REPORT

Identification of the Gene Encoding the Enzyme Deficient in Mucopolysaccharidosis IIIC (Sanfilippo Disease Type C)

Xiaolian Fan, Huiwen Zhang, Sunqu Zhang, Richard D. Bagshaw, Michael B. Tropak, John W. Callahan, and Don J. Mahuran

Mucopolysaccharidosis IIIC (MPS IIIC), or Sanfilippo C, represents the only MPS disorder in which the responsible gene has not been identified; however, the gene has been localized to the pericentromeric region of chromosome 8. In an ongoing proteomics study of mouse lysosomal membrane proteins, we identified an unknown protein whose human homolog, TMEM76, was encoded by a gene that maps to 8p11.1. A full-length mouse expressed sequence tag was expressed in human MPS IIIC fibroblasts, and its protein product localized to the lysosome and corrected the enzymatic defect. The mouse sequence was used to identify the full-length human homolog (*HGSNAT*), which encodes a protein with no homology to other proteins of known function but is highly conserved among plants and bacteria. Mutational analyses of two MPS IIIC cell lines identified a splice-junction mutation that accounted for three mutant alleles, and a single base-pair insertion accounted for the fourth.

Mucopolysaccharidosis IIIC (MPS IIIC [MIM 252930], or Sanfilippo C), first described by Kresse et al.,[1] is an autosomal recessive disorder characterized by the lysosomal storage of heparin and heparan sulfate. Clinically, MPS IIIC is similar to the other three forms of MPS III (A, B, and D). MPS III has a prevalence in Australia of 1 in 66,000 births, with the IIIC form accounting for 1 in 1,407,000 (MPS Australia). Symptoms can vary in severity, in some cases including severe CNS involvement but only mild somatic disease.[2] Death usually occurs during the late teens. The defective enzyme in MPS IIIC is heparin acetyl-CoA:α-glucosaminide N-acetyltransferase, EC 2.3.1.78 (N-acetyltransferase).[3] Identification of the gene (*HGSNAT*) encoding this enzyme would facilitate mutational analyses, to elucidate the molecular basis of the disease and to suggest potential clinical approaches for its treatment.

N-acetyltransferase is a lysosomal membrane protein whose function is to acetylate the nonreducing, terminal α-glucosamine residue of intralysosomal heparin or heparan sulfate, converting it into a substrate for luminal α-N-acetyl glucosaminidase. Therefore, N-acetyltransferase catalyzes the only synthetic reaction known to occur in the lysosome. To do this, the enzyme uses a cytosolic cofactor, acetyl–coenzyme A (acetyl-CoA).[3] Thus, its substrate and cofactor are separated by the lysosomal membrane. The mechanism by which it overcomes this spatial problem is controversial. One model suggests a ping-pong mechanism involving an initial acetylation reaction of the enzyme in the cytosol, followed by a transfer of the acetyl group to the intralysosomal heparin-α-glucosamine residue.[4,5] An alternative model proposes that the enzyme operates via a random-order ternary-complex mechanism; that is, there is no acetylated-enzyme intermediate.[6] The enzyme has proven very difficult to purify but is believed to be a dimer of 120 kDa subunits containing Asn-linked oligosaccharides; that is, concanavalin A-binding, with a pI of ~7.4, based on {^{14}C}acetyl-CoA labeling experiments.[4] It is also postulated that the 120-kDa subunit may only be the catalytic subunit of a protein complex whose other unidentified members are also needed for functionality.[4]

Mouse lysosomal membrane proteomics.—Lysosomes from mutant (Beige) and normal mice were isolated, and their membrane proteins were specifically extracted by methods similar to those reported for normal rat,[7] for a quantitative proteomics[8] study (authors' unpublished data). The cleavable ICAT (cICAT) reagent kit was obtained from Applied Biosystems. Procedures were performed according to the Applied Biosystems protocol. The pools of lysosomal membrane proteins from each group of mice were separately labeled, at their Cys residues, with isotopically heavy (^{13}C) or light (^{12}C) cICAT reagents and were digested by trypsin. The ICAT-Cys-containing peptides were isolated by avidin affinity chromatography. After cleavage, fractions were analyzed by online reverse-phase liquid chromatography-electrospray ionization–tandem mass spectrometry (MS/MS).

Construction of a mammalian expression vector containing the full-length mouse homolog of human TMEM76 (HGSNAT).—The full-length cDNA for mouse *hgsnat* was amplified by PCR from an IMAGE clone (BC024084), with use of primers 5′-CACCATGGGCGCGGGCCCGGCG-3′ and 5′-TCCGATTTTCCAAAACAGCTTCTTTTTATAGAGC-ACG-3′, and was inserted into the mammalian expression vector pcDNA3.1d TOPO/V5His (Invitrogen) in frame with the V5-epitope tag. The fidelity of the cloned DNA was confirmed by sequencing.

From the Research Institute, Metabolism Program, The Hospital for Sick Children (X.F.; H.Z.; S.Z.; R.D.B.; M.B.T.; J.W.C.; D.J.M.), and the Departments of Laboratory Medicine and Pathobiology (R.D.B.; D.J.M.) and Biochemistry (J.W.C.), University of Toronto, Toronto

Received June 14, 2006; accepted for publication July 27, 2006; electronically published August 23, 2006.

Address for correspondence and reprints: Dr. Don J. Mahuran, The Hospital for Sick Children, 555 University Avenue, Room 9146A, Toronto, Ontario M5G 1X8, Canada. E-mail: hex@sickkids.ca

Am. J. Hum. Genet. 2006;79:738–744. 0002-9297/2006/7904-0017$15.00

Table 1. Primers Used to Construct a Full-Length Human cDNA Encoding N-Acetyltransferase

Oligonucleotide Number	Nucleotide Sequence	Comment
103154	GCAGCGGGCAGGCAAG	16 nt of the sense strand at the 5′ end of exon 1
103379	GTGATAACAGCATTCAGATTTCCAGTAGACGGTC	34 nt of the antisense strand at the 3′ end of exon 2
103667	GACCTGCTACTGGAAATCTGAATGCTGTTATCAC	34 nt of the antisense strand at the 3′ end of exon 2 (complementary to oligonucleotide 103379)
EJ653	CCAGCAGCACATCTCAGTGGGAGCC	25 nt of the 3′ UTR antisense strand, very close to the coding region
103666	GGCCGTCGACGCAGCGGGCAGGCAAG	Adds an *Sal*I site to the 5′ end of oligonucleotide 103154
58961	GGGCGGCCGCCACATCTCAGTGGGAGCC	Adds an *Not*I recognition site at the 5′ end of EJ653

Transfection of HeLa cells and immunofluorescence staining and microscopy.—HeLa cells were plated onto glass coverslips and were transiently transfected using Lipofectamine 2000, according to the manufacturer's directions. The cells were fixed with 4% paraformaldehyde 24 h after transfection and were immunostained as described elsewhere.[9] LAMP1 (H4A3) antibody was from the Developmental Studies Hybridoma Bank at the University of Iowa. Anti-V5 rabbit-polyclonal antibody was from Sigma-Aldrich. Chicken anti-mouse Alexafluor 594 and chicken anti-rabbit Alexafluor 488 antibodies (Molecular Probes) were used as secondary antibodies for immunofluorescence. Confocal images were from an LSM 510 META laser scanning confocal microscope (Zeiss) equipped with a Plan-Apochromat 63 × (NA 1.4) oil DIC objective. Contrast and brightness of images were adjusted using Adobe Photoshop. The N-acetyltransferase activity was also determined for pools of HeLa cells transfected with the V5-His6–tagged and untagged mouse IMAGE clone and our full-length human *HGSNAT* cDNA (see below). These activities were compared with the endogenous levels of untransfected HeLa cells (see below).

Generation of a full-length human HGSNAT *cDNA*.—The cDNA of the human *HGSNAT* gene (which, when defective, causes MPS IIIC) was constructed through RT-PCR. Dimethyl sulfoxide (10%) was kept in both the RT and PCR, because of high GC-rich content at the 5′ region of the cDNA.

The total RNA was extracted from HeLa cells using RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. The first-strand cDNA was synthesized using Thermoscript (Invitrogen) as the reverse transcriptase. The RNA, primer EJ653 (table 1), and deoxyribonucleotide triphosphates were mixed and heated at 95°C for 5 min. After cooling to 60°C, the prewarmed mixture of enzyme and dithiothreitol was added, and the reaction was processed at 60°C for 50 min. The reaction was stopped by heating at 85°C for 5 min, and the RNA was digested by RNase H at 37°C for 1 h. The synthesized first-strand cDNA was used as template in the PCR or was kept at −20°C for further use.

The double-stranded fragment containing the last 34 nt of exons 2–18 was amplified by PCR. In this reaction, ProofStart (Qiagen) was used as the polymerase, the above first-strand cDNA as template, and oligonucleotides 103667 and EJ653 as primers (table 1). In the cycling program, the conditions used were 94°C for 30 s for denaturing, 60°C for 30 s for annealing, and 72°C for 1 min 40 s for extension. After 35 cycles, the final 72°C extension step was performed over a 10-min period. The double-stranded fragment containing exons 1–2 was amplified by the same technique, except that oligonucleotides 103154

Table 2. Exon-Flanking PCR Primers in the Human *HGSNAT* Gene

Region Spanned	Primer (5′→3′) Forward	Primer (5′→3′) Reverse	Amplimer (bp)
5′ Exon 1–5′ IVS1	CTTCTTTCCTGGCCAGCTC	GGAGACGCTGCGCTCGTA	390
3′ IVS1–5′ IVS2	GGAAGCAACTGTTCACACGA	AGCCCCAAAGCAAACAGAAC	249
3′ IVS2–5′ IVS4	GTCCTCCAGCAATCAACAG	TGCAAATGTGAAAATCAGCAG	573
3′ IVS3–5′ IVS4[a]	GGTTGGAATACAGATTTGGAGA	TGGTGCTCATCAATGACCTAA	260
3′ IVS4–5′ IVS5	CTTCCCCTCACTGGTTTTCA	GGTGACAGCTGCCCCTTC	249
3′ IVS5–5′ IVS6	GGAAACTATGCTGTCAGAGTTGA	CAACCTGGCCTTCCTCAATA	393
3′ IVS6–5′ IVS7	GGGCCACAGAGCGAGACT	TGGAGAGGCTGCAGAGTTTA	400
3′ IVS7–5′ IVS8	GCATGATTGCACCTTGTGAG	GTGATTCTCCTGCCCTGAAG	250
3′ IVS8–5′ IVS9	CCTCCCCTGGGTTTACTTTC	TGGGTCAAGTATGAGAACTATCCA	297
3′ IVS9–5′ IVS10	CCCATGGGGCTATATTCTGA	AATGCTCACAGTTAGACAACCA	490
3′ IVS10–5′ IVS11	TGTCTTCCCCTTTAGGAAACAA	TTAAGCCTGGGATGTTGAGG	489
3′ IVS11–5′ IVS13	CCAAAAAGGCCAAGAAATGA	GGCATGGTTTAACGTTTGCT	1,155
3′ IVS12–5′ IVS13[a]	AGCGATTCTCCTGCATTAGC	GGCATGGTTTAACGTTTGCT	492
3′ IVS13–5′ IVS14	GGAGTGTATTCAGGTTTGTATTTGG	GGCCCATAGCACAAGAGAGA	296
3′ IVS14–5′ IVS17	CCCTGTAATCCCAGCACCTA	CTAGGGCAGTGGCTCAAGAC	1,242
3′ IVS17–5′ end of the 3′ UTR (exon 18)	CTGGTTTCAAGAATTAAATAGATGG	CAGTGTCTTCCCCCAGAGTC	397

[a] Additional PCRs were done to verify the mutations.

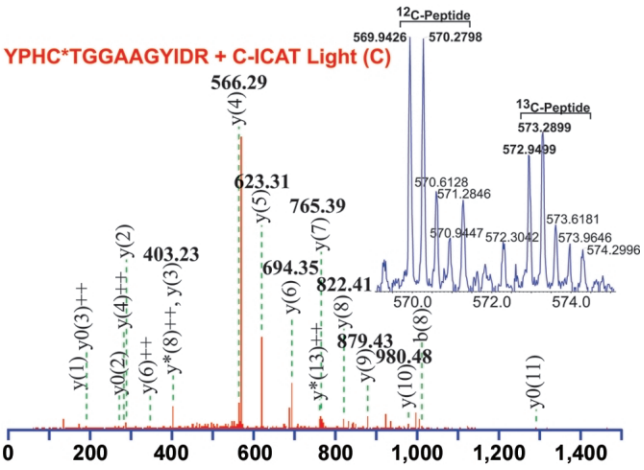

Figure 1. Identification of D8Ertd354e by MS/MS. The lysosomal membrane proteins from mutant (Beige) and normal mice were labeled with either isotopically heavy (^{13}C) or light (^{12}C) cICAT reagent, were digested with trypsin, and were affinity purified. The upper right inset is the MS pattern of a pair of these labeled peptides. The main spectrogram shows the MS/MS fragmentation pattern for the Cys-ICAT light–labeled peptide. The obtained peptide sequence is shown at the upper left. *C indicates the labeled Cys residue.

and 103379 (table 1) were used as primers and 72°C for 30 s was used for extension.

The full-length double-stranded cDNA of the *HGSNAT* gene was amplified by the same technique, except that the 72°C extension step was for 2 min. In this PCR, the above two overlapping PCR fragments—instead of the first-strand cDNA—were used as the template, and oligonucleotides 103666 and 58961 were used as primers (table 1). Finally, the amplified full-length cDNA was subcloned into the expression plasmid pCMVsport6 (Invitrogen) through its *Sal*I and *Not*I sites. The sequences of the two fragments and the final full-length cDNA were confirmed by automatic sequencing.

Transfections and expression of N-acetyltransferase in patient fibroblasts.—Fibroblasts from patients with MPS IIIC were transfected by electroporation (Human Dermal Fibroblast Nucleofector Kit [Amaxa]), according to the manufacturer's instructions. In brief, ~10^6 cells were suspended in 100 μl of nucleofector solution and 2 μg of plasmid DNA containing our V5-His6–tagged or untagged mouse homolog (positive control), or a candidate full-length human cDNA-encoding N-acetyltransferase was added, and electroporation was done using program U-23. Electroporated cells were then replated and harvested 20 h after transfection. Levels of N-acetyltransferase and β-hexosaminidase (Hex) were then determined from whole-cell homogenates. Hex activities were used to normalize the activities obtained from the N-acetyltransferase, HGSNAT.

Assay for N-acetyltransferase activity.—The method we used to assay N-acetyltransferase activity was based on the original work of Voznyi et al.[10] This is a coupled assay with use of the substrate 4-methylumbelliferyl-β-D-glucosaminide (MU-βGlcNH$_2$). Although heparan sulfate contains only α-glucosamine residues, Voznyi et al.[10] found that N-acetyltransferase is active toward either the α- or β-anomer. In the presence of acetyl-CoA, N-acetyltransferase converts MU-βGlcNH$_2$ to MU-βGlcNAc, which is a substrate for Hex, which results in the production of free fluorescent MU. If whole-cell homogenates are being assayed, their endogenous Hex is sufficient for the assay. In brief, cells were washed (in 2 × PBS), were resuspended in 0.2 M phosphate/0.1 M citrate buffer (pH 5.5) containing 0.2% heat-inactivated BSA, and were freeze-thawed six times. Then, 10 μl of the homogenate was mixed with 10 μl of 12 mM

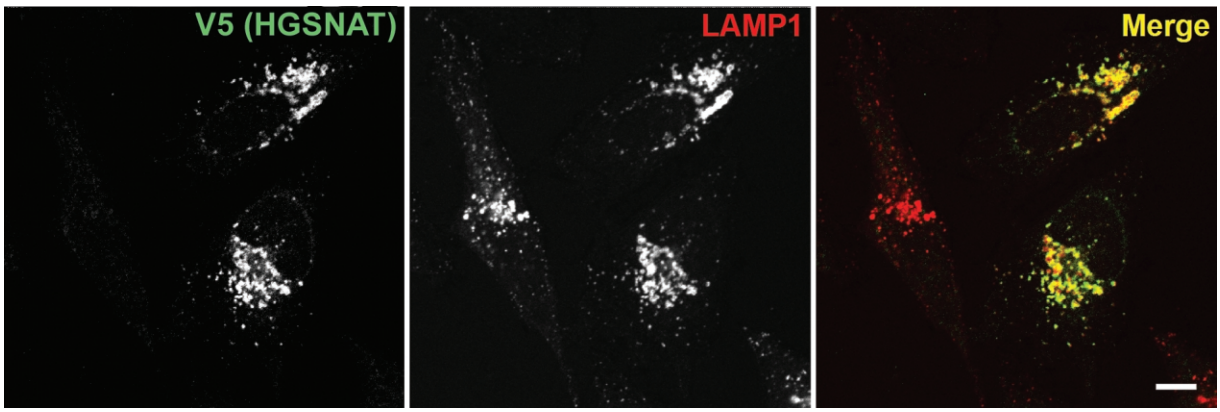

Figure 2. The epitopically expressed mouse D8Ertd354e (N-acetyltransferase) localizes to lysosomes in HeLa cells. C-terminal V5-epitope–tagged mouse N-acetyltransferase cDNA was transiently expressed in HeLa cells. The fixed cells were processed for indirect immunofluorescence confocal microscopy. N-acetyltransferase was detected with V5-antibody (*green*), whereas lysosomes were detected with LAMP-1 (H4A3) antibody (*red*). The merged images show substantial overlap between the signals (*yellow*), which indicates colocalization. Bar = 10 μm.

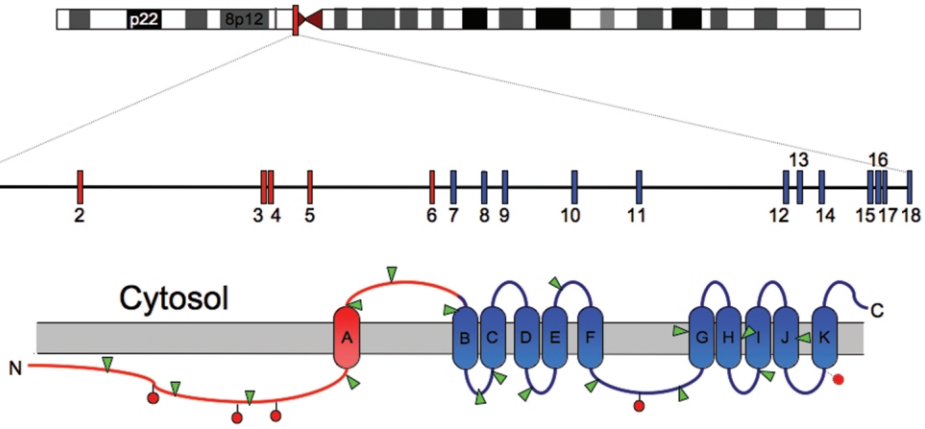

Figure 3. Chromosomal localization and structure of the human *HGSNAT* gene and the predicted topology of the enzyme it encodes, N-acetyltransferase, with its 11 TMDs (A–K) after the cleavage of the signal peptide. The exons and their encoded protein domain, shown in red, are unique to higher metazoans. Red circles on the protein diagram represent the positions of putative Asn-linked oligosaccharide sites, and the green triangles delineate the sections of the protein encoded by each of the 18 exons in *HGSNAT*.

acetyl-CoA and 10 μl of 3 mM MU-βGlcNH$_2$ in 0.2 M phosphate/0.1 M citrate buffer (pH 5.5) containing 0.25% Triton X-100. The reaction mixture was incubated at 37°C for 17 h. Reaction was terminated by adding 200 μl MAP (pH >10), and the fluorescence of the reaction-product, MU, was read. The endogenous Hex activity was determined separately, as described elsewhere,[11] to normalize the N-acetyltransferase activity

Mutational analyses.—Fibroblasts established from two unrelated patients (4172 and 8861)—who are from the same rural area in eastern Canada and who received biochemical and clinical diagnoses of MPS IIIC—were analyzed for mutations in their *HGSNAT* genomic DNA and cDNA. The genomic structure of the *HGSNAT* gene was established by the sequence of our cDNA clone, combined with mouse full-length cDNA and the UCSC Genome Bioinformatics databases. The genomic DNA from the patients' fibroblasts and one control cell line were isolated using DNeasy Tissue Kit (Qiagen). We amplified all exons and intron-exon boundaries in *HGSNAT*. PCR primers were designed using the Primer3 program (table 2). PCR products were purified, either by QiaQuick columns or from an excised gel band, with use of the QIAquick Gel Extraction Kit (Qiagen) and were bidirectionally sequenced with BigDye Terminator V3.1 cycle sequencing kit and 3100 Genetic Analyzer (Applied Biosystems) by ACGT Corp. We amplified the cDNA from the two patients and a control individual, using oligonucleotides 5′-ATGGATCAGGCTTTGCTACTCATTC-3′ and 5′-CTCCTCCATAATTGACAAAGACC-3′ as primers. The product, which spanned exons 2–8, was purified and sequenced. We also confirmed the presence of the splice-site mutation (see below) by restriction digestion of a 574-bp genomic PCR fragment, generated using oligonucleotides 5′-TGGTGCTCATCAATGACCTAA-3′ and 5′-GAGTCCAAAAGCAGGGAAGC-3′ as primers, with *Bsi*WI (data not shown), performed according to the manufacturer's instructions (New England Biolabs).

Northern-blot analysis.—Northern blotting was done on a Human Multiple Tissue Northern Blot (Clontech). The cDNA probe for human *HGSNAT* was PCR amplified from the cDNA with oligonucleotides 103667 and EJ653 as primers (see above). The probe for human actin was supplied by the manufacturer. The Multiple Tissue Northern Blot was probed with ^{32}P-cytidine–labeled cDNA, according to the manufacturer's directions.

Identification of the mouse homolog of human N-acetyltransferase.—As part of an unrelated proteomics study comparing mutant (Beige) and normal mouse lysosomal membrane proteins, one pair of ICAT-labeled–Cys-containing tryptic peptides, detected by MS (fig. 1 [inset]), was identified by MS/MS as D8Ertd354e (fig. 1), the mouse homolog of human TMEM76. Since this protein had no known function or identifiable functional domain, originated from lysosomal membranes, and its human homolog mapped to 8p11.1, we considered TMEM76 a strong candidate for the elusive MPS IIIC enzyme. Several apparently full-length mouse ESTs (mRNAs) were found in the Human Genome Browser Gateway database (position: human chromosome 8 43114000–43178000). One, BC024084 (GeneID 52120), was available as an IMAGE clone (5101064).

The IMAGE clone encoding mouse D8Ertd354e was transiently transfected into human MPS IIIC fibroblasts. The N-acetyltransferase activity in pools of transfected MPS IIIC cells was ~200-fold higher than in untransfected cells. A V5-His6 epitope tag was placed at the C-terminus of the IMAGE clone and was similarly expressed in HeLa and human MPS IIIC fibroblasts. Immunofluorescence microscopy of transfected HeLa cells demonstrated that the protein product colocalized with LAMP-1 (fig. 2), and N-acetyltransferase assays of both MPS IIIC fibroblasts (~100-fold increase) and HeLa cells (~2-fold increase) demon-

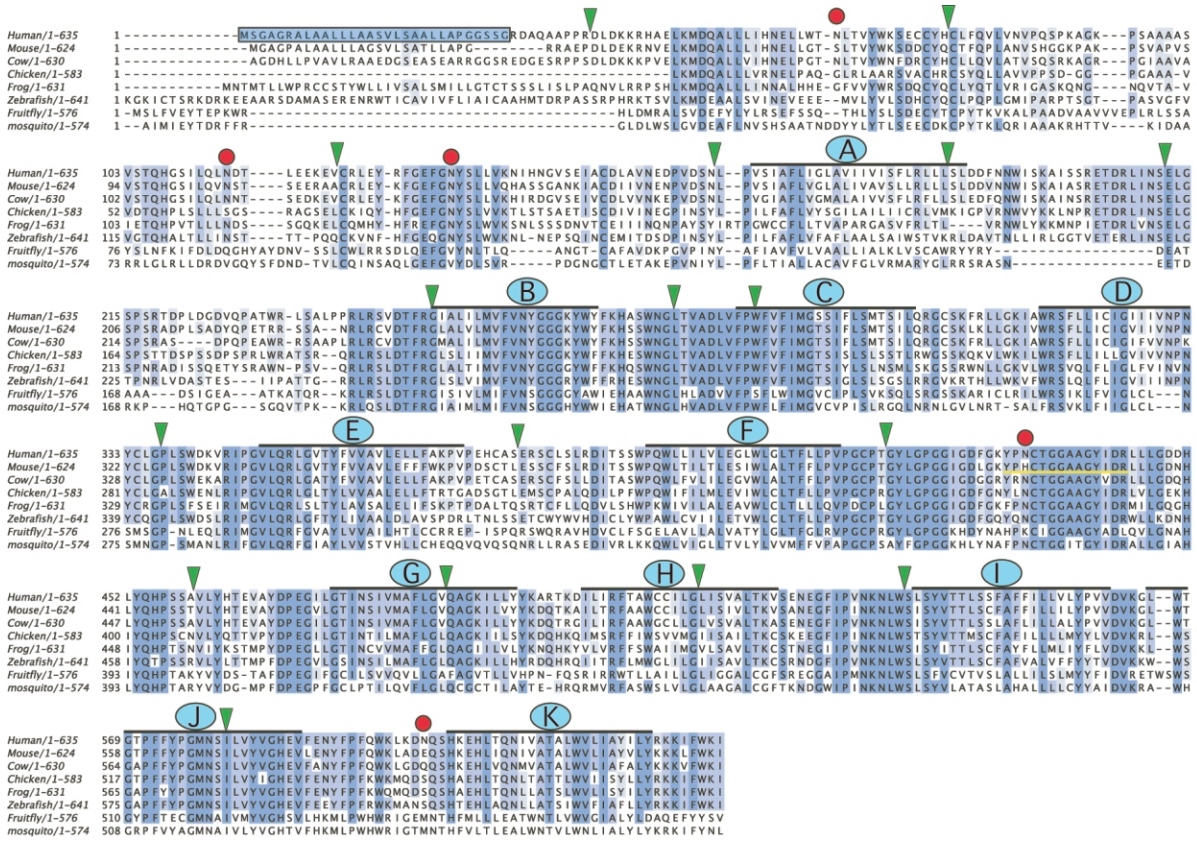

Figure 4. Aligned primary sequences of N-acetyltransferase from higher metazoans. The predicted human 30-residue, cleavable signal peptide is shown in a cyan-filled rectangle. Red circles and green triangles represent putative Asn-linked oligosaccharide sites and delineate the portions of the protein encoded by each of the 18 exons of the gene, respectively. Black bars indicate the 11 predicted TMDs (A–K), and the yellow underlined sequence in the mouse is the tryptic peptide we initially identified by MS/MS (fig. 1).

strated that the tag did not substantially interfere with function. These data also suggest that (*a*) the protein may not be a subunit in a multimeric enzyme complex, since many such multimers cannot form if one of their subunits is supplied through the transfection of cells with a cDNA derived from another species—for example, human Hex A (α-β heterodimer)[12]—and (*b*) the C-terminus of the protein likely faces the cytosol, since epitope tags usually do not survive inside lysosomes.[13]

Identifying the full-length human cDNA homolog of mouse D8Ertd354e (N-acetyltransferase) and deducing the human gene structure.—Comparison of the functional mouse sequence with existing human sequences identified it as XP_372038, first seen at the National Center for Biotechnology Information (NCBI) database on January 22, 2004. However, this sequence and two other revisions lacked the correct full-length N-terminus, as defined by the mouse sequence. On March 1, 2006, the fourth revision, XP_372038.4 (Gi:89028035), was added to the database. That version contained an N-terminus convincingly similar to the mouse to be considered full length. Using these data and comparing them with genomic sequences from contig NC_000008.9, we deduced that the *TMEM76* gene, renamed "*HGSNAT*" (Gene ID 138050), is 62.4 kb in length and contains 18 exons (fig. 3). In the course of mutational analyses (see below), we confirmed the exon- and intron-exon–junction sequences in the database by PCR (table 2).

Just as full-length human ESTs were apparently difficult to obtain, we were unsuccessful in generating a full-length cDNA by a single RT-PCR (see above). The 5′ section of the human transcript is very GC rich, ~80%, which may explain these difficulties. We transiently expressed our full-length human construct in MPS IIIC fibroblasts and HeLa cells and observed ~100- and ~2-fold increases, respectively, in N-acetyltransferase activity, as compared with untransfected cells of the same type.

Characterization of the human N-acetyltransferase protein.—The deduced C-terminal half of *HGSNAT* (exons 7–18) is highly conserved across species, including plants and bacteria. However, the N-terminus, extending to transmembrane domain (TMD) B, is found only in metazoans (figs. 3 and 4). Neither of these regions has homology with other functional domains identified elsewhere, including those from proteins known to bind CoA or acetyl-CoA. Thus, the *HGSNAT*-encoded N-acetyltransferase represents

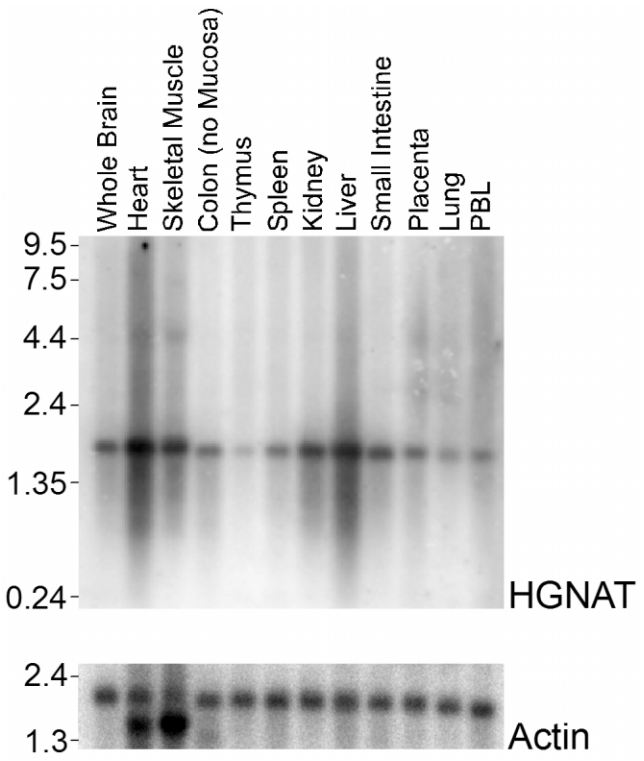

Figure 5. Northern-blot analysis of the *HGSNAT* transcripts in various human tissues. PBL = peripheral-blood leukocytes.

the human member of a new family of enzymes. The deduced primary sequence of human N-acetyltransferase predicts a 635-aa protein containing a 30-aa, cleavable, N-terminal signal peptide and 11 TMDs (fig. 3), as predicted by TMHMM Server v.2.0. This suggests an orientation with the 135-aa N-terminal domain inside the endoplasmic reticulum (ER) and, ultimately, the lysosome and the short 8-aa C-terminal domain facing the cytosol (consistent with the survival of the C-terminal V5-His6 epitope tag; see above). After signal-peptide cleavage, the predicted Mr and pI are ~68,000 and 8.4, respectively. There are also five consensus sites for Asn-linked glycosylation. On the basis of our topology predictions, all of these sites would be synthesized inside the lumen of the ER and thus would be available for glycosylation (fig. 3). If N-linked oligosaccharides were present at all five sites, they could add another ~10 kDa to the protein's apparent molecular mass. The pI of the protein could also be lowered and the apparent Mr further raised if any of these oligosaccharides were to have a high sialic acid content—for example, the polylactosaminoglycans found on LAMP-1 and LAMP-2.[14]

Characterization of mutant HGSNAT *alleles in two MPS IIIC–affected patient cell lines*.—Mutational analyses were initially performed by RT-PCR, oligonucleotide 103379, and oligonucleotide EJ653 (table 1) (nucleotides 189–1946). As compared with normal samples, samples from patient 1 contained only a lower-weight cDNA band, and the cDNA from patient 2 separated into two main bands. One corresponded to the lower weight band seen in patient 1, and the other was of apparent normal length. Further PCRs indicated that the size difference occurred between nucleotides 189 (exon 2) and 822 (exon 8). The cDNA bands detected in both patients' samples were of diminished intensity (data not shown). Results of sequencing of the cDNA from patient 1 indicated that it lacked exon 4, which resulted in the generation of a premature stop codon shortly downstream of the new exon 3/exon 5 splice site (data not shown). Premature stop codons are often associated with unstable mRNA.[15] Sequencing of each exon and exon-intron junction was undertaken to confirm the sequence present in the NCBI database and to identify the mutations in the two patients' cells. Automated sequencing of exon 4 and the beginning of intron 4 detected a G→A substitution in the first nucleotide of intron 4 of the *HGSNAT* gene from patient 1 (data not shown). Since a GT sequence at the beginning of each intron is required for splicing,[16] this substitution explains the missing exon 4 sequence in the patient's cDNA (data not shown). This substitution also interrupts a *Bsi*WI restriction site (CGTACG), which allows us to confirm the sequence change by amplifying and digesting a 574-bp PCR fragment spanning the mutation. These data confirm that patient 1 is homozygous for the splice-junction mutation and demonstrate that patient 2 is heterozygous for the same mutation (data not shown). Further sequencing of the genomic DNA from patient 2 revealed the expected IVS-4 +1G→A substitution and the emergence of two overlapping sequences in exon 13 (data not shown). The latter observation is explained by the insertion of a single G after nucleotide A1344, resulting in a frameshift after Gly448 and the generation of a premature stop codon 21 codons downstream.

The determination of the size and tissue distribution of HGSNAT *transcripts*.—Transcripts of HGSNAT were analyzed by northern blotting. These data revealed a major mRNA species of ~2.1 kb that is present in all tissues examined (fig. 5) but is highest in heart, skeletal muscle, liver, and kidney. The database predicts a transcript size of 5.1 kb that includes a large 3′ UTR of ~3 kb, encoded by exon 18. Such a large exon is rare, and our northern blot suggests that it may be present only as a minor species—for example, in heart and skeletal muscle (fig. 5).

Acknowledgments

This work was supported by Canadian Institutes of Health Research grants MOP67193 (to J.W.C. and D.J.M.) and MOP10435 (to D.J.M.).

Web Resources

Accession numbers and URLs for data presented herein are as follows:

Human Genome Browser Gateway, http://genome.ucsc.edu/cgi-bin/hgGateway

MPS Australia, http://www.mpssociety.org.au/table_of_diseases.htm

NCBI, http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi (for hypothetical human protein LOC138050 [accession numbers XP_372038 and XP_372038.4] and *HGSNAT* [GeneID 138050])

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for MPS IIIC)

Primer3, http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi

TMHMM Server, http://www.cbs.dtu.dk/services/TMHMM/ (for version 2.0)

UCSC Genome Bioinformatics, http://genome.ucsc.edu/

References

1. Kresse H, Von Figura K, Klein U (1978) A new biochemical subtype of the Sanfilippo syndrome: characterization of the storage material in cultured fibroblasts of Sanfilippo C patients. Eur J Biochem 92:333–339
2. Neufeld EF, Muenzer J (2001) The mucopolysaccharides. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic and molecular bases of inherited disease. McGraw-Hill, New York, pp 3421–3452
3. Klein U, Kresse H, von Figura K (1978) Sanfilippo syndrome type C: deficiency of acetyl-CoA:α-glucosaminide N-acetyltransferase in skin fibroblasts. Proc Natl Acad Sci USA 75:5185–5189
4. Ausseil J, Landry K, Seyrantepe V, Trudel S, Mazur A, Lapointe F, Pshezhetsky AV (2006) An acetylated 120-kDa lysosomal transmembrane protein is absent from mucopolysaccharidosis IIIC fibroblasts: a candidate molecule for MPS IIIC. Mol Genet Metab 87:22–31
5. Bame KJ, Rome LH (1986) Genetic evidence for transmembrane acetylation by lysosomes. Science 233:1087–1089
6. Meikle PJ, Whittle AM, Hopwood JJ (1995) Human acetyl-coenzyme A:alpha-glucosaminide N-acetyltransferase: kinetic characterization and mechanistic interpretation. Biochem J 308:327–333
7. Bagshaw RD, Mahuran DJ, Callahan JW (2005) A proteomic analysis of lysosomal integral membrane proteins reveals the diverse composition of the organelle. Mol Cell Proteomics 4:133–143
8. Heck AJ, Krijgsveld J (2004) Mass spectrometry-based quantitative proteomics. Expert Rev Proteomics 1:317–326
9. Bagshaw RD, Callahan JW, Mahuran DJ (2006) The Arf-family protein, Arl8b, is involved in the spatial distribution of lysosomes. Biochem Biophys Res Commun 344:1186–1191
10. Voznyi Ya V, Karpova EA, Dudukina TV, Tsvetkova IV, Boer AM, Janse HC, van Diggelen OP (1993) A fluorimetric enzyme assay for the diagnosis of Sanfilippo disease C (MPS III C). J Inherit Metab Dis 16:465–472
11. Tommasini R, Mahuran DJ, Lowden JA (1985) Discrimination of the isozymes of human placental hexosaminidase by kinetic parameter estimation. Can J Biochem Cell Biol 63:219–224
12. Brown CA, Mahuran DJ (1993) β-Hexosaminidase isozymes from cells co-transfected with α and β cDNA constructs: analysis of α subunit missense mutation associated with the adult form of Tay-Sachs disease. Am J Hum Genet 53:497–508
13. Hou Y, Vocadlo D, Leung A, Withers S, Mahuran D (2001) Characterization of Glu and Asp residues in the active site of human β-hexosaminidase B. Biochemistry 40:2201–2209
14. Carlsson SR, Fukuda M (1990) The polylactosaminoglycans of human lysosomal membrane glycoproteins lamp-1 and lamp-2: localization on the peptide backbones. J Biol Chem 265:20488–20495
15. Muhlrad D, Parker R (1994) Premature translational termination triggers mRNA decapping. Nature 370:578–581
16. Krawczak M, Reiss J, Cooper DN (1992) The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. Hum Genet 90:41–54