Analysis of the Biogenesis of Heparan Sulfate Acetyl-CoA:α-Glucosaminide N-Acetyltransferase Provides Insights into the Mechanism Underlying Its Complete Deficiency in Mucopolysaccharidosis IIIC

Received for publication, May 6, 2010, and in revised form, July 5, 2010. Published, JBC Papers in Press, July 22, 2010, DOI 10.1074/jbc.M110.141150

Stéphanie Durand1,3, Matthew Feldhammer1,3, Eric Bonneil², Pierre Thibault5, and Alexey V. Pshehotsky1,3,6,2

From the 1Department of Medical Genetics, CHU Sainte-Justine, and Departments of 2Biochemistry and 3Pediatrics, University of Montreal, Montreal H3T 1C5, the 4Institute of Research in Immunology and Cancer, University of Montreal, Montreal H3C 3J7, and the 5Department of Anatomy and Cell Biology, Faculty of Medicine, McGill University, Montreal H3A 2B2, Canada

Heparan sulfate acetyl-CoA:α-glucosaminide N-acetyltransferase (HGSNAT) catalyzes the transmembrane acetylation of heparan sulfate in lysosomes required for its further catabolism. Inherited deficiency of HGSNAT in humans results in lysosomal storage of heparan sulfate and causes the severe neurodegenerative disease, mucopolysaccharidosis IIIC (MPS IIIC). Previously we have cloned the HGSNAT gene, identified molecular defects in MPS IIIC patients, and found that all missense mutations prevented normal folding and trafficking of the enzyme. Therefore characterization of HGSNAT biogenesis and intracellular trafficking became of central importance for understanding the molecular mechanism underlying the disease and developing future therapies.

In the current study we show that HGSNAT is synthesized as a catalytically inactive 77-kDa precursor that is transported to the lysosomes via an adaptor protein-mediated pathway that involves conserved tyrosine- and dileucine-based lysosomal targeting signals in its C-terminal cytoplasmic domain with a contribution from a dileucine-based signal in the N-terminal cytoplasmic loop. In the lysosome, the precursor is cleaved into a 29-kDa N-terminal α-chain and a 48-kDa C-terminal β-chain, and assembled into active ~440-kDa oligomers. The subunits are held together by disulfide bonds between at least two cysteine residues (Cys123 and Cys434) in the lysosomal luminal loops of the enzyme. We speculate that proteolytic cleavage allows the nucleophile residue, His269, in the active site to access the substrate acetyl-CoA in the cytoplasm, for further transfer of the acetyl group to the terminal glucosamine on heparan sulfate. Altogether our results identify intralysosomal oligomerization and proteolytic cleavage as two steps crucial for functional activation of HGSNAT.

The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S4 and Table S1.

1 Both authors contributed equally to this work.
2 To whom correspondence should be addressed: Service de génétique médicale, CHU Sainte-Justine, 3175 Côte Ste-Catherine, Montréal, Québec H3T 1C5, Canada. Tel.: 514-345-4931/2736; Fax: 514-345-4801; E-mail: alexei.pchejetski@umontreal.ca.
3 The work was supported by research grants from the March of Dimes Foundation, the Sanfilippo Children’s Research Foundation, and Canadian Institutes for Health Research Grant MOP84430 (to A. V. P.), and operating grants from the National Science and Engineering Research Council (NSERC) (to P. T.). The Institute of Research in Immunology and Cancer receives infrastructure support funds from the Fonds de la Recherche en Santé du Québec.
4 The abbreviations used are: LMP, lysosomal membrane protein; MPS IIIC, mucopolysaccharidosis IIIC; ER, endoplasmic reticulum; 4MU-βGalNAc, 4-methylumbelliferyl-β-D-galactosaminide; bis-Tris, 2-(bis(2-hydroxyethyl)amino)-2-(hydroxymethyl)propane-1,3-diol; TEMED, N,N,N’,N’-tetramethylethylenediamine; MOPS, 4-morpholinopropanesulfonic acid; TAP, tandem affinity purification; CBP, calmodulin-binding peptide; PFO, perfluorooctanoic acid; HGSNAT, heparan sulfate acetyl-CoA:α-glucosaminide N-acetyltransferase.
Biogenesis of HGSNAT

HGSNAT, the only known lysosomal enzyme that is not a hydrolase, transfers an acetyl group from cytoplasmically derived acetyl-CoA to terminal N-glucosamine residues of heparan sulfate within the lysosomes (10). This reaction is critical for degradation of heparan sulfate because there is no enzyme that can act on the unacetylated glucosamine molecule. Since the discovery of the gene, 54 HGSNAT sequence variants have been identified including 13 splice-site mutations, 11 insertions and deletions, 8 nonsense, 18 missense mutations, and 4 polymorphisms (reviewed in Ref. 11 and also see the online data base). Although the HGSNAT transcripts with abnormal splicing, frame shifts, and premature stop codons are most likely rapidly degraded via the nonsense-mediated mRNA decay pathway (11, 12), the effect of missense mutations remained unclear until we expressed 17 mutant proteins in cultured human fibroblasts and COS-7 cells and studied their folding, targeting, and activity. We found that all studied missense mutations in HGSNAT caused misfolding of the enzyme, which was abnormally glycosylated and not targeted to the lysosome, but retained in the endoplasmic reticulum (ER). For at least 5 of the studied mutations (N273K, R344C, R344H, S518F, and S541L) the active conformation could be stabilized by the competitive inhibitor of HGSNAT, glucosamine, resulting in part of the enzyme pool being properly processed and targeted to the lysosomes (12). These results showed that lysosomal targeting of HGSNAT is essential for expression of enzymatic activity, consistent with the possibility that the enzyme may undergo in the lysosome additional post-translational modifications, obtain appropriate oligomeric structure and/or interact with other proteins necessary for its activation, but the direct evidence for this was absent. In the current study we applied site-directed mutagenesis, metabolic labeling, immunocytochemistry, size-exclusion chromatography, and electrophoresis to study the catalytic mechanism, oligomeric organization, and biogenesis of HGSNAT and show that intralysosomal oligomerization and proteolytic cleavage are essential for its functional activation.

EXPERIMENTAL PROCEDURES

Constructs and Site-directed Mutagenesis—The previously described HGSNAT-TAP construct (pCTAP-HGSNAT-long) was used as a template for site-directed mutagenesis (12). Mutations were introduced using the QuikChange Lightning kit (Stratagene) and primers designed with QuikChange Primer Design Program. The sequence of the primers is shown in supplemental Table S1. For all constructs except the pCTAP-HGSNAT-short, the coding sequence included an extra 84 base pairs on their 5′ end encoding 28 amino acids as for the sequence first described for the HGSNAT gene (6), but for clarity the used nomenclature reflects a unified numbering system based on the sequence of GenBank entries NM_152419.2 and NG_009552.1, as in Refs. 5, 11, 13 and 14. Throughout the text and figures, wild type refers to HGSNAT-TAP-long, unless indicated.

Cell Culture and Transfection—COS-7 cells were cultured in Eagle’s minimal essential medium supplemented with 10% (v/v) fetal calf serum (Wisent). Transfections were carried out in 10-cm dishes using Lipofectamine LTX (Invitrogen) as described by the manufacturer, and cells were harvested 42–48 h later unless specified otherwise in the figure legends. Inhibition of lysosomal processing was achieved by incubating cells with 50 mM ammonium chloride (NH₄Cl), starting 6 h post-transfection and lasting for 24 h. Inhibition of de novo protein synthesis was then achieved by replacing the NH₄Cl-supplemented medium with fresh medium containing 7 μM cycloheximide (Sigma) and culturing the cells for up to 72 h.

Enzyme Assays—N-Acetyltransferase enzymatic activity was measured using the fluorogenic substrate 4-methylumbelliferyl β-d-glucosaminide (4MU-βGlcN, Moscerdam, Rotterdam, The Netherlands), as previously described by He et al. (15). The reaction mixture containing 5 μl of cell homogenate, 5 μl of 10 mM acetyl-CoA, and 5 μl of 3 mM 4MU-βGlcN in McIlvain buffer (100 mM sodium citrate, 200 mM sodium phosphate, pH 5.7) was incubated at 37 °C for 3 h. The reaction was terminated by adding 1.98 ml of 0.5 M Na₂CO₃/NaHCO₃, pH 10.7, and fluorescence was measured and used to calculate the specific activity. The protein concentration was measured according to the method of Bradford (16) using a commercially available reagent (Bio-Rad).

Western Blotting—Cell homogenates were sonicated and boiled for 3 min in lithium dodecyl sulfate sample buffer (Invitrogen) in the presence or absence of 25 mM DTT. Proteins were resolved by SDS-polyacrylamide gel electrophoresis using either precast NuPAGE 4–12% bis-Tris gels (Invitrogen, Figs. 1, 2B, and 6) or casted 8% bis-Tris polyacrylamide gels (see Ref. 17 and Figs. 2C and 3–5) and electrotransferred onto a nitrocellulose membrane. Both resolving and stacking layers of bis-Tris polyacrylamide gels were prepared using 1.25 M bis-Tris HCl, pH 6.8, and a 30% (v/v) solution of acrylamide/bis-acrylamide (29:1). 10% (w/v) ammonium persulfate and TEMED were resolved by SDS-polyacrylamide gel electrophoresis using the method of Bradford (16) using a commercially available reagent (Bio-Rad).

Pulse-Chase Experiments—Pulse-chase experiments were performed using cells transfected with constructs pCTAP-HGSNAT-long, pCTAP-HGSNAT-short, and pCTAP-HGSNAT-M29A as described above. Twenty-four hours post-transfection, medium was removed and cells were placed in 5 ml of starvation medium (Met/Cys-free DMEM) for 1 h. Cells were then radiolabeled for 2 h with 250 μCi of [35S]Met and [35S]Cys (PerkinElmer Life Sciences) dissolved in the same medium. After two washes with PBS, the cells were chased for the indicated times with 5 ml of Eagle’s minimal essential medium supplemented with 2 mM cold Met and 2 mM cold Cys. Cell pellets were stored at −20 °C until all time points were collected. Metabolically labeled HGSNAT was affinity purified from cell pellets as follows. Pellets were resuspended in 1 ml of lysis buffer (40 mM Tris-HCl, 300 mM KCl, pH 7.5, 0.1% Nonidet P-40, 1 mM PMSE) supplemented with Sigma protease inhibitor mixture (P8340, 10 μl/ml of cell suspension). The homogenates

31234 JOURNAL OF BIOLOGICAL CHEMISTRY
were sonicated for 5 s and gently shaken at 4 °C overnight to solubilize proteins. The suspension was then centrifuged at 13,000 × g for 30 min. The supernatant was first passed through 0.2 ml of avidin-agarose resin (Sigma A9207), then affinity purification of TAP-tagged HGSNAT was performed using streptavidin and calmodulin resins (Stratagene) according to the manufacturer’s protocol. Elution from calmodulin resin was done by boiling in 1 resin volume of lithium dodecyl sulfate sample buffer (Invitrogen) containing 25 mM DTT. Eluates were resolved by SDS-PAGE; gels were fixed in 30% MeOH, 10% acetic acid and incubated in 10 gel volumes of Fluoro-Hance autoradiography enhancer (Research Products International) prior to drying. The dried gels were exposed to a phosphor storage screen for 1 week and then scanned using a Storm 840 molecular imager (GE Healthcare).

Confocal Immunofluorescence Microscopy—Immortalized control human skin fibroblasts or COS-7 cells were transfected with pCTAP-HGSNAT-long or plasmids coding for the HGSNAT mutants as described above. Forty-two-hour post-transfection, cells were washed with ice-cold PBS, fixed with 4% paraformaldehyde, 4% sucrose in PBS for 5 min, and then rinsed 3 times with PBS. Cells were permeabilized by 0.25% Triton X-100 for 10 min and blocked for 1 h in 3% horse serum and 0.1% Triton X-100. Cells were either co-stained with rabbit anti-CBP (Immunochemistry Consultants Laboratory; 1:400) and mouse monoclonal anti-calnexin (Millipore; 1:250) antibodies in 3% horse serum or with anti-CBP antibodies and mouse monoclonal antibodies against human LAMP-2 (Developmental Studies Hybridoma Bank; 1:100). Cells were then counterstained with Oregon Green 488-conjugated anti-rabbit IgG antibodies and Texas Red-conjugated goat anti-mouse antibodies (Molecular Probes; 1:1000). Slides were studied on a Zeiss LSM510 inverted confocal microscope. Images were processed using the LSM image browser software (Zeiss) and Photoshop (Adobe).

Active Site Labeling with [14C]Acetyl-CoA—Cell pellets from COS-7 cells transfected with either pCTAP-HGSNAT-long or pCTAP-HGSNAT-H269A were resuspended in McIlvain buffer (see above), sonicated, and incubated in the presence of 2 mM [acetyl-1-14C]acetyl-CoA (American Radiolabeled Chemicals) and 8 mM cold AcCoA at 37 °C for 30 min. Reactions were then stopped by precipitating the protein with 10% TCA. Pellets were washed with acetone and resuspended in Laemmli buffer (62 mM Tris-HCl, pH 6.8, 12.5% glycerol, 1% SDS, 0.05% bromophenol blue, 25 mM DTT). 14C-Labeled proteins were separated by SDS-PAGE. After Coomassie Brilliant Blue staining of the gel, the lanes were cut into 23 slices and analyzed for radioactivity via scintillation counting.

Mass Spectrometry Analyses—HGSNAT was affinity purified from COS-7 cells transfected with the pCTAP-HGSNAT-long construct as described above. The affinity purified HGSNAT protein extract was separated by SDS-PAGE, stained with colloidal silver, and gel bands were excised and digested with trypsin and Glu endopeptidase (Promega). The tryptic peptides were analyzed by nano-liquid chromatography mass spectrometry (nano-LC-MS) using an Eksigent nano-LC system (Eksigent, Dublin CA) interfaced to an LTQ-Orbitrap mass spectrometer (ThermoFisher Scientific, Waltham, MA) via a nanoelectrospray ionization source. LC separations were performed using custom made C18 pre-column (5 mm × 300 μm inner diameter, Jupiter 3 μm, C18) and an analytical column (10 cm × 150 μm inner diameter, Jupiter 3 μm C18). Data-dependent acquisition mode was enabled and each Orbitrap survey scan (resolution 60,000) was followed by three MS/MS scans on the LTQ linear ion trap mass spectrometer. The normalized collision energy was set to 25%. Mass calibration was used as an internal lock mass (protonated (Si(CH3)2O))6+; m/z 445.12057) and provided mass accuracy within 5 ppm for all nano-LC-MS/MS experiments. The centroided MS/MS data were merged into single peak-list files and searched with the Mascot search engine version 2.10 (Matrix Science, UK) against the mouse HGSNAT sequence using Glu endopeptidase, trypsin, and semi-tryptic cleavage sites. Parent ion and fragment ion mass tolerances were both set at ±0.02 and 0.4 Da, respectively. Assignment of His acetylation sites were validated through manual inspection of relevant MS/MS spectra.

Size Exclusion Chromatography—Proteins from transfected cells were solubilized as described for the pulse-chase experiments and the supernatant was loaded onto a FPLC Superdex 200 column (total volume 25 ml) equilibrated with 10 mM Tris, pH 7.4, 0.15 M NaCl, 0.1% Triton X-100, and 0.02% NaN3. The column was calibrated with thyroglobulin (669 kDa), ferritin (440 kDa), catalase (240 kDa), aldolase (158 kDa), and chymotrypsinogen A (25 kDa). The elution was performed at a flow rate of 0.5 ml/min and 0.5-ml fractions were collected and assayed for N-acetyltransferase activity. The assay mixture contained 50 μl of fraction, 10 μl of McIlvain buffer (see above), 5 μl of 10 mM acetyl-CoA, 5 μl of 3 mM 4MU-βGlcN in McIlvain buffer, and 2.5 μl of total soluble glycoprotein fraction from human placenta (19). The incubation was performed at 37 °C for 1 h and terminated by the addition of 1.93 ml of 0.5 mM Na2CO3/NaHCO3, pH 10.7.

Perfluorooctanoic Acid PAGE—Perfluorooctanoic acid (PFO) PAGE was performed as described by Ramjesingh et al. (18). Transfected cell pellets were suspended in PFO-PAGE buffer (50 mM Tris, 4% NaPFO (Oakwood Products), 10% glycerol, 0.0025% bromphenol blue, pH 8.0), sonicated, incubated at room temperature for 30 min, and centrifuged at 10,000 × g for 5 min. Approximately 20 μg of protein was loaded on each lane of freshly poured 7% Tris glycite gels containing 0.1% NaPFO, which were run at 1 mA overnight and then at 10 mA for another 1.5 h, with buffer containing 25 mM Tris, 192 mM glycine, and 0.1% NaPFO, pH 8.5. Transfer and Western blot were performed as described above.

RESULTS

HGSNAT Can Use Either of Its Two ATG Start Sites in Vitro and Retains Its N-terminal Signal Peptide—The cDNA sequence of human HGSNAT was previously predicted to contain two putative start sites at a distance of 84 bp from each other, each preceded by a Kozak sequence (5, 6). The first 43 nucleotides were based on genomic sequence (GenBank accession number NG_009552.1), nucleotides 44–202 were present in a spliced EST (DR000652.1) and nucleotides 203–1992 in mRNA sequence XM_372038.2. Direct amplification of the predicted full-length HGSNAT cDNA was, however, unsuc-
cessful (5, 6) and further publications adopted the nomenclature where the second ATG was considered to be the initiation site (11–14, 21). In the current work we experimentally assessed which of the predicted sites initiates the synthesis of a functional protein. Three constructs were prepared and expressed in COS-7 cells. The first two defined as “pCTAP-HGSNAT-long” and “pCTAP-HGSNAT-long-M29A” contained HGSNAT cDNA starting from the first ATG but in the latter the second ATG start site was changed to encode an Ala residue thus preventing simultaneous synthesis of both long and short forms of the enzyme. The third construct “pCTAP-HGSNAT-short” lacked the first 84 bp of the cDNA so protein synthesis could only start from the second ATG site. All constructs expressed HGSNAT with a C-terminal TAP tag consisting of a high affinity streptavidin-binding peptide and a CBP to allow purification of the recombinant protein using successively applied streptavidin-resin and calmodulin-resin affinity purification steps or its detection with anti-CBP antibodies (18, 19). All three constructs produced a full size enzymatically active protein, although cells transfected with the pCTAP-HGSNAT-long construct had higher levels of N-acetyltransferase activity and HGSNAT protein as compared with cells transfected with the short construct and the M29A mutant (Fig. 1, A and B). Short-term metabolic labeling with [35S]Met/Cys of cells transfected with the three constructs followed by a 0–30-min chase revealed that the pCTAP-HGSNAT-long construct was able to elicit synthesis from both ATG start sites as seen by the appearance of doublet bands of 77 and 74 kDa (Fig. 1C), whereas the cells transfected with pCTAP-HGSNAT-long-M29A and pCTAP-HGSNAT-short constructs expressed 77- and 74-kDa forms of the protein, respectively. The size difference of ~3 kDa observed between the long and short forms of HGSNAT corresponds to the calculated size of the 28-amino acid sequence fragment between the two putative start sites suggesting that the N-terminal signal peptide of the HGSNAT precursor is not cleaved upon entrance into the ER as previously proposed (5, 6). This finding was further supported by data obtained through tandem mass spectrometry analysis of the affinity purified 77-kDa form of HGSNAT protein, which identified tryptic peptides (R)AGMSGAGR, (R)ALAAALLAASVLSALLAPGSSGR, and (R)DAQAAPPRDLDK present in the putative signal peptide (supplemental Fig. SI).

Previously, we have shown that homogenates of the cells transfected with active wild type HGSNAT, but not of cells transfected with inactive mutants retained in the ER contained a C-terminal fragment of HGSNAT with a molecular mass of ~48 kDa (12). To understand whether this form represents a product of intralysosomal maturation of HGSNAT, we treated cells transfected with HGSNAT-expressing plasmid with 50 mM NH4Cl, a procedure known to inhibit lysosomal acidification and therefore the activity of lysosomal proteases (20, 21). Indeed, Western blot analysis showed that cells treated with 50 mM NH4Cl contained only a 77-kDa HGSNAT precursor and its 160-kDa dimer, whereas untreated cells contained the precursor and the 48-kDa band. Moreover, inhibition of lysosomal acidification by NH4Cl completely abolished HGSNAT activity, suggesting that the proteolytic cleavage is essential for the activation of the enzyme (Fig. 2, A and B). Lysosomal processing was restored following replacement of the medium containing NH4Cl with normal medium supplemented with 7 μM cycloheximide to inhibit de novo protein synthesis, resulting in the appearance of the 48-kDa band with a proportional increase in N-acetyltransferase activity (Fig. 2, A and B).

To confirm these results and understand whether the N-terminal part of HGSNAT remains associated with the mature protein, we performed 2-h metabolic labeling of the cells transfected with either pCTAP-HGSNAT-long or pCTAP-HGSNAT-short constructs with [35S]Met/Cys followed by a prolonged chase for up to 72 h. This analysis revealed that the 77-kDa precursor predominant after the 2-h pulse is gradually cleaved into 48- and 29-kDa protein chains (Fig. 2C) present as expected in a ~1:1 molar ratio. The 29-kDa chain presumably contains the N-terminal part of the enzyme as it is undetectable with the antibodies specific against the C-terminal CBP tag, whereas the 48-kDa form contains the C terminus and is detected by the antibodies (Fig. 1B). Interestingly, densitomet-
ric analysis revealed that processing of the “short” precursor form of HGSNAT starting from the Met at position 29 occurs faster than that of the “long” form starting from the first Met (not shown), however, it is not clear at this point whether this has any biological relevance because the cells transfected with pCTAP-HGSNAT-long-M29A and pCTAP-HGSNAT-short plasmids show similar levels of enzymatic activity (Fig. 1A).

To confirm the results of metabolic labeling and to approximate the position of the cleavage site we affinity purified the 77- and 48-kDa forms of HGSNAT, resolved them by SDS-PAGE, digested by combined action of Glu endopeptidase and trypsin, and analyzed tryptic peptides from both fragments by LC-MS/MS. We found (supplemental Fig. SI) that the 48-kDa fragment contained peptides from the C-terminal part of the protein but missed peptides (R)AAGMSGAGR, (R)AALLLAASVLASALLAPGGSSGR, (R)DAQAAPPR, (R)DAQAAPPRDLDK, (E)LKMDQALLIHNE, (K)MDQALLIHNE, (K)SECCYHCLFQVLNVPSKP, (E)FGNYSLLVK, and (R)LLLSLDFNNWISK from the N-terminal part of HGSNAT that were identified by the MS analysis of the 77-kDa precursor. In contrast, 29-kDa fragment contained N-terminal peptide (R)DAQAAPPRDLDK but none of the peptides in the C-terminal part (supplemental Fig. SI). Tryptic peptides from the first cytoplasmic loop of HGSNAT, (R)ETDRLINSELGSPSR and (R)TDPLDGDVQPATWR, were identified in both 48- and 29-kDa fragments. This together with the calculated Mr and predicted glycosylation pattern of HGSNAT suggested that proteolytic cleavage occurs within the end of the first and/or the beginning of the second luminal domain of the enzyme.

Cysteine Residues within the Lysosomal Luminal Domains Are Essential for Oligomerization and Activity of HGSNAT—Our previous observation that HGSNAT shows much higher Mr on Western blot under non-reducing conditions 4 suggested that its oligomeric structure may be supported by disulfide bonds between cysteine residues. To define which residues are important for HGSNAT oligomerization, we expressed HGSNAT mutants with cysteines changed to serines (supplemental Fig. SII). Cys75 was excluded for lack of conser-

4 S. Durand, unpublished data.
**Biogenesis of HGSNAT**

**A**

| HGSNAT sequence (621–635) | LIAYILYKKIPFKWI |
|--------------------------|------------------|
| del624−635               | LIA---------------|
| L622A_1,622A             | AAAAAAYKKIPFKWI  |
| I625A_1,626A             | LIAAYILYKKIPFKWI |
| Y627A                    | LIAYLARKIPFKWI   |
| HGSNAT sequence (204−209) | EDTAEL |
| L208A,1209A              | EDTAEL |

**B**

![Immunoblot analysis of HGSNAT.](image)

**C**

**D**

![Western blot analysis of the del624–635 mutant.](image)

**FIGURE 4. Identification of HGSNAT regions important for lysosomal targeting.** COS-7 cells were transfected with HGSNAT-TAP plasmids bearing mutations in the predicted targeting signals and studied by immunohistochemistry and confocal microscopy or analyzed for N-acetyltransferase activity and by Western blot. A, mutations introduced in HGSNAT sequence. B, transfected cells were fixed and stained with mouse monoclonal anti-LAMP-2 antibodies (red) and rabbit polyclonal anti-CBP antibodies (green). Slides were studied on a Zeiss LSM510 inverted confocal microscope. Magnification ×630. C, N-acetyltransferase activity in cell homogenates shown relative to the wild type (mean of three or four independent experiments; error bars represent 5D). D, representative Western blots of cell homogenates. NT, non-transfected cells.

**LysoSomal Targeting of HGSNAT Occurs via a C-terminal Signal YIL-YRKKIPFKWI**—Lysosomal targeting of membrane proteins occurs via adaptor protein-mediated pathways (reviewed in Ref. 23). The crucial step of this mechanism is the interaction of the µ-subunits of the adaptor protein complexes with either tyrosine-based (NPXY or YXXØ, where Ø represents a bulky hydrophobic residue consensus motifs), or dileucine-based ([DE]-XXX[L] and DXX[L]) consensus motifs in the cytosolic domains of lysosomal membrane proteins (24, 25). Because the consensus dileucine-based signal can be found in the first cytoplasmic loop of HGSNAT (amino acids 204–209) and the tyrosine-based signal in its cytoplasmic tail (amino acids 624–635), we used site-directed mutagenesis to identify residues involved in its targeting. In particular the putative crucial leucine and isoleucine residues in a predicted dileucine targeting motif were mutated to alanines (L208A/I209A), and those in the cytoplasmic tail were either mutated or deleted (Fig. 4A). Most importantly, deletion of 12 residues at the C terminus of HGSNAT (del624–635) resulted in the protein being localized to the plasma membrane, in contrast to the wild type enzyme that co-localizes with the lysosomal marker LAMP2 (Fig. 4B). The L208A/I209A mutant displayed both lysosomal and plasma membrane localization (Fig. 4B), whereas all other mutants showed lysosomal localization similar to that of the wild type enzyme (not shown).

Western blot analysis of the del624–635 mutant revealed the absence of the 48-kDa HGSNAT peptide chain and a complete lack of enzymatic activity (Fig. 4, C and D), which could be expected of a HGSNAT mutant that is mistargeted to the plasma membrane and does not undergo intralysosomal maturation. At the same time, the mutant enzyme showed proper glycosylation (not shown) suggesting that the deletion does not cause its misfolding in the ER. The L208A/I209A mutant displayed a reduced amount of the processed 48-kDa form, and was partially mistargeted to the plasma membrane, but had levels of enzymatic activity comparable with that of the wild type protein (Fig. 4, C and D). Mutants with alanine substitutions within the cytoplasmic tail all displayed enzymatic activity but showed unusually high variability between the parallel measurements and had a reduced level of the 48-kDa HGSNAT form on the Western blot.

**His**²⁶⁹ Is the Active Site Residue Accepting the Acetyl Group in the Process of Catalysis—Previously Bame and Rome (26) showed that HGSNAT is specifically inhibited by N-bromosuc-
cinimide (26), which allowed them to speculate that the enzyme active site contains a histidine residue presumably serving as an acceptor for the acetyl group (26–30). To verify this experimentally and to determine which His residue(s) is involved in the catalysis we performed site-directed mutagenesis of all histidine residues outside of the predicted transmembrane domains (supplemental Fig. III). All mutants were analyzed for enzymatic activity, banding pattern on Western blot, and intracellular localization. Of the 11 HGSNAT mutants three (H269A, H451A, and H605A) showed a complete loss of enzymatic activity (Fig. 5A). Two inactive mutants, H451A and H605A, also lacked the processed 48-kDa form and showed reduced molecular mass of the precursor (Fig. 5B) compatible with their underglycosylation and improper folding in the ER previously reported for the majority of HGSNAT missense mutations identified in MPS IIIC patients (12). Indeed both mutants colocalized with the ER marker calnexin indicating that they were retained in the ER (see representative image on Fig. 5C). In contrast, H269A was the only inactive mutant that showed intralysosomal processing on Western blot (Fig. 5B) and lysosomal localization (Fig. 5C), suggesting that the mutation affects its catalytic activity but not its folding. To confirm this we transfected COS-7 cells with plasmids expressing the wild type enzyme and H269A mutant and performed active site labeling by incubating the cell homogenates with 2 mM [14C]acetyl-CoA for 20 min. Labeled proteins were then precipitated by trichloroacetic acid and resolved by SDS-PAGE. After fixing, gel lanes were cut into 23 bands and analyzed using a scintillation counter. Lanes containing wild type HGSNAT had radioactive peaks in the areas of 60 and 80 kDa, which were completely absent in the case of the H269A mutant (Fig. 5D), further confirming that His269 undergoes acetylation and therefore represents the active site residue. Although the His269 containing tryptic peptide predicted to be extremely hydrophobic could never be identified by the MS/MS analysis of purified recombinant HGSNAT, we detected acetylation of two other His residues (His143 and His605) but this modification was constitutive and did not depend of the presence of the AcCoA in the reaction mixture (supplemental Fig. SIV).

Mature HGSNAT Is Assembled into 440-kDa Oligomers—Our previous data obtained by gel filtration analysis of semi-purified HGSNAT from human placenta suggested that the enzyme forms a high molecular weight protein complex (27). To confirm these results and to understand whether the detected form represents a HGSNAT oligomer or its complex with other proteins we analyzed cells expressing TAP-tagged HGSNAT by FPLC gel filtration. To understand whether the organization of the processed active form is different from that of the enzyme precursor, we also analyzed cells transfected with...
Biogenesis of HGSNAT

plasmids expressing the mutant del624–635, which lacks lysosomal targeting, and H269A, which lacks enzymatic activity but is correctly targeted and processed.

Wild type HGSNAT eluted as a single peak in a volume corresponding to a 440-kDa protein (Fig. 6A). A small amount (<10% of activity) also eluted at a void volume of the column and most probably represented protein aggregates. Western blot analysis of the elution fractions showed that the 440-kDa peak contained both the HGSNAT precursor and processed forms appearing as 77- and 48-kDa bands, respectively, under reducing conditions (the 29-kDa chain cannot be detected on Western blots because it is not recognized by the antibodies). Under non-reducing conditions we observe 160-kDa HGSNAT dimers together with a 120-kDa immunoreactive band always present in a ratio of 2:1. Our attempts to study the composition of a 120-kDa band by excising it from the gel and analyzing by the second SDS-PAGE under reducing conditions failed because once denatured under non-reducing conditions the enzyme looses its ability to be further reduced by DTT (data not shown). It is possible that the 120-kDa form that could be made of the 77-kDa precursor and the 48-kDa processed chain is produced exclusively in vitro as a result of incomplete maturation of HGSNAT. However, our previous data showed that treatment of HGSNAT purified from human placenta and rodent liver with [14C]acetyl-CoA resulted in radioactive labeling of the 120-kDa protein suggesting that this form is also present in vivo (27). The active site mutant H269A showed gel filtration and Western blot patterns similar to that of the wild type, whereas the del624–635 mutant was present only in aggregates (Fig. 6B). Both mutants are devoid of N-acetyltransferase activity (Figs. 4C and 5B) but H269A localizes to the lysosome and is processed, whereas del624–635 is directed to the plasma membrane and is not processed, suggesting that the active oligomers of the enzyme are formed in the lysosome. The results obtained by gel filtration were confirmed using PFO-PAGE analysis (18). This technique uses the PFO ability to solubilize and charge a protein for electrophoretic separation while maintaining its native oligomeric state. Analysis by PFO-PAGE of COS-7 cells expressing wild type HGSNAT and correctly targeted and processed active site mutant H269A shows bands of 440 and ~220 kDa. At the same time the mistargeted del624–635 mutant is assembled only in the form of high molecular weight aggregates that appear as the band in the upper part of the gel (Fig. 6C).

DISCUSSION

In the current study we characterize the sequence of events crucial for functional activation of HGSNAT. First, by either deleting or mutating the two predicted N-terminal Met residues of the enzyme, expressing mutants in COS-7 cells, and studying their activity and processing, we demonstrate that both sites can produce correctly targeted, functional enzymes. Interestingly, the shorter peptide starting from the second Met and missing 28 residues at the N terminus is processed 2-fold faster. The transfection of cells with the plasmid containing both start sites resulted in about a 2-fold higher level of HGSNAT activity and protein suggesting that the translation can simultaneously start from both sites. Our results show that

FIGURE 6. Oligomeric composition of HGSNAT active form. COS-7 cells were transfected with HGSNAT-TAP-short (WT), del624–635, or H269A plasmids. Protein extracts were analyzed by size exclusion chromatography on a FPLC Superdex 200 column or by PFO-PAGE. A, N-acetyltransferase activity in the elution fractions from the Superdex 200 column loaded with wild-type (WT) enzyme. Elution volumes of protein size standards are indicated above the graph: thyroglobulin (669 kDa), ferritin (440 kDa), catalase (240 kDa), aldolase (158 kDa), and chymotrypsinogen A (25 kDa). B, elution fractions were analyzed by Western blot using anti-CBP antibody, in reducing (+DTT) or non-reducing (−DTT) conditions for the WT and reducing conditions for the mutants. C, cell homogenates were suspended in PFO-PAGE sample buffer, separated on 7% PFO-PAGE, and analyzed by Western blot using anti-CBP antibody. Positions of the commercial molecular mass markers (260, 160, and 110 kDa), thyroglobulin (669 kDa), and ferritin (440 kDa) are indicated.
the signal peptide is not cleaved upon translocation into the ER, which is also observed for several lysosomal transmembrane proteins including the cation channel, TRP-ML1, which deficiency is responsible for mucolipidosis IV (31), receptor for lysosomal mannos-6-phosphate-independent targeting LIMP II, responsible for myoclonus epilepsy and glomerulosclerosis (32–34), and CLN3 protein, which deficiency causes neuronal ceroid lipofuscinosis type 3 (34). Also similarly to many lysosomal membrane proteins, the HGSNAT precursor is targeted to the lysosomes via the adaptor protein complex-mediated pathway. This involves tyrosine- and/or dileucine-based conserved amino acid motifs in the C-terminal 16-amino acid fragment, TALWVLIAYLRRKK, with help of the dileucine-based motif in the second cytoplasmic loop of the enzyme. Upon the arrival to the lysosome the precursor is cleaved into 2 protein chains, a 29-kDa N-terminal α-chain and a 48-kDa C-terminal β-chain. This processing is a prerequisite for the expression of catalytic activity, because inhibition of lysosomal acidification abolishes both processing and the enzymatic activity of HGSNAT. The MS/MS analysis of the purified 48-kDa HGSNAT fragment suggests that cleavage may happen within the second lysosomal luminal loop of the enzyme containing His269 residue (supplemental Figs. S1 and SIII). Our experiments identify this residue as the nucleophile of the HGSNAT active site because its conversion to Ala completely abolishes both activity and acetylation of HGSNAT while maintaining its correct targeting and processing. Our data therefore directly confirm the previous hypothesis of Bame and Rome (28) who showed that acetyl-CoA is not stable in the lysosomal environment and proposed that HGSNAT catalyzes the acetylation of heparan sulfate in the lysosome via a ping-pong (double displacement) transmembrane reaction, with covalent linkage of the acetyl group to the enzyme as the intermediate (10, 29, 30). By analogy with other lysosomal enzymes synthesized as inactive precursors and then activated in the lysosome by cleavage into a two-chain form (reviewed in Ref. 35) it is tempting to speculate that the cleavage may be necessary to provide access for the His269 residue in the active site of the enzyme to the substrate acetyl-CoA in the cytoplasm, but structural studies are necessary to confirm this directly. Interestingly, MS/MS analysis of the purified recombinant HGSNAT also detected acetylation of two other His residues (His143 and His665) but this modification is constitutive and does not depend of the presence of the AcCoA in the reaction mixture.

Gel filtration analysis combined with PFO- and SDS-PAGE revealed that the mature enzyme forms of the 440-kDa oligomers presumably containing six α- and six β-chains. The presence of both the processed HGSNAT and its precursor within the HGSNAT 440-kDa peak separated by FPLC gel filtration is likely caused by co-elution of the oligomers composed of processed HGSNAT and 77-kDa precursors. However, it is also possible that because the process of proteolytic maturation of HGSNAT takes up to 72 h some species of enzyme oligomers are composed of a precursor and proteolytically processed peptides. Peptide chains within the HGSNAT oligomers are held together by disulfide bonds. Site-directed mutagenesis reveals that two residues, Cys123 and Cys344, located in the first and fourth lysosomal luminal domains of HGSNAT are involved in its oligomerization and most likely form a disulfide bond with each other because conversion of either one of them into a Ser residue results in a dramatic reduction of the level of oligomers and appearance of monomers under non-reducing conditions. Because both gel filtration and PFO-PAGE analysis of the enzyme show 440-kDa oligomers as the predominant form of the enzyme we speculate that it may represent the native oligomeric organization of the enzyme resembling that of the potassium channel in insulin-secreting β-cells and other transporters where the transmembrane pore is formed by association of membrane-spanning domains from several monomers (36–38).

Altogether our current results identify intralysosomal oligomerization and proteolytic cleavage as two steps crucial for functional activation of HGSNAT. Previously we showed that all 17 missense mutations in HGSNAT identified in the MPS IIIC patients caused misfolding of the enzyme, which was not targeted to the lysosome but retained in the ER (12). All mutants were completely missing proteolytic cleavage, which we now show to be associated with activation of HGSNAT thus explaining the complete deficiency of HGSNAT in the patients’ cells and their severe clinical phenotype.

Acknowledgments—We thank Drs. Martin Hrebícek for stimulating discussions and Mila Ashmarina for critically reading the manuscript.

REFERENCES

1. Ruivo, R., Anne, C., Sagné, C., and Gassier, B. (2009) Biochim. Biophys. Acta 1793, 636–649
2. Dell’Angelica, E. C., Mullins, C., Caplan, S., and Bonifacino, J. S. (2000) FASEB J. 14, 1265–1278
3. Callahan, J. W., Bagshaw, R. D., and Mahuran, D. J. (2009) J. Proteomics 72, 23–33
4. Saftig, P., and Klumperman, J. (2009) Nat. Rev. Mol. Cell Biol. 10, 623–635
5. Fan, X., Zhang, H., Zhang, S., Bagshaw, R. D., Tropak, M. B., Callahan, J. W., and Mahuran, D. J. (2006) Am. J. Hum. Genet. 79, 738–744
6. Hrebícek, M., Mrázová, L., Seyranetepe, V., Durand, S., Roslin, N. M., Nosková, L., Hartmannová, H., Ivánková, R., Cízková, A., Poupetová, H., Sikora, J., Urinovská, J., Straneyček, V., Zeman, J., Lepage, P., Roquis, D., Verner, A., Auer, J., Seeley, C. E., Maire, I., Poorthuis, B. J., van de Kamp, J., van Dijck, O. P., Wevers, R. A., Hudson, T. J., Fujiwara, T. M., Majewski, J., Morgan, K., Knoch, S., and Pshezhetsky, A. V. (2006) Am. J. Hum. Genet. 79, 807–819
7. Bartsocas, C., Gröbe, H., van de Kamp, J. I., van Figura, K., Kresse, H., Klein, U., and Giesberts, M. A. (1979) Eur. J. Pediatr. 130, 251–258
8. Kakki, E. D., Muenzer, J., Tiller, G. E., Waber, L., Belmont, J., Passage, M., Izikowicz, B., Phillips, J., Doroshov, R., Walot, I., Höft, R., and Neufeld, E. F. (2001) N. Engl. J. Med. 344, 182–188
9. Valstar, M. J., Ruijter, G. J., van Dijck, O. P., Poorthuis, B. J., and Wijburg, F. A. (2008) J. Inherit. Metab. Dis. 31, 104–111
10. Bame, K. I., and Rome, L. H. (1985) J. Biol. Chem. 260, 11293–11299
11. Feldhammer, M., Durand, S., Mrázová, L., Boucher, R. M., Laframboise, R., Steinfeld, R., Wraith, J. E., Michelakakis, H., van Dijck, O. P., Ribeck, M., Knoch, S., and Pshezhetsky, A. V. (2009) Hum. Mutat. 30, 918–925
12. Feldhammer, M., Durand, S., and Pshezhetsky, A. V. (2009) PLoS One 4, 7434
13. Ruijter, G. J., Valstar, M. J., van de Kamp, J. M., van der Helm, R. M., Durand, S., van Dijck, O. P., Wevers, R. A., Poorthuis, B. J., Pshezhetsky, A. V., and Wijburg, F. A. (2008) Mol. Genet. Metab. 93, 104–111
Biogenesis of HGSNAT

14. Fedele, A. O., Filocamo, M., Di Rocco, M., Sersale, G., Lubke, T., di Natale, P., Cosma, M. P., and Ballabio, A. (2007) Hum. Mutat. 28, 523
15. He, W., Voznyi, YaV., Huijmans, J. G., Geilen, G. C., Karpova, E. A., Dudukina, T. V., Zaremba, J., Van Diggelen, O. P., and Kleijer, W. J. (1994) Prenat. Diagn. 14, 17–22
16. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
17. Graham, D. R., Garnham, C. P., Fu, Q., Robbins, J., and Van Eyk, J. E. (2005) Proteomics 5, 2309–2314
18. Ramjeesingh, M., Huan, L. J., Garami, E., and Bear, C. E. (1999) Biochem. J. 342, 119–123
19. Pshezhetsky, A. V., and Potier, M. (1993) Biochem. Biophys. Res. Commun. 195, 354–362
20. Swaney, D. L., McAlister, G. C., and Coon, J. J. (2008) Nat. Methods 5, 959–964
21. Coutinho, M. F., Lacerda, L., Prata, M. J., Ribeiro, H., Lopes, L., Ferreira, C., and Alves, S. (2008) Clin. Genet. 74, 194–195
22. Rath, A., Glidewick, M., Nadeau, V. G., Chen, G., and Deber, C. M. (2009) Proc. Natl. Acad. Sci. U.S.A. 106, 1760–1765
23. Bonifacino, J. S., and Traub, L. M. (2003) Annu. Rev. Biochem. 72, 395–447
24. Guarnieri, F. G., Arterburn, L. M., Penno, M. B., Cha, Y., and August, J. T. (1993) J. Biol. Chem. 268, 1941–1946
25. Pearse, B. M., Smith, C. J., and Owen, D. J. (2000) Curr. Opin. Struct. Biol. 10, 220–228
26. Bame, K. J., and Rome, L. H. (1986) J. Biol. Chem. 261, 10127–10132
27. Ausseil, J., Landry, K., Seyrantepe, V., Trudel, S., Mazur, A., Lapointe, F., and Pshezhetsky, A. V. (2006) Mol. Genet. Metab. 87, 22–31
28. Rome, L. H., Hill, D. F., Bame, K. J., and Crain, L. R. (1983) J. Biol. Chem. 258, 3006–3011
29. Bame, K. J., and Rome, L. H. (1986) Science 233, 1087–1089
30. Freeman, C., Clements, P. R., and Hopwood, J. J. (1983) Biochem. Int. 6, 663–671
31. Kiselevsky, K., Chen, I., Rbaibi, Y., Oberdick, D., Tjon-Kon-Sang, S., Shcheynikov, N., Muallem, S., and Soyombo, A. (2005) J. Biol. Chem. 280, 43218–43223
32. Berkovic, S. F., Dibbens, L. M., Oshlack, A., Silver, J. D., Katerelos, M., Vears, D. F., Lüllmann-Rauch, R., Blanz, J., Zhang, K. W., Stankovich, J., Kilins, R. M., Dowling, J. P., Andermann, E., Andermann, F., Faldini, E., D’Hooge, R., Vadlamudi, L., Macdonell, R. A., Hodgson, B. L., Bayly, M. A., Savige, J., Mulley, J. C., Smyth, G. K., Power, D. A., Saftig, P., and Bahlo, M. (2008) Am. J. Hum. Genet. 82, 673–684
33. Reczek, D., Schwaek, M., Schröder, J., Hughes, H., Blanz, J., Jin, X., Brondyk, W., Van Patten, S., Edmunds, T., and Saftig, P. (2007) Cell 131, 770–783
34. Vesa, J., Chin, M. H., Oelgeschläger, K., Isosomppi, J., Dell’Angelica, E. C., Jalanko, A., and Peltonen, L. (2002) Mol. Biol. Cell 13, 2410–2420
35. Erickson, A. H. (1989) J. Cell. Biochem. 40, 31–41
36. Hibino, H., Imanabe, A., Furutani, K., Murakami, S., Findlay, I., and Kuraschi, Y. (2010) Physiol. Rev. 90, 291–366
37. Green, W. N., and Millar, N. S. (1995) Trends Neurosci. 18, 280–287
38. Jones, P. M., O’Mara, M. L., and George, A. M. (2009) Trends Biochem. Sci. 34, 520–531