVHR RSSETSKDSS TSLDAIVEFL NYHNEVDIVG NIQATSPCLH
 Human  AKQFGAQVHR RSSEVSKDSS TSLDAIIEFL NYXNEXDIVG NIQATSXCLH
E.coli  AESYGADFLL RPEELATDKA SSFEFIKHAL SIYTDYESFA LLQPTSPFRD

        151                                                 200
Murine  PTDLQKVAEM IRE.EGYDSV FSVVRRHQFR WSEIQKGVRE VTEPLNLNPA
 Human  PTDLQKVAEM IRE.EGYDSX FSVVRRHQFR WSEIQKGVRE VTEPLNLNPA
E.coli  STHIIEAVKL YQTLEKYQCV VSVTRSN..K PSQIIRPLDD YSTLSFFDLD

        201                                                 250
Murine  KRPRRQDWDG ELYENGSFYF A.KRHLIEMG YLQGGKMAYY EMRAEHSVDI
 Human  KRPRRQDWDG ELYENGSFYF A.KRHLIEMG YLQGGKWHTT KCELEHSVDI
E.coli  YSKYNRNSIV EYHPNGAIFI ANKQHYLHTK HFFGRYSLAY IMDKESSLDI

        251                                                 300
Murine  DVDIDWPIA. ..EQRVLR.. .......... .......... ..........
 Human  DVDIDWPIA. ..EQRVLR.. .......... .......... ..........
E.coli  DDRMDFELAI TIQQKKNRQK IDLYQNIHNR INEKRNEFDS VSDITLIGHS

        301                                                 350
Murine  .FGYFGKEKL KEIKLLVCNI DG........ CLTNGHIYVS GDQKEIISYD
 Human  .YGYFGKEKL KEIKLLVCNI DG........ CLTNGHIYVS GDQKEIISYD
E.coli  LFDYWDVKKI NDIEVNNLGI AGINSKEYYE YIIEKELIVN FGEFVFIFFG

        351                                                 400
Murine  VKDAIGISLL KKSGIEVRLI SERACSKQTL SALKLDCKTE VSVSDKLATV
 Human  VKDAIGISLL KKSGIEVRLI SERACSKQTL SSLKLDCKME VSVSDKLAVV
E.coli  TND.IVVSDW KKEDTLWYLK KTCQYIKKKN AASKIYLLSV PPVFGRIDRD

        401                                                 450
Murine  DEWRKEMGLC WKEVAYLGNE VSDEECLKRV GLSAVPADAC SGAQKAVGYI
 Human  DEWRKEMGLC WKEVAYLGNE VSDEECLKRV GLSGAPADAC SYAQKAVGYI
E.coli  NRIINDLNSY LRENVDFAKF ISLDHVLK.. .......DSY GNLNKMYTYD

        451                             483
Murine  C...KCSGGR GAIREFAEHI FLLIEKVNNS CQK
 Human  C...KCNGGR GAIREFAEHI CLLMEKVNNS CQK
E.coli  GLHFNSNGYT VLENEIAEIV K......... ...
```

Figure 1. Homology alignment of murine (top row), human (middle row), and *E. coli* (bottom row) CMP-sialic acid synthase amino acid sequences.

of the recombinant murine enzyme expressed in 3T3 cells confirmed this localization [22]. Analysis of the human CMP-SAS sequence with the pSORT algorithm also predicted the protein would localize to the nucleus because of the nuclear localization sequence (KRPR) starting at amino acid 200. To determine if the recombinant human enzyme localized to the nucleus of insect cells, a nuclear separation using the Pierce NE-PER™ kit was performed. A protein band from the AcCMP-SAS infection was found predominantly in the nuclear fraction (Figure 3, lane 5) but not in the nuclear fraction of a control infection (Figure 3, lane 8). In contrast, recombinant human sialic acid phosphate synthase (SAS) localized in the cytoplasmic fraction (Figure 3, lane 2) but not in the nuclear fraction (Figure 3, lane 7) as observed previously for the native enzyme in mammalian cells [26]. Using a thiobarbituric acid (TBA) assay based on a previously described procedure [25], the nuclear fraction of AcCMP-SAS-infected but not AcSAS-infected cells showed CMP-SAS activity (data not shown). The protein band in the nuclear fraction following a detergent separation [23] was also

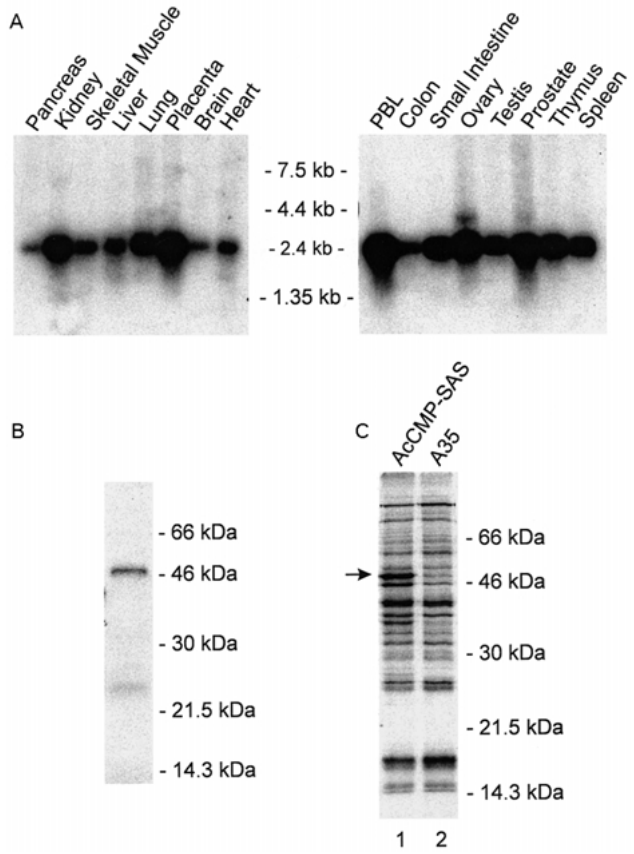

Figure 2. (A) Autoradiogram of Northern blot probing for *Cmp-Sas* mRNA from the indicated human tissues. PBL = peripheral blood lymphocyte. (B) Autoradiogram of *in vitro* products from transcription and translation of PCR product using primers for the *Cmp-Sas* gene with T7 promoter as described in *Materials and Methods*. The baculovirus transfer vector pA2 containing *Cmp-Sas* was used as the PCR template. (C) Autoradiogram of ^{35}S pulse labeling experiment showing AcCMP-SAS (lane 1) and A35 negative control (lane 2) infected cell lysates. CMP-SAS band is indicated by arrow.

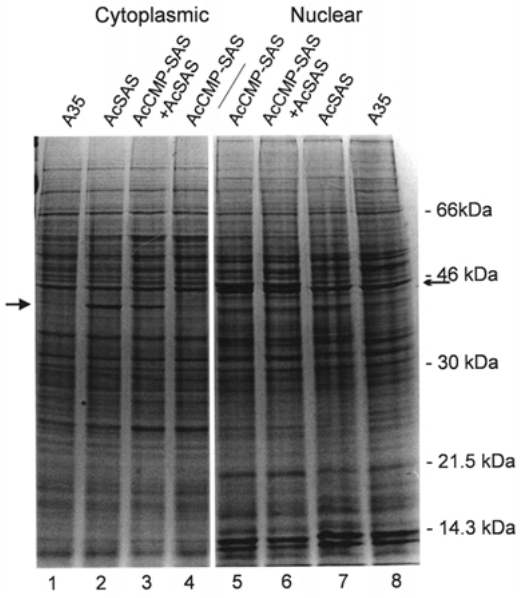

Figure 3. Coomassie-stained SDS-PAGE separation of cytoplasmic (lanes 1–4) and nuclear (lanes 5–8) fractions prepared using the Pierce NE-PER^{TM} kit. Lysates from A35 (control)-infected (lanes 1 and 8), AcSAS-infected (lanes 2 and 7), AcSAS- and AcCMP-SAS- co-infected (lanes 3 and 6), and AcCMP-SAS-infected (lanes 4 and 5) Sf9 cells were used to obtain the nuclear and cytoplasmic fractions. SAS is observed in the cytoplasmic fraction (molecular weight indicated by large arrow), and CMP-SAS is predominantly found in the nuclear fraction (molecular weight indicated by small arrow).

transferred to a PVDF membrane for *N*-terminal protein sequencing. Interestingly, the sequence was identical to amino acids 29 through 33, as predicted from the gene sequence, indicating cleavage of the first 28 amino acids of the nuclear localized protein.

CMP-sialic acid synthase activity

The function of the putative gene product was determined *in vitro* using cell lysates incubated with the precursors *N*-acetylneuraminic acid and cytidine triphosphate (CTP) with the modified TBA assay. CMP-SA synthetic ability was observed only in Sf9 cells following AcCMP-SAS infection and was not detectable in control infected or uninfected cells (data not shown). Furthermore, lysates from Sf9 cells infected with AcCMP-SAS had 60 times the CMP-Neu5Ac synthetic ability (300 pmol/min/μg lysate protein) of CHO cell lysates.

The substrate specificity of the human enzyme was tested using an *in vitro* assay. A nuclear fraction was taken, and the fractions incubated with Neu5Ac, *N*-glycolylneuraminic acid (Neu5Gc), and KDN (2-keto-3-deoxy-D-*glycero*-D-*galacto*-nononic acid). After lyophilization and 10,000 molecular weight cut-off (MWCO) filtration, the samples were analyzed using high performance anion exchange chromatography (HPAEC) with UV monitoring [14], and the resulting chromatograms are shown in Figure 4. Human CMP-sialic acid synthase has significant CMP-sialic acid synthetic ability using Neu5Ac, Neu5Gc, or KDN as substrates with CTP. Significant CMP-sialic acid synthetic ability was not observed in the corresponding control infected lysates.

Completion of the CMP-sialic acid pathway in insect cells

The production of Neu5Ac in insect cells has been accomplished by infection with the recombinant human sialic acid phosphate synthase virus (AcSAS) in concert with feeding of the precursor ManNAc [15]. Co-infection of AcSAS and AcCMP-SAS with concomitant ManNAc feeding was attempted in order to produce and activate sialic acids in insect cells. Proteins with the molecular masses and cellular locations expected for SAS and CMP-SAS are seen in lysates

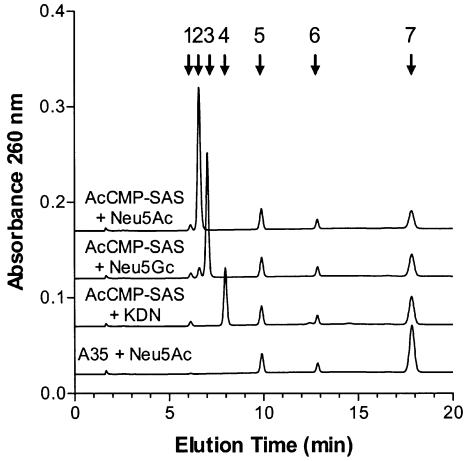

Figure 4. Chromatograms showing the formation of CMP-sialic acids from *in vitro* assays. The sugar nucleotides were detected using UV absorption after the reaction mixture was treated and separated by HPAEC chromatography. The top three chromatograms represent assays in which AcCMP-SAS lysates were incubated with the free sialic acids as labeled (from top to bottom: Neu5Ac, Neu5Gc, and KDN). The lower chromatogram shows the assay in which A35-infected (control) lysates were incubated with Neu5Ac. The identities of the peaks were determined using standards and are: 1 CMP; 2 CMP-Neu5Ac; 3 CMP-Neu5Gc; 4 CMP-KDN; 5 AMP (internal standard); 6 CDP; 7 CTP. A small CMP-Neu5Ac peak is observed in the CMP-Neu5Gc assay because of Neu5Ac contamination in the Neu5Gc source.

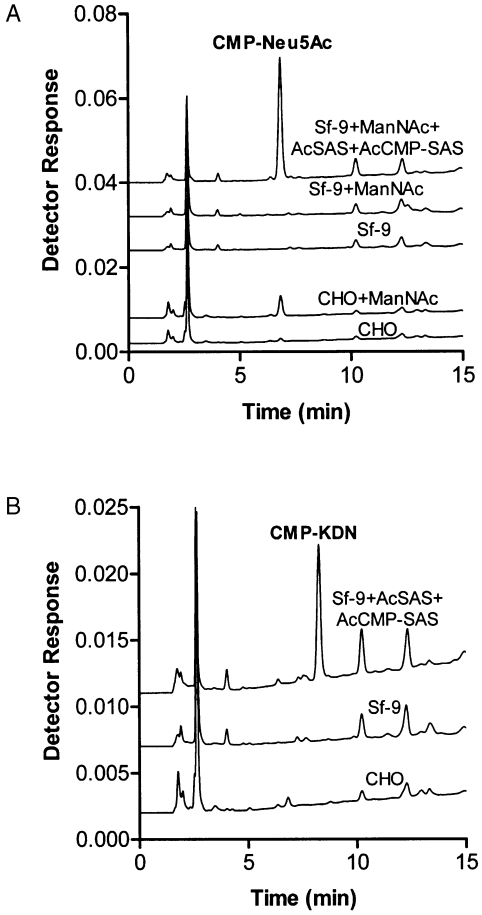

Figure 5. Chromatograms illustrating presence of CMP-sialic acids in CHO and Sf9 cell lines with the indicated medium supplementation and/or baculovirus infection strategies. The CMP-sialic acids were separated by HPAEC and detected by UV absorbance. The original chromatogram values have been divided by sample protein concentration to normalize the chromatograms. (A) Intracellular CMP-Neu5Ac levels in CHO and Sf9 under indicated conditions. CMP-Neu5Ac standard eluted at 6.8 min. (B) Intracellular CMP-KDN levels under indicated conditions. CMP-KDN standard eluted at 8.3 min.

of cells infected with each virus individually and together (Figure 3).

The HPAEC procedure that was used to determine CMP-sialic acid levels in the *in vitro* assays was used to measure CMP-Neu5Ac levels in samples from several different infections (Figure 5). The levels of CMP-Neu5Ac produced on a protein basis are listed in Table 1. In agreement with previous work [16], CHO cells are shown to have measurable CMP-Neu5Ac levels which are augmented upon feeding of 10 mM ManNAc (Figure 5A). Uninfected Sf9 cells grown in serum-free medium do not have detectable CMP-Neu5Ac levels with or without ManNAc feeding. High levels of CMP-Neu5Ac are produced in insect cells only when all three components of the sialic acid pathway are present: the precursor sugar ManNAc, the sialic acid phosphate synthase from AcSAS, and the CMP-sialic acid synthase from AcCMP-SAS. CMP-Neu5Ac levels in the insect cells fed with ManNAc are approximately 6 times those seen in CHO with ManNAc feeding and 30 times those seen without ManNAc feeding.

When ManNAc feeding is omitted from an AcSAS and AcCMP-SAS infection, a CMP-Neu5Ac peak is not observed but a peak eluting at 8.5 min is detected (Figure 5B). This peak has the same elution time as authentic CMP-KDN (gift of Dr. Yasuhiro Kajihara, Yokohama City University, Japan) and is found in quantities of 4 pmol/μg protein. Further confirmation of the peak as CMP-KDN was obtained by acid hydrolysis of the eluted peak followed by fluorescent labeling with 1,2-diamino-4,5-methylene-dioxybenzene dihydrochloride (DMB) and HPLC separation. The resulting fluorescent compound eluted at the same time as authentic KDN derivatized with DMB (data not shown).

Discussion

The baculovirus expression system offers several advantages, including the capacity to generate significant levels of post-translationally modified recombinant proteins in a eucaryotic cell line. However, one of the drawbacks of recombinant baculovirus expression is its inability to produce glycoproteins with a significant percentage of complex *N*-glycans, especially those terminating in sialic acid residues. Sialic acids have been associated with numerous *in vivo* biological processes

Table 1. Intracellular CMP-Neu5Ac levels of CHO and Sf9 cells using the indicated feeding strategy and infection scheme. The CMP-Neu5Ac content of cell lysates was determined by HPAEC.

Sample	*CMP-Neu5Ac (pmol/μg protein)*
CHO	0.3
CHO + ManNAc	1.8
Sf9	< 0.02
Sf9 + ManNAc	< 0.02
Sf9 + A35 + ManNAc	0.02
Sf9 + AcSAS + ManNAc	0.02
Sf9 + AcCMP-SAS + ManNAc	0.1
Sf9 + A35 + AcSAS	< 0.02
Sf9 + A35 + AcCMP-SAS	< 0.02
Sf9 + AcSAS + AcCMP-SAS	0.02
Sf9 + A35 + AcSAS + ManNAc	0.08
Sf9 + A35 + AcCMP-SAS + ManNAc	0.04
Sf9 + AcCMP-SAS + AcSAS + ManNAc	10.

including development [27], pathogen infectivity [28], ligand-receptor interactions [28], cancer metastasis [29], and the elimination of non-sialylated glycoproteins by the asialoglycoprotein receptor [7].

The inability of insect cells to generate sufficient pools of the donor nucleotide sugar CMP-Neu5Ac is likely to be one reason for their deficiency in glycoprotein sialylation. Sf9 cells lack sizeable pools of either this nucleotide sugar or the alternative nucleotide sugar, CMP-KDN, when grown in serum-free medium. Furthermore, the supplementation of ManNAc does not overcome this limitation, indicating a bottleneck in the conversion of ManNAc to CMP-Neu5Ac. This limitation may be due to the absence of essential genes or due to the inactivity of these genes in insect cell cultures.

However, this bottleneck to sialylation of glycoconjugates can be overcome by substrate feeding and engineering the CMP-SA synthesis genes into insect cells. In the present study, the human CMP-sialic acid synthase gene has been identified and expressed in an active form in insect cells. The enzyme localizes to the nucleus of insect cells as observed in the native mammalian host suggesting that the encoded nuclear localization signals are functional in insect cells as well. Moreover, the protein isolated from the nuclear fraction retains enzyme activity, but its first 28 amino acids are truncated as determined by *N*-terminal protein sequencing analysis. The cause of the cleavage of the nuclear-localized CMP-SA synthase is unknown as is the relevance of this cleavage to the nuclear localization process in insect cells.

By feeding ManNAc along with expression of CMP-SAS and SAS, CMP-Neu5Ac levels more than five times higher than those observed in ManNAc-supplemented CHO cells are achieved in Sf9 cells grown in serum-free medium. Some low-level endogenous sialic acid metabolic activity may be present in insect cells as seen with the addition of ManNAc and just one component of the pathway (Table 1). However, CMP-Neu5Ac levels are increased more than 100-fold by the addition of all three components. In order to complete sialylation, the CMP-Neu5Ac must be transported into the Golgi where a sialyltransferase can incorporate the Neu5Ac to an acceptable target oligosaccharide. Expression of recombinant galactose transferase, sialyltransferases, and a CMP-sialic acid transporter may also be required to achieve sialylation in insect cell culture. Recently, Jarvis and co-workers detected sialylation of recombinant glycoproteins when heterologous galactosyltransferase and sialyltransferase were co-expressed in Sf9 cells [11]. The sialylation was accomplished in serum-bearing medium, and serum has a significant sialic acid content [30].

In addition to its CMP-Neu5Ac synthetic ability, the human enzyme also has broad substrate specificity forming CMP-Neu5Gc and CMP-KDN from the corresponding free sialic acids. A CMP-KDN synthase has been previously reported from trout which has lower activity for producing CMP-Neu5Ac and CMP-Neu5Gc [31]. The apparent high activity of the human enzyme toward Neu5Gc compared to Neu5Ac (Figure 4) is particularly significant. While the gene encoding the CMP-sialic acid hydroxylase enzyme in humans does not make active enzyme due to a frame-shift mutation [32–33], human glycoconjugates with Neu5Gc have nonetheless been reported in the literature [34]. The activity of the human enzyme towards Neu5Gc could explain the presence of Neu5Gc-containing glycostructures in humans if Neu5Gc is provided in the diet as has been suggested [34].

Interestingly, significant levels of CMP-KDN are produced in insect cells co-expressing CMP-SAS and SAS without ManNAc feeding. In all other samples listed in Table 1, CMP-KDN was not detected at the limit of 0.04 pmol/μg protein of CMP-KDN. Previously, we observed the production of KDN when insect cells expressed SAS in the presence or absence of ManNAc feeding [15]. In the absence of SAS expression, KDN was not detected in insect cell lysates. Since CMP-KDN was generated in this study, both the SAS and CMP-SAS enzymes accept the substrates for generating deaminated sialic acid forms. Furthermore, it was not necessary to supplement the insect cell medium for the generation of CMP-KDN. We determined previously that the substrate for the SAS enzyme is mannose-6-phosphate [15], and therefore this metabolite must be present in insect cells grown in serum-free medium. More significant is the observation that the CMP-KDN is generated in the absence of detectable levels of CMP-Neu5Ac. This is the first report of a cellular system that can produce substantial levels of activated KDN without activated Neu5Ac. Furthermore, Sf9 cells natively have less than 20 pmol/μg protein of potentially immunogenic Neu5Gc (free sugar or sugar nucleotide) as measured by DMB derivatization with or without 10,000 MWCO. In contrast, similar experiments with CHO cells showed 90 and 200 pmol Neu5Gc/μg protein with and without 10,000 MWCO filtration, respectively. Since at

least one sialyltransferase recognizes both CMP-KDN as well as CMP-Neu5Ac [35], this expression system may eventually allow the exclusive KDN sialylation of glycoproteins. KDN was first identified in 1986 [36], and its effects on function and activity of a sialylated glycoprotein are largely unknown. However, elevated KDN:Neu5Ac levels have been observed in tissues during biological events such as cancer and development [37]. In addition, glycoproteins sialylated by KDN rather than Neu5Ac may be less susceptible to sialidases [36]. Such proteins could have longer *in vivo* circulatory half-lives and potentially greater pharmaceutical value if they are not immunogenic.

The capacity to enhance sialylation by engineering the CMP-SA pathway may have significance beyond the baculovirus expression system. Reports have suggested that limitations in levels of activated sialic acids may exist in mammalian cell lines as well [38]. Other expression systems such as yeast and plant cells may also lack some of the genes involved in generating CMP-SA, so the pathway may be engineered into other recombinant hosts as well. In these cases, the sialyltransferases and the enzymes involved in sugar nucleotide transport may also be needed. Ultimately, the number of cellular species that can generate complex "mammalian-like" glycoforms during recombinant protein expression may be increased through the metabolic engineering of the CMP-SA synthetic and transfer pathways.

Acknowledgments

The authors acknowledge Yasuhiro Kajihara for the generous gift of CMP-KDN, Akemi Suzuki for the generous gift of CMP-Neu5Gc, Dave Parmelee for *N*-terminal protein sequencing, and Donald Jarvis for review of the manuscript. NSF grants BES9814100 from the Metabolic Engineering Program and DGE9843635 from the Graduate Research Fellowship (to S.L.) were used to fund this research.

References

1 Goochee CF, Gramer MJ, Andersen DC, Bahr JB, Rasmussen JR, The oligosaccharides of glycoproteins: bioprocess factors affecting oligosaccharide structure and their effect on glycoprotein properties, *Bio/technology* **9**, 1347–1355 (1991).

2 Altmann F, Staudacher E, Wilson IB, März L, Insect cells as hosts for the expression of recombinant glycoproteins, *Glycoconjugate J* **16**, 109–123 (1999).

3 Hsu TA, Takahashi N, Tsukamoto Y, Kato K, Shimada I, Masuda K, Whiteley EM, Fan JQ, Lee YC, Betenbaugh MJ, Differential *N*-glycan patterns of secreted and intracellular IgG produced in *Trichoplusia ni* cells, *J Biol Chem* **272**, 9062–9070 (1997).

4 Ailor E, Takahashi N, Tsukamoto Y, Masuda K, Rahman BA, Jarvis DL, Lee YC, Betenbaugh MJ, *N*-glycan patterns of human transferrin produced in *Trichoplusia ni* insect cells: effects of mammalian galactosyltransferase, *Glycobiology* **10**, 837–847 (2000).

5 Jarvis DL, Finn EE, Modifying the insect cell *N*-glycosylation pathway with immediate early baculovirus expression vectors, *Nature Biotech* **14**, 1288–1292 (1996).

6 Hollister JR, Shaper JH, Jarvis DL, Stable expression of mammalian beta 1,4-galactosyltransferase extends the *N*-glycosylation pathway in insect cells, *Glycobiology* **8**, 473–480 (1998).

7 Grossmann M, Wong R, The NG, Tropea JE, East-Palmer J, Weintraub BD, Szkudlinski MW, Expression of biologically active human thyrotropin (hTSH) in a baculovirus system: effect of insect cell glycosylation on hTSH activity *in vitro* and *in vivo*, *Endocrinology* **138**, 92–100 (1997).

8 Malykh YN, Krisch B, Gerardy-Schahn R, Lapina EB, Shaw L, Schauer R, The presence of *N*-acetylneuraminic acid in Malpighian tubules of larvae of the cicada *Philaenus spumarius*, *Glycoconjugate J* **16**, 731–739 (1999).

9 Roth J, Kempf A, Reuter G, Schauer R, Gehring WJ, Occurrence of sialic acids in *Drosophila melanogaster*, *Science* **256**, 673–675 (1992).

10 Davidson DJ, Fraser JF, Castellino FJ, Oligosaccharide processing in the expression of human plasminogen cDNA by Lepidopteran insect (*Spodoptera frugiperda*) cells, *Biochemistry* **29**, 5584–5590 (1990).

11 Hollister JR, Jarvis DL, Engineering lepidopteran insect cells for sialoglycoprotein production by genetic transformation with mammalian beta 1,4-galactosyltransferase and alpha 2,6-sialyltransferase genes, *Glycobiology* **11**, 1–9 (2001).

12 Angata T, Varki A, Cloning, characterization, and phylogenetic analysis of siglec-9, a new member of the CD33-related group of siglecs. Evidence for co-evolution with sialic synthesis pathways, *J Biol Chem* **275**, 22127–22135 (2000).

13 Hooker AD, Green NH, Baines AJ, Bull AT, Jenkins N, Strange PG, James DC, Constraints on the transport and glycosylation of recombinant IFN-gamma in Chinese hamster ovary and insect cells, *Biotechnol Bioeng* **63**, 559–572 (1999).

14 Tomiya N, Ailor E, Lawrence SM, Betenbaugh MJ, Lee YC, Determination of nucleotides and sugar nucleotides involved in protein glycosylation by high-performance anion-exchange chromatography: sugar nucleotide contents in cultured insect cells and mammalian cells, *Anal Biochem* **293**, 129–137 (2001).

15 Lawrence SM, Huddleston KA, Pitts LR, Nguyen N, Lee YC, Vann WF, Coleman TA, Betenbaugh MJ, Cloning and expression of the human *N*-acetylneuraminic acid phosphate synthase gene with 2-keto-3-deoxy-D-*glycero*- D-*galacto*-nononic acid biosynthetic ability, *J Biol Chem* **275**, 17869–17877 (2000).

16 Gu X, Wang DI, Improvement of interferon-gamma sialylation in Chinese hamster ovary cell culture by feeding of *N*-acetylmannosamine, *Biotechnol Bioeng* **58**, 642–648 (1998).

17 Mahal LK, Yarema KJ, Lemieux GA, Bertozzi CR, Chemical approaches to glycobiology: engineering cell surface sialic acids for tumor targeting. In *Sialobiology and Other Novel Forms of Glycosylation*, Lee YC, Inoue Y, Troy FA (eds), Gakushin Publishing, Osaka, 1999, pp. 273–280.

18 Reutter W, Keppler O, Pawlita M, Schuler C, Horstkorte R, Schmidt C, Hoppe B, Lucka L, Stehling P, Mickeleit M, Biochemical engineering by new *N*-acetylmannosamines of sialic acid creates new biological characteristics and technical tools of its *N*-acyl side chain. In *Sialobiology and Other Novel Forms of Glycosylation*, Lee YC, Inoue Y, Troy FA (eds), Gakushin Publishing, Osaka, 1999, pp. 281–288.

19 Kean EL, Sialic acid activation, *Glycobiology* **1**, 441–447 (1991).
20 Roseman S, Enzymatic synthesis of cytidine-5′-monophosphosialic acids, *Proc Natl Acad Sci USA* **48**, 437–441 (1962).
21 Zapata G, Vann WF, Aaronson W, Lewis MS, Moos M, Sequence of the cloned *Escherichia coli* K1 CMP-*N*-acetylneuraminic acid synthetase gene, *J Biol Chem* **264**, 14769–14774 (1989).
22 Munster AK, Eckhardt M, Potvin B, Muhlenhoff M, Stanley P, Gerardy-Schahn R, Mammalian cytidine 5′-monophosphate *N*-acetylneuraminic acid synthetase: a nuclear protein with evolutionarily conserved structural motifs, *Proc Natl Acad Sci USA* **95**, 9140–9145 (1998).
23 Miyamoto C, Smith GE, Farrell-Towt J, Chizzonite R, Summers MD, Ju G, Production of human c-*myc* protein in insect cells infected with a baculovirus expression vector, *Mol Cell Biol* **5**, 2860–2865 (1985).
24 Potvin B, Raju TS, Stanley P, *lec32* is a new mutation in Chinese hamster ovary cells that essentially abrogates CMP-*N*-acetylneuraminic acid synthetase activity, *J Biol Chem* **270**, 30415–30421 (1995).
25 Vann WF, Silver R, Abeijon C, Chang K, Aaronson W, Sutton A, Finn C, Lindner W, Kotsatos M, Purification, properties, and genetic location of *Escherichia coli* cytidine 5′-monophosphate *N*-acetylneuraminic acid synthetase, *J Biol Chem* **262**, 17556–17562 (1987).
26 Van Rinsum J, Van Dijk W, Hooghwinkel GJ, Ferwerda W, Subcellular localization and tissue distribution of sialic acid-forming enzymes. *N*-acetylneuraminate-9-phosphate synthase and *N*-acetylneuraminate 9-phosphatase, *Biochem J* **223**, 323–328 (1984).
27 Cunningham BA, Hoffman S, Rutishauser U, Hemperly JJ, Edelman GM, Molecular topography of the neural cell adhesion molecule N-CAM: surface orientation and location of sialic acid-rich and binding regions, *Proc Natl Acad Sci USA* **80**, 3116–3120 (1983).
28 Schauer R, Kelm S, Reuter G, Roggentin P, Shaw L, Biochemistry and role of sialic acids. In *Biology of the Sialic Acids*, Rosenberg A (eds), New York, Plenum Press, 1995, pp. 7–67.
29 Fukuda M, Possible roles of tumor-associated carbohydrate antigens, *Cancer Res* **56**, 2237–2244 (1996).
30 Hara S, Yamaguchi M, Takemori Y, Furuhata K, Ogura H, Nakamura M, Determination of mono-*O*-acetylated *N*-acetylneuraminic acids in human and rat sera by fluorometric high-performance liquid chromatography, *Anal Biochem* **179**, 162–166 (1989).
31 Terada T, Kitazume S, Kitajima K, Inoue S, Ito F, Troy FA, and Inoue Y, Synthesis of CMP-deaminoneuraminic acid (CMP-Kdn) using the CTP:CMP-3-deoxynonulosonate cytidylyltransferase from rainbow trout testis. Identification and characterization of a CMP-Kdn synthetase, *J Biol Chem* **268**, 2640–2648 (1993).
32 Chou H, Takematsu H, Diaz S, Iber J, Nickerson E, Wright KL, Muchmore EA, Nelson DL, Warren ST, Varki A, A mutation in human CMP-sialic acid hydroxylase occurred after the *Homo-Pan* divergence, *Proc Natl Acad Sci USA* **95**, 11751–11756 (1998).
33 Irie A, Koyama S, Kozutsumi Y, Kawasaki T, Suzuki A, The molecular basis for the absence of *N*-glycolylneuraminic acid in humans, *J Biol Chem* **273**, 15855–15871 (1998).
34 Varki A, Diversity at the 5-carbon position of sialic acids. In *Sialobiology and Other Novel Forms of Glycosylation*, Lee YC, Inoue Y, Troy FA (eds), Gakushin Publishing, Osaka, 1999, pp. 281–288.
35 Angata T, Matsuda T, Kitajima K, Synthesis of neoglycoconjugates containing deaminated neuraminic acid (KDN) using rat liver alpha-2,6-sialyltransferase, *Glycobiology* **8**, 277–284 (1998).
36 Nadano D, Iwasaki M, Endo S, Kitajima K, Inoue S, Inoue Y, A naturally occurring deaminated neuraminic acid, 3-deoxy-D-*glycero*-D-*galacto*-nonulosonic acid (KDN). Its unique occurrence at the nonreducing ends of oligosialyl chains in polysialoglycoprotein of rainbow trout eggs, *J Biol Chem* **261**, 11550–11557 (1986).
37 Inoue S, Lin SL, Chang T, Wu SH, Yao CW, Chu TY, Troy FA, Inoue Y, Identification of free deaminated sialic acid (2-keto-3-deoxy-D-*glycero*-D-*galacto*-nononic acid) in human red blood cells and its elevated expression in fetal cord red blood cells and ovarian cancer cells, *J Biol Chem* **273**, 27199–27204 (1998).
38 Pels Rijcken WR, Overdijk B, Van den Eijnden DH, Ferwerda W, The effect of increasing nucleotide-sugar concentrations on the incorporation of sugars into glycoconjugates in rat hepatocytes, *Biochem J* **305**, 865–870 (1995).

Received 8 May 2001; revised and accepted 12 July 2001