

# Molecular basis of variant pseudo-Hurler polydystrophy (mucopolidosis IIIc)

Annick Raas-Rothschild,<sup>1,2</sup> Valerie Cormier-Daire,<sup>2</sup> Ming Bao,<sup>3</sup> Emmanuelle Genin,<sup>4</sup> Remi Salomon,<sup>2</sup> Kevin Brewer,<sup>3</sup> Marsha Zeigler,<sup>1</sup> Hanna Mandel,<sup>5</sup> Steve Toth,<sup>6</sup> Bruce Roe,<sup>6</sup> Arnold Munnich,<sup>2</sup> and William M. Canfield<sup>3</sup>

<sup>1</sup>Department of Human Genetics, Hadassah Hebrew University Hospital, Jerusalem 91120, Israel

<sup>2</sup>Unité de Recherches sur les Handicaps Génétiques de l'Enfant, Institut National de la Santé et de la Recherche Médicale (INSERM), Unite 393, Hôpital des Enfants Malades, 75015 Paris, France

<sup>3</sup>W.K. Warren Medical Research Institute and the Department of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104, USA

<sup>4</sup>INSERM U155, Château de Longchamp, Bois de Boulogne, 75016 Paris, France

<sup>5</sup>Department of Pediatrics, Rambam Hospital, Haifa 35254, Israel

<sup>6</sup>Department of Chemistry, University of Oklahoma, Norman, Oklahoma 73019, USA

Address correspondence to: William M. Canfield, University of Oklahoma Health Sciences Center, 975 NE 10th Street, Stanton L. Young Biomedical Research Center 411, Oklahoma City, Oklahoma 73104, USA.  
Phone: (405) 271-8147; Fax: (405) 271-3910; E-mail: william-canfield@ouhsc.edu.

Annick Raas-Rothschild, Valerie Cormier-Daire, and Ming Bao contributed equally to this work.

Received for publication November 17, 1998, and accepted in revised form January 24, 2000.

Mucopolidosis IIIc, or variant pseudo-Hurler polydystrophy, is an autosomal recessive disease of lysosomal hydrolase trafficking. Unlike the related diseases, mucopolidosis II and IIIa, the enzyme affected in mucopolidosis IIIc (*N*-Acetylglucosamine-1-phosphotransferase [GlcNAc-phosphotransferase]) retains full transferase activity on synthetic substrates but lacks activity on lysosomal hydrolases. Bovine GlcNAc-phosphotransferase has recently been isolated as a multisubunit enzyme with the subunit structure  $\alpha_2\beta_2\gamma_2$ . We cloned the cDNA for the human  $\gamma$ -subunit and localized its gene to chromosome 16p. We also showed, in a large multiplex Druze family that exhibits this disorder, that MLIIC also maps to this chromosomal region. Sequence analysis of the  $\gamma$ -subunit cDNA in patients from 3 families identified a frameshift mutation, in codon 167 of the  $\gamma$  subunit, that segregated with the disease, indicating MLIIC results from mutations in the phosphotransferase  $\gamma$ -subunit gene. This is to our knowledge the first description of the molecular basis for a human mucopolidosis and suggests that the  $\gamma$  subunit functions in lysosomal hydrolase recognition.

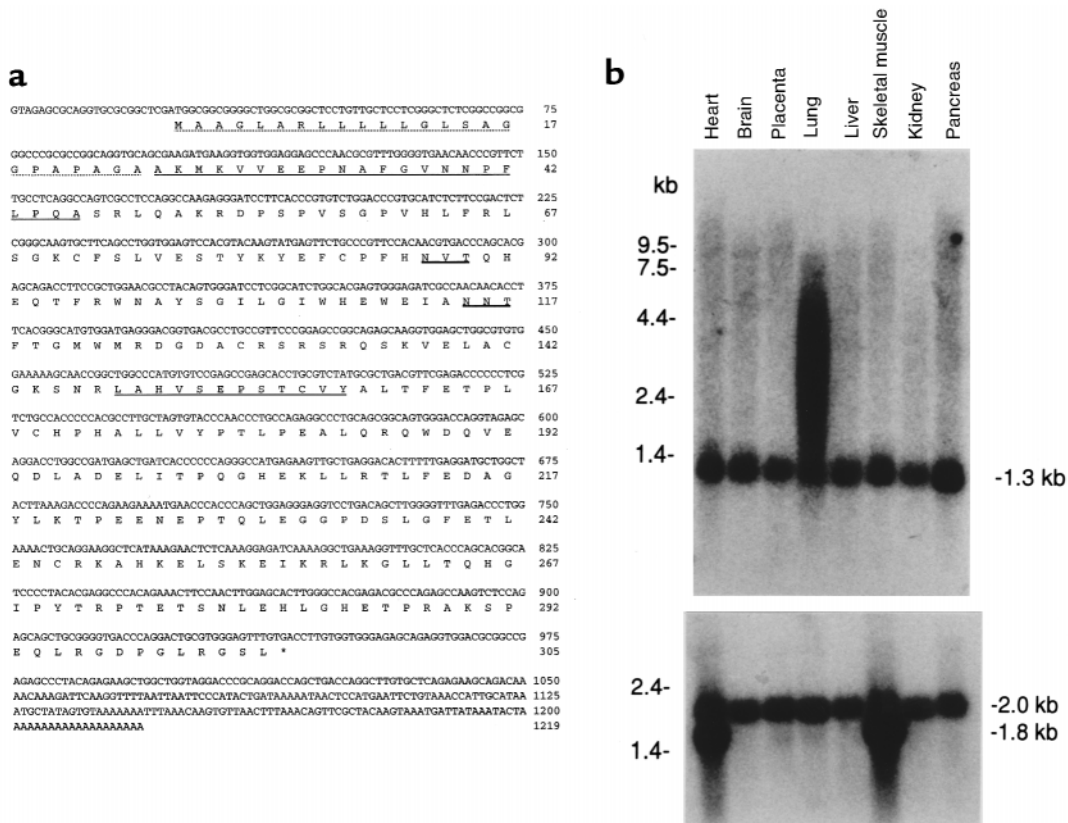
*J. Clin. Invest.* 105:673–681 (2000).

## Introduction

Mucopolidoses II, IIIa, and IIIc are 3 related diseases of lysosomal enzyme trafficking. Although rare diseases, they have been the subject of extensive study in the 20 years since their first description (1). The study of these diseases revealed the previously unsuspected process of lysosomal enzyme targeting (2, 3) and was crucial for the discovery of the mannose 6-phosphate-dependent lysosomal enzyme trafficking pathway (4, 5; for a recent review, see ref. 6). The trafficking of most lysosomal hydrolases in higher eucaryotes is mediated by an M6P-dependent pathway, in which asparagine-linked oligosaccharides on newly synthesized lysosomal hydrolases are specifically and uniquely modified to contain a terminal M6P. The initial and determining step in the biosynthesis of the M6P modification is catalyzed by the enzyme UDP-*N*-acetylglucosamine:lysosomal enzyme *N*-acetylglucosamine-1-phosphotransferase (GlcNAc-phosphotransferase). GlcNAc-phosphotransferase catalyzes the transfer of GlcNAc 1-phosphate from UDP-GlcNAc to specific  $\alpha$ 1,2-linked mannoses on lysosomal hydrolases

(7). The second enzyme in the pathway, *N*-Acetylglucosamine-1-phosphodiester  $\alpha$ -*N*-Acetylglucosaminidase, removes the covering GlcNAc, generating a terminal M6P. Lysosomal enzymes bearing the M6P modification then bind to 1 of 2 M6P-receptors in the *trans*-Golgi and are transferred to the lysosome (5, 8). In the absence of GlcNAc-phosphotransferase, lysosomal enzymes fail to acquire M6P, enter the secretory pathway, and are secreted from the cell.

GlcNAc-phosphotransferase activity is deficient in 2 distinct autosomal recessive human diseases. Mucopolidosis III (MLIII; pseudo-Hurler polydystrophy; MIM 252600) was first described as a lysosomal storage disease with morphological changes similar to Hurler syndrome (9). Pseudo-Hurler polydystrophy was distinguished from Hurler syndrome because of slower progression, absence of organomegaly, and absence of mucopolysacchariduria. Typical clinical findings include stiffness of the hands and shoulders, claw-hand deformity, scoliosis, short stature, coarse facies, and mild mental retardation. Radiographically, severe dysostosis



**Figure 1**

(a) Nucleotide and deduced amino acid sequence of the human GlcNAc-phosphotransferase  $\gamma$ -subunit cDNA. The predicted protein sequence is shown below the DNA sequence. The putative signal peptide is indicated by the dotted underline. Sequences homologous to the bovine  $\gamma$  subunit are indicated by single underline. Two potential sites for N-linked glycosylation are indicated by double underline. (b) Northern blot analysis of GlcNAc-phosphotransferase  $\gamma$  subunit. The  $\gamma$ -subunit cDNA was random labeled and hybridized to a human multiple tissue Northern blot (CLONTECH); each lane contained 2  $\mu\text{g}$  of poly(A)<sup>+</sup> RNA from the indicated tissue. The filter was washed at 65°C in 0.1  $\times$  SSC, 0.1% SDS, and exposed for 18 hours. In a, transcripts of 1.3 kb are identified in all tissues with additional larger transcripts identified in lung; in b, after hybridization, the blot was stripped and rehybridized with a probe for human  $\beta$ -actin mRNA as a control for loading.

multiplex of the hip is characteristic and frequently disabling (10). The clinical diagnosis can be confirmed by finding elevated serum lysosomal enzyme levels and/or decreased lysosomal enzyme levels in cultured fibroblasts. Alternatively, GlcNAc-phosphotransferase can be directly assayed in cultured fibroblasts (11, 12). GlcNAc-phosphotransferase activity is decreased in MLIII and absent in mucopolipidosis II (MLII; inclusion-cell [I-cell] disease; MIM 252500), a more severe disease that is generally fatal within the first decade. The deficiency of GlcNAc-phosphotransferase is believed to be the primary genetic deficit in both MLIII and MLII, although this has not yet been demonstrated at the molecular level (11, 13). Cultured MLIII fibroblasts, like MLII fibroblasts, demonstrate the I-cell phenotype with prominent cytoplasmic inclusions, which represent lysosomes containing ingested material that cannot be degraded because of the failure of lysosomal enzyme trafficking.

The relationship between the diseases MLII and MLIII has not been determined. Initially, it was proposed that the 2 disorders differed only in the amount of residual GlcNAc-phosphotransferase activity. Alternatively, the molecular basis for these 2 diseases may be distinct.

Support for this later proposal has been obtained by genetic complementation studies demonstrating complementation of MLII by some MLIII fibroblasts (14).

MLIII is itself genetically heterogeneous. Heterokaryon analysis has been used to divide MLIII into 3 distinct complementation groups (15, 16). Complementation group A (classic MLIII) is characterized by moderately decreased GlcNAc-phosphotransferase activity measured on all substrates. Complementation group B was represented by a single patient with atypical clinical features; the relationship of this patient to the other diseases is uncertain. Complementation group C (variant or Iranian variant MLIII) is characterized by deficient activity when assayed using lysosomal glycoproteins as acceptors, but normal activity when assayed with the simple acceptor,  $\alpha$ -methylmannoside (16, 17). The variant MLIII phenotype is confined to the MLIIIC complementation group, and it seems these 2 designations refer to the biochemical and genetic manifestations of the same primary defect (14).

Bovine GlcNAc-phosphotransferase has recently been purified as a 54-kDa complex composed of 3 different subunits ( $\alpha_2$ ,  $\beta_2$ , and  $\gamma_2$ ) (18). GlcNAc-phosphotransferase is the product of 2 genes, 1 encoding the  $\alpha$  and  $\beta$

subunits, the second encoding the  $\gamma$  subunit (W.M. Canfield, unpublished study). With the identification of GlcNAc-phosphotransferase as a multisubunit enzyme, we sought to define the relationship between the subunits identified biochemically, the complementation groups, and the mucopolidoses. We now report isolation of the human GlcNAc-phosphotransferase  $\gamma$ -subunit cDNA and localization of the gene to chromosome 16p. Homozygosity mapping with polymorphic microsatellite markers resulted in the mapping of variant MLIII to chromosome 16p. Direct sequence analysis of the  $\gamma$ -subunit cDNA in variant MLIII patients from 3 families demonstrate a single base insertion resulting in a frameshift and early termination. These results indicate that MLIIIC can result from mutations in the GlcNAc-phosphotransferase  $\gamma$ -subunit gene.

### Methods

**Amino acid sequencing.** Bovine GlcNAc-phosphotransferase (1.5 mg) was isolated by mAb immunoaffinity chromatography as described previously (18). Lubrol was removed by chromatography on a 0.2-mL column of POROS 50 HQ equilibrated with 0.025 M Tris HCl

(pH 7.4), followed by elution in buffer containing 0.5 M NaCl. The enzyme was denatured, reduced, and alkylated, and the subunits were resolved by chromatography on a 1.0  $\times$  30 cm column of Superose 6 equilibrated with 5% formic acid. An aliquot was subjected to automated Edman degradation in an Applied Biosystems Model 492 protein sequencer (Perkin-Elmer Applied Biosystems, Foster City, California, USA). A second aliquot was subjected to digestion with trypsin, and a tryptic peptide was isolated by reverse phase microbore HPLC and similarly sequenced.

**cDNA cloning.** A full-length cDNA for the  $\gamma$  subunit was assembled by ligating an 85-bp EcoRI-ApaI fragment amplified from a human brain cDNA library in pSPORT1 (Life Technologies, Grand Island, New York, USA) with an SP6 promoter primer and 5'-GTAGACGCAGGTGCTC-CGCTCG-3', with a 1,147-bp ApaI-Not I fragment from EST 48250 into EcoRI/Not I cut pcDNA3.

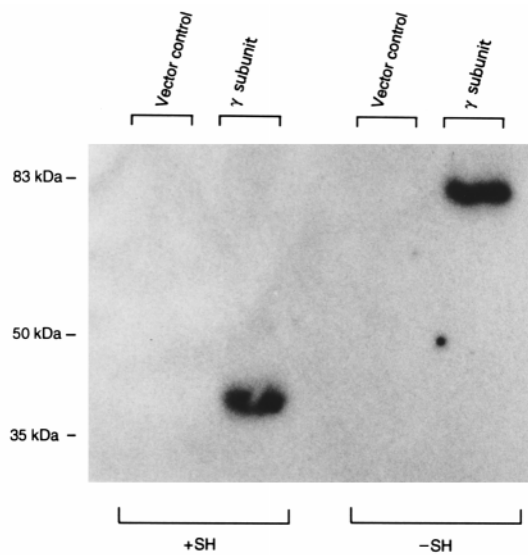
**DNA sequence analysis.** I.M.A.G.E. clone 48250 was sequenced as described previously using *Thermus aquaticus* (Taq) DNA polymerase, universal forward, reverse DNA, or custom sequencing primers and the Perkin-Elmer Cetus dye-labeled Taq terminators

**Table 1**

Pairwise lod-scores between MLIII (pedigree 1) and 5 polymorphic DNA markers on chromosome 16p

Locus name	Recombination fraction ( $\theta$ )									Zmax	$\theta$
		0	0.01	0.05	0.10	0.20	0.30	0.40			
<b>D16S521</b> (AFMa139wgl)	Sibs 1	-1.74	-1.03	-0.42	-0.17	0.01	0.04	0.03	0.04	0.04	0.3
	Sibs 2	$-\infty$	-3.23	-1.33	-0.63	-0.12	0.02	0.03	0.03	0.03	0.4
	Sibs 3	-0.58	-0.01	0.47	0.60	0.59	0.45	0.24	0.6	0.1	
	Sibs 4	1.02	0.99	0.88	0.74	0.47	0.25	0.10	1.02	0	
	Sibs 5	-0.75	-0.46	-0.10	0.01	0.02	0.00	-0.01	0.02	0.2	
	<b>Total</b>	$-\infty$	<b>-3.74</b>	<b>-0.51</b>	<b>0.55</b>	<b>0.96</b>	<b>0.76</b>	<b>0.40</b>	0.96	0.2	
<b>D16S3024</b> (AFMa134xcl)	Sibs 1	1.92	1.88	1.74	1.56	1.17	0.77	0.37	1.92	<b>0</b>	
	Sibs 2	2.65	2.59	2.36	2.08	1.51	0.95	0.44	2.65	<b>0</b>	
	Sibs 3	2.76	2.70	2.47	2.18	1.59	1.00	0.45	2.76	<b>0</b>	
	Sibs 4	0.90	0.87	0.77	0.65	0.43	0.24	0.10	0.90	<b>0</b>	
	Sibs 5	1.83	1.79	1.62	1.42	1.00	0.50	0.26	1.83	<b>0</b>	
	<b>Total</b>	<b>10.04</b>	<b>9.83</b>	<b>8.98</b>	<b>7.89</b>	<b>5.70</b>	<b>3.57</b>	<b>1.62</b>	<b>10.04</b>	<b>0</b>	
<b>D16S3070</b> (AFMa353yhl)	Sibs 1	0.99	0.96	0.86	0.74	0.50	0.29	0.11	0.99	<b>0</b>	
	Sibs 2	0.32	0.31	0.26	0.21	0.12	0.05	0.01	0.32	<b>0</b>	
	Sibs 3	0.38	0.36	0.31	0.24	0.14	0.07	0.02	0.38	<b>0</b>	
	Sibs 4	1.04	1.02	0.91	0.79	0.54	0.31	0.13	1.04	<b>0</b>	
	Sibs 5	1.09	1.06	0.93	0.77	0.49	0.25	0.08	1.09	<b>0</b>	
	<b>Total</b>	<b>3.81</b>	<b>3.70</b>	<b>3.27</b>	<b>2.75</b>	<b>1.78</b>	<b>0.97</b>	<b>0.36</b>	<b>3.81</b>	<b>0</b>	
<b>D16S3082</b> (AFMb070yg5)	Sibs 1	1.60	1.56	1.41	1.21	0.82	0.44	0.13	1.6	<b>0</b>	
	Sibs 2	2.51	2.46	2.24	1.96	1.40	0.86	0.37	2.51	<b>0</b>	
	Sibs 3	2.64	2.59	2.36	2.08	1.51	0.96	0.44	2.64	<b>0</b>	
	Sibs 4	0.95	0.92	0.81	0.67	0.41	0.21	0.08	0.95	<b>0</b>	
	Sibs 5	1.69	1.64	1.44	1.19	0.73	0.35	0.11	1.69	<b>0</b>	
	<b>Total</b>	<b>9.39</b>	<b>9.17</b>	<b>8.26</b>	<b>7.11</b>	<b>4.87</b>	<b>2.81</b>	<b>1.13</b>	<b>9.39</b>	<b>0</b>	
<b>D16S3072</b> (AFMb015wa9)	Sibs 1	-1.73	-1.03	-0.49	-0.29	-0.14	-0.07	-0.03	-0.03	0.4	
	Sibs 2	1.20	1.18	1.11	0.99	0.72	0.44	0.18	1.20	0	
	Sibs 3	0.66	0.63	0.55	0.44	0.26	0.13	0.05	0.66	0	
	Sibs 4	-1.65	-0.93	-0.37	-0.16	-0.04	-0.01	-0.00	-0.01	0.3	
	Sibs 5	-0.67	-0.38	-0.00	0.12	0.13	0.08	0.03	0.13	0.2	
	<b>Total</b>	<b>-2.20</b>	<b>-0.52</b>	<b>0.80</b>	<b>1.10</b>	<b>0.94</b>	<b>0.57</b>	<b>0.23</b>	1.10	0.1	

To reduce the computational complexity, the pedigree was divided into 5 sibships that were analyzed individually as described in Methods.



**Figure 2**

Expression of human GlcNAc-phosphotransferase  $\gamma$  subunit in COS cells. Membrane fractions of COS cells transfected with either the epitope-tagged  $\gamma$ -subunit cDNA or a vector control were analyzed by immunoblot. The samples were analyzed by 7.5% SDS-PAGE under reducing and nonreducing conditions followed by transfer to a nitrocellulose membrane, which was probed with mAb HPC4 and  $^{125}\text{I}$ -labeled goat anti-mouse IgG. Disulfide bonds were reduced with 10%  $\beta$ -mercaptoethanol in samples electrophoresed in lanes labeled +SH. The migration positions of molecular mass markers (in kDa) are indicated on the left.

(Perkin-Elmer Corp., Norwalk, Connecticut, USA) (19). The reactions were incubated for 30 cycles in a Perkin-Elmer Cetus DNA Thermocycler 9600, and, after removal of excess primers by filtration through Sephadex G-50, the fluorescent-labeled nested fragment sets were resolved by electrophoresis on an ABI 373A Automatic Sequencer modified with a Stretch Upgrade (Perkin-Elmer Applied Biosystems). After base calling with the ABI Analysis Software (version 2.1), the analyzed data were transferred to a Sun Sparcstation II (Sun Microsystems, Palo Alto, California, USA), and the electrophoretograms were edited using the TED trace editor program. Overlapping sequences and contigs were analyzed using the XGAP program (20, 21). Sufficient data were analyzed to obtain at least 3-fold coverage for each base.

**Multiple-tissue Northern blot analysis of GlcNAc-phosphotransferase  $\gamma$ -subunit expression.** A human multiple tissue Northern blot (CLONTECH Laboratories, Palo Alto, California, USA) was probed with a 1.15-kb Not I-Hind III fragment from the  $\gamma$ -subunit gene labeled with  $^{32}\text{P}$  by random hexamer priming with Klenow fragment. The blot was washed with  $0.1 \times \text{SSC}$ , 0.1% SDS at  $65^\circ\text{C}$ , and exposed to Biomax film (Eastman Kodak Co., Rochester, New York, USA) at  $-70^\circ\text{C}$  for 3 hours.

**Patients.** Three consanguineous families from Northern Israel were studied. Family 1 of Druze ancestry included 28 individuals (18 healthy and 10 affected

with MLIII). Family 2 (6 healthy and 2 affected) and family 3 (5 healthy and 2 affected) were of Arab origin.

The clinical diagnosis of MLIII was based on typical clinical, radiographic, and biochemical features. Clinical criteria included stiffness of the hands and shoulders, claw-hand deformity, scoliosis, short stature, coarse facies, and mild mental retardation. X-ray criteria included dysostosis multiplex with progressive destruction of the hip joint. Biochemical criteria included an at least 10-fold increase in serum lysosomal enzyme levels (Arylsulfatase A,  $\beta$ -hexosaminidase,  $\alpha$ -mannosidase) and/or low intracellular lysosomal enzyme levels in cultured fibroblasts.

**Genotyping.** DNA was isolated from peripheral blood lymphocytes by standard methods. Genotyping was performed with the PCR-based hypervariable microsatellite technique (22). PCR reactions contained genomic DNA (100 ng), Taq polymerase (1 unit), 50 pmol of each primer, 6.25 nmol of each deoxynucleotide triphosphate, 50 mM KCl, 10 mM Tris-HCl [pH 8], 1.5 mM  $\text{MgCl}_2$ , and 0.1% gelatin, in a final volume of 25  $\mu\text{L}$ .

PCR were performed in a thermocycler (Perkin-Elmer 480) with a 10-minute initial denaturation and 30 cycles of denaturation ( $94^\circ\text{C}$ , 40 seconds), annealing ( $55^\circ\text{C}$ , 40 seconds), and elongation ( $72^\circ\text{C}$ , 40 seconds) and a final elongation ( $72^\circ\text{C}$ , 10 minutes). An aliquot of the amplification reaction was denatured (10 minutes,  $94^\circ\text{C}$ ) and loaded on a 6% polyacrylamide denaturing gel. After transfer to Hybond N+ membranes (Amersham Pharmacia Biotech, Piscataway, New Jersey, USA) and crosslinking in 0.4 N NaOH, membranes were hybridized with a  $(\text{CA})_{12}$  peroxidase-labeled probe (Amersham ECL kit) for at least 8 hours at  $42^\circ\text{C}$ . Blots were washed once in  $0.3 \times \text{SSC}$ , 0.1% SDS,  $42^\circ\text{C}$  for 15 minutes, then twice in  $0.2 \times \text{SSC}$  solution for 20 minutes each time, and hybridizing bands were revealed using an Amersham ECL kit.

**Linkage analyses.** Two-point linkage analyses were performed with the MLINK linkage package (version 5.1) (23, 24). Autosomal recessive inheritance with complete penetrance, an assumed disease allele frequency of 0.0001, and CEPH allele frequencies were used. Because of the excessive computational requirements created by the large number of inbreeding loops, pedigree 1 was split into 5 sibships that were analyzed individually and lod scores were summed, thus reducing the overall statistical power of the pedigree. Sibships 1–5 consisted of VI:1–8; V:2–10; VI:9–14; VIII:1–6, and V:17–25. Inbreeding loops were broken at mothers as described in the Linkage software manual (23, 24). Homozygosity mapping was performed with MAP-MAKER/HOMOZ (25). Haplotypes were constructed by minimizing the number of recombination events. A multiple map was constructed with the pedigree divided into 5 sibships with the 5 relevant markers using the HOMOZ program.

**Sequence analysis of mutant  $\gamma$ -subunit cDNAs.** Total RNA was isolated from cultured skin fibroblasts as described using QIAGEN kit (QIAGEN Inc., Valencia,

California, USA) and reverse transcribed with oligo dT using the GeneAmpkit (Perkin-Elmer) to generate cDNA, which was amplified in 3 segments with the following primer pairs:

5'-GAGCGCAGGTGCGCGGCTCGA-3';  
 5'-GCGTTCCAGCGGAAGGTCTGC-3';  
 5'-GTTCCACAACGTGACCCAGCACGAGCA-3';  
 5'-GTGATCAGCTCATCGGCCAGGT-3';  
 5'-GCAGCGGCAGTGGGACCAGGT-3';  
 and 5'-TCCACCTCTGCTCTCCACCA-3'.

PCR was performed after a 10-minute 94°C initial denaturation and consisted of 30 cycles (denaturation 94°C, 40 seconds; annealing 65°C, 40 seconds; extension, 72°C, 40 seconds).

Amplification products were loaded onto a 2% Nusieve-agarose gel, purified by phenolchloroform extraction, and recovered by ethanol precipitation. Both strands were sequenced using the fluorescent dideoxy terminator method (26). Fluorometric sequences were analyzed using the Applied Biosystem 373A DNA Sequencer according to the manufacturer's protocol.

**Single-strand conformation polymorphism.** Single-strand conformation polymorphism (SSCP) analysis (27) of PCR-amplified genomic DNA was performed with the forward primer 5'-AGCACCTGCGTC-TACGCGCT-3' and backward primer 5'-GGGGTGGG-GAACGATCACAT-3', generating a 68-bp product. The product was labeled by adding 20 μCi α-[P<sup>33</sup>]-dCTP to the PCR reaction. PCR products were denatured at 95°C for 10 minutes and then loaded onto a

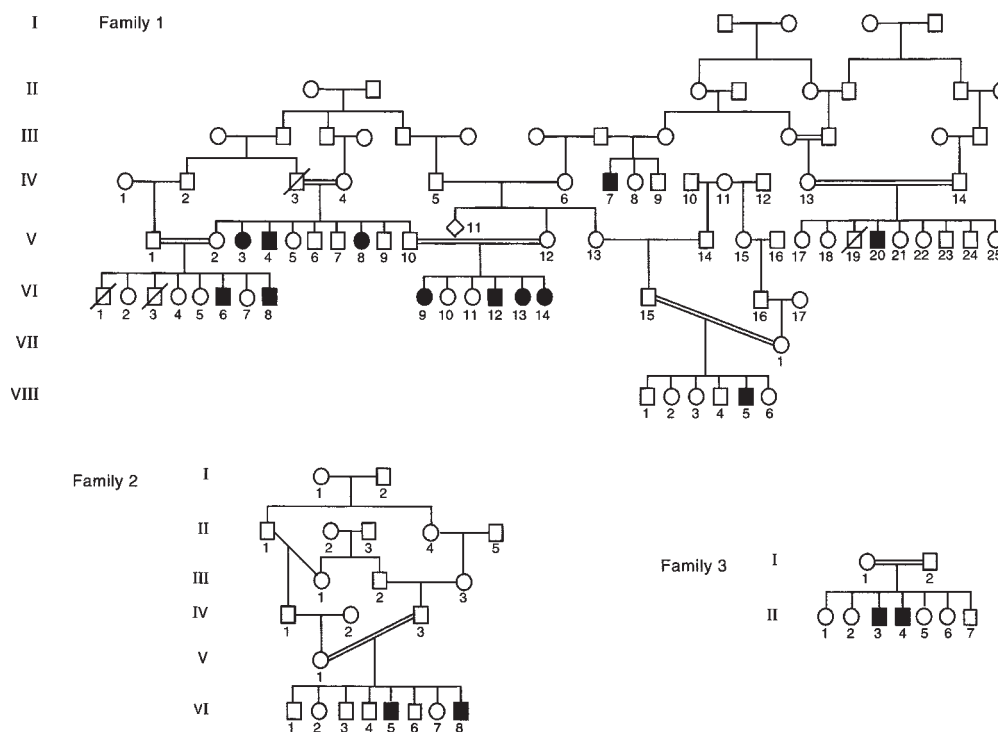
hydrolink MDE gel (Bioprobe Systems, Montreuil, France) and electrophoresed at 3 W for 10 hours in 0.6 TBE. Gels were dried and autoradiographed.

**GlcNAc-phosphotransferase assays.** GlcNAc-phosphotransferase transfer to α-methylmannoside was determined with β-[<sup>32</sup>P]-UDP-GlcNAc as described previously (28). Transfer to uteroferrin was performed by incubation of a solubilized membrane fraction at 37°C for 16 hours with 40 μM uteroferrin and 10 μCi [<sup>32</sup>P]-UDP-GlcNAc. The uteroferrin was then precipitated with 10% trichloroacetic acid, resolved by electrophoresis on a 12% SDS-PAGE gel, and incorporated radioactivity quantitated by Phosphorimager (Molecular Dynamics, Sunnyvale, California, USA). Protein was quantitated with the micro BCA assay (Pierce Chemical Co., Rockford, Illinois, USA), and results were expressed as pmol GlcNAc-1-phosphate incorporated per milligram of protein relative to a normal human fibroblast control.

**Accession numbers.** The GlcNAc-phosphotransferase γ-subunit cDNA has been deposited in GenBank with the accession number 179091.

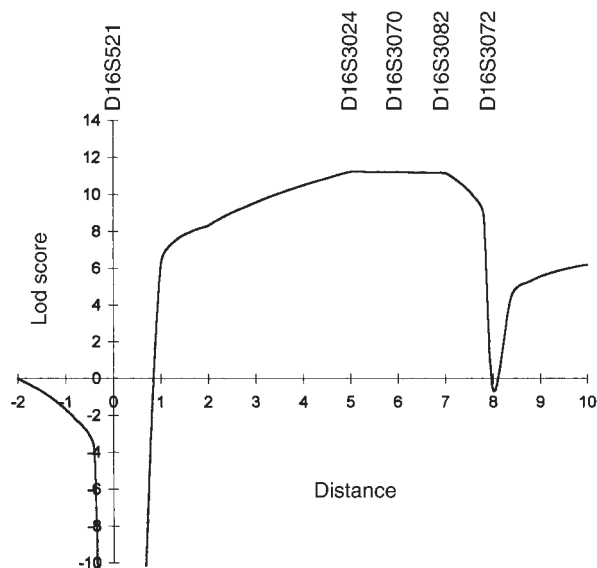
## Results

**Cloning and characterization of the human GlcNAc-phosphotransferase γ-subunit cDNA.** Purified bovine GlcNAc-phosphotransferase was reduced and alkylated, and the γ-subunit was isolated by gel filtration. The sequence of the NH<sub>2</sub>-terminus, AKMKVVEEPNTFGLNPNFLPQT-SRLQP, and a tryptic peptide, LAHVSEPSTCVY, were determined by automated Edman degradation.



**Figure 3**

Pedigrees of 3 Families with MLIII. Family 1 is of Druze origin; families 2 and 3 are of Arab origin. Circles denote female family members; squares, male family members; solid symbols, affected family members; and symbols with a slash, deceased family members.



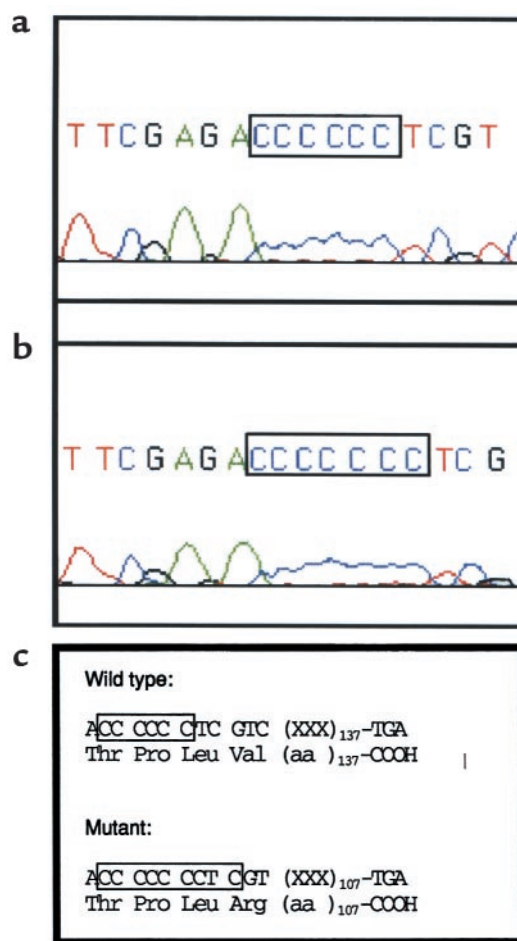
**Figure 4**  
Multipoint linkage analysis of the MLIII locus using DNA markers D16S521-D16S3024-D16S3070-D16S3082-D16S3072 with MAP-MAKER/HOMOZ. Genetic distances between the 5 marker loci on the abscissa are expressed in cM and based on the 1996 Génethon human genetic linkage map.

When the peptide sequences were used to search the databases using the program TBLASTN (29), 3 human expressed sequence tags (ESTs) were identified. ESTs 29315 and 179293 contain a 99-bp common sequence, a portion of which codes for a 20/22 amino acid sequence match with the bovine  $\gamma$ -subunit NH<sub>2</sub>-terminal sequence. EST 280314 contains a 12/12 amino acid match with the tryptic peptide sequence. A relationship between EST 29314 and EST280314 was demonstrated by PCR using a forward primer, 5'-GCGAAGATGAAGGTGGTGGAGGACC-3', corresponding to the NH<sub>2</sub>-terminal protein sequence and a reverse primer, 5'-GTAGACGCAGGTGCTCCGCTCG-3', corresponding to the tryptic peptide sequence using a human brain cDNA library as template. A 402-bp product was obtained that, when sequenced, demonstrated that EST 29315 and 280314 were derived from a common cDNA and that a 45-bp sequence bridges the gap.

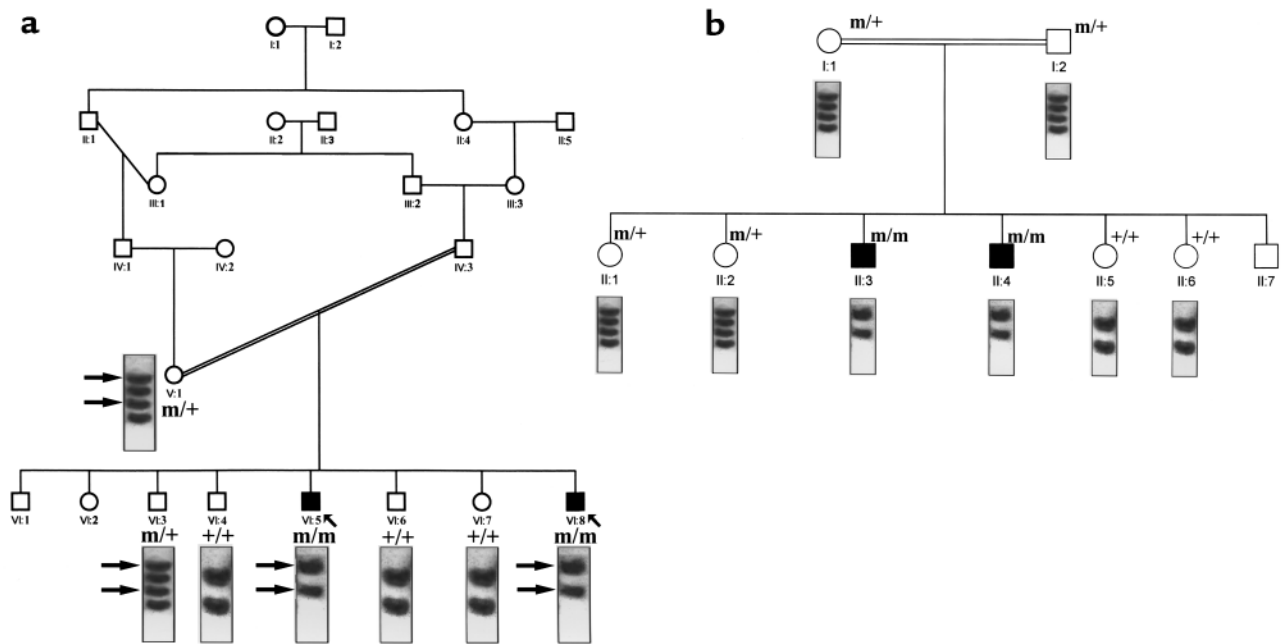
I.M.A.G.E. clone 48250 from which the 280314 EST sequence was determined was obtained and sequenced demonstrating an 1,151-bp insert containing an almost full-length cDNA for the  $\gamma$  subunit missing only the initiator methionine from the coding portion. The remaining 5' portion of the cDNA was obtained by PCR of a human brain cDNA library using an SP6 promoter primer as forward primer, and the previously described reverse primer yielded a 569-bp product that, upon cloning and sequencing, contained the 23 bp 5' nontranslated sequence and the codon for the initiator methionine.

*Sequence analysis of the GlcNAc-phosphotransferase  $\gamma$ -subunit cDNA.* The sequence of the 1,219-bp cDNA containing a 20-bp poly(A) tail is shown in Figure 1a. A

predicted open reading frame of 915 bp begins with a potential initiator ATG codon at base 24 and ends with a TGA termination codon at base 939. The cDNA lacks a consensus polyadenylation signal. The first ATG in the sequence fulfills the consensus sequence described for an initiator methionine (30). After the initiator methionine is a 24-amino acid signal peptide followed by a consensus signal peptidase cleavage site (31), just before the experimentally identified NH<sub>2</sub>-terminal alanine. The sequence predicts a mature protein of 281 amino acids with a molecular weight of 31,827. A search for protein motifs with Pustell Protein Matrix (MacVector, version 4.5.3; Genetics Computer Group, Madison, Wisconsin, USA) identified only 2 N-glycosylation sites. The sequence predicts that the mature protein contains 7 cysteine, the odd number being consistent with the disulfide-linked  $\gamma$ -subunit homodimers found in bovine GlcNAc-phosphotransferase.



**Figure 5**  
Sequence analysis of the mutant and wild-type  $\gamma$ -subunit cDNA. RNA was isolated from cultured fibroblasts, reverse transcribed, and a 300-bp portion of the  $\gamma$ -subunit cDNA amplified by PCR as described in Methods. The amplified fragment was resolved by agarose gel electrophoresis, excised, and directly sequenced using the PCR primers as primers and fluorescent terminators. (a) Wild-type sequence. (b) Mutant sequence. (c) Effect of the insertion on the  $\gamma$ -subunit protein.



**Figure 6**

Linkage of the  $\gamma$ -subunit mutation to MLIII in families 2 and 3 by SSCP analysis. A 680-bp region of the  $\gamma$ -subunit gene containing the  $\gamma$ -subunit mutation was amplified, denatured, electrophoresed, and detected by autoradiography as described in Methods. The SSCP pattern of each patient tested is given below the pedigree symbol. (a) Family 2; (b) family 3. m, mutant allele; +, wild-type allele.

A search of the BLOCKS database (32), and the Swiss Prot GenBank, and dbEST databases using the BLASTN or BLASTP algorithms (29) identified no significant similarities to genes of known function. However, a near identity (> 95% similarity) was found to a number of human ESTs isolated from 7 independent cDNA libraries.

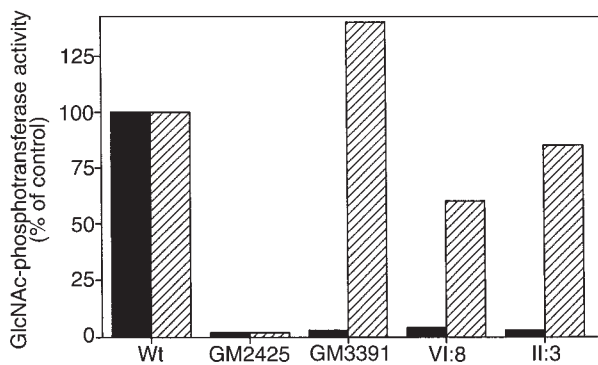
A multiple-tissue Northern blot probed with a full-length  $\gamma$ -subunit cDNA showed a single 1.3-kb transcript in all 8 human tissues examined (Figure 1b). The size of this transcript indicates that the 1,189-bp cDNA clone we isolated represents the full-length mRNA. Additional transcripts up to 7.5 kb were found in lung, consistent with alternate splicing, incomplete processing, or variation in polyadenylation.

**Biochemical characterization of the  $\gamma$ -subunit gene product in COS cells.** Transient expression of the GlcNAc-phosphotransferase  $\gamma$ -subunit was obtained by subcloning an epitope-tagged version of the cDNA in the vector pRcRSV (33). This construct replaces the native signal peptide with the human transferrin signal peptide (MRLAVGALLVCAVLGLCLA) and adds the HPC4 calcium-dependent epitope (EDQVDPRLIDGK) to the NH<sub>2</sub>-terminus of the mature  $\gamma$  subunit. As negative controls, COS cells were also transfected in parallel with the vector alone. Forty-eight hours after transfection, the cell membranes were isolated and the membrane proteins were solubilized, subjected to SDS-PAGE, and transferred to a nitrocellulose membrane. The epitope-tagged  $\gamma$  subunit was detected with mouse mAb HPC4 against the epitope, [<sup>125</sup>I]rabbit anti-mouse

IgG, and autoradiography (Figure 2). In COS cells transfected with the epitope-tagged  $\gamma$ -subunit cDNA but not the vector alone, a 35-kDa protein is expressed that forms disulfide-linked homodimers similar to the bovine GlcNAc-phosphotransferase  $\gamma$  subunit.

**Chromosomal localization of the human GlcNAc-phosphotransferase  $\gamma$ -subunit gene.** EST T30150 had previously been localized to chromosome 16p (D16S3024-D16S3070) by radiation hybrid mapping with the Genebridge 4 radiation hybrid panel (34). These results indicated the  $\gamma$ -subunit gene is also localized to chromosome 16p.

**Linkage of MLIII to chromosome 16p.** To test whether MLIII results from mutations in the  $\gamma$ -subunit gene, linkage analysis with 5 polymorphic microsatellite markers from chromosome 16p was performed in a large multiplex MLIII family from northern Israel (Figure 3, family 1). Because of the computational complexity of the multiple inbreeding loops, family 1 was analyzed as 5 separate sibships. Table 1 shows pairwise lod scores between microsatellite markers of chromosome 16 and the disease locus. A maximum pairwise lod score was obtained with probe AFMa134xcl at locus D16S3024 ( $Z_{\max} = 10.04$  at  $\theta = 0$ ). Positive lod scores were also obtained for polymorphic markers at loci D16S3070 and D16S3082 ( $Z_{\max} = 3.81$  at  $\theta = 0$ , and  $Z_{\max} = 9.39$  at  $\theta = 0$ , respectively). Multipoint linkage analysis performed to estimate the position of the MLIII gene gave a peak lod score of 11.59 at locus D16S3024. The maximum likelihood estimate for location of the MLIII locus was defined by loci D16S521 and D16S3072 in a genetic interval of 7 cM (Figure 4).



**Figure 7**

Enzymatic characterization of the mutant GlcNAc-phosphotransferase. Normal human skin fibroblasts (Wt); and skin fibroblasts from patients with classical MLIII (GM2425) and variant MLIII (GM3391); and patient VI:8 (pedigree 1), patient II:3 (pedigree 3), were cultured and a solubilized membrane fraction prepared as described in Methods. GlcNAc-phosphotransferase was assayed using both the lysosomal enzyme uteroferin and  $\alpha$ -methylmannoside as acceptors and calculated as pMol GlcNAc-1-phosphate transferred per milligram of cell protein per hour. The data are expressed as a percentage of the activity observed in the wild-type cell line.

*Identification of a GlcNAc-phosphotransferase  $\gamma$ -subunit mutation in MLIII.* The  $\gamma$ -subunit cDNA was amplified by RT-PCR using 3 sets of primers in the probands of family 1 (VI:8) and family 3 (II:3, II:4). The products were identified by electrophoresis on agarose gels and excised, and both strands were directly sequenced using the PCR primers and fluorescent dideoxy terminators. Insertion of a single cytosine at codon 167 (ACC CCC CTC  $\rightarrow$  ACC CCC CCT; Figure 5, a and b) resulted in a frame shift and predicts a premature translation termination, 107 bp downstream from the insertion (Figure 5c).

*Linkage of MLIII to the  $\gamma$ -subunit insertion by SSCP analysis.* A 68-bp segment, encompassing the insertion already identified here, was amplified from genomic DNA for SSCP analysis. The mutation segregated with the disease in the 3 families except for individual VIII:5 (family 1), whose SSCP profile was normal (Figure 6). SSCP study of 75 unrelated healthy individuals did not show the abnormal profile. These results indicate the  $\gamma$ -subunit mutation is the molecular basis for MLIII in these 3 families.

The SSCP profile was normal in VIII:5 and his parents. Although this patient was homozygous at the loci studied, his haplotype was different than the one segregating in the rest of family 1 and is probably derived from individual IV:11. GlcNAc-phosphotransferase assays showed that this patient belonged to complementation group A (data not shown). This indicates both MLIII complementation groups (A and C) are segregating in family 1 (see later here). Sequence analysis of the  $\gamma$ -subunit cDNA from patient VIII:5 did not show the cytosine insertion. These data suggest that this patient carries a different mutation, probably in the  $\alpha$  or  $\beta$  subunit.

*Enzymatic characterization of the mutant GlcNAc-phosphotransferase.* To classify the disease resulting from the  $\gamma$ -subunit mutation as either classical or variant MLIII, GlcNAc-phosphotransferase assays were performed. A solubilized membrane fraction was prepared from cultured skin fibroblasts from patients VI:8 (pedigree 1), II:3 (pedigree 3), 2 reference MLIII lines (GM2425; MLI-IIA and GM3391; MLIIC), and controls. In MLIII(A), GlcNAc-phosphotransferase activity is equally reduced with either  $\alpha$ -methylmannoside or lysosomal enzymes as acceptors. In MLIII(C), the GlcNAc-phosphotransferase activity with  $\alpha$ -methylmannoside as acceptor is retained, but it is defective with lysosomal enzymes as acceptors. As shown in Figure 7, extracts from MLIII(A) and MLIII(C) cells are similarly deficient in GlcNAc-phosphate transfer to the lysosomal enzyme, uteroferin. The extracts, however, differ in the ability to transfer GlcNAc-phosphate to the simple acceptor,  $\alpha$ -methylmannoside. GM2425 is deficient (~2% activity), whereas GM3391 demonstrates approximately 140% of normal activity with the simple acceptor. When membrane fractions from patients VI:8 (pedigree 1) and II:3 (pedigree 3), were examined, little transfer to uteroferin was observed, but 60–85% of normal activity in the  $\alpha$ -methylmannoside assay was found. These results indicate the  $\gamma$ -subunit mutation results in a variant or MLIII(C) type defect.

## Discussion

The molecular basis for the distinctive properties of the mutant GlcNAc-phosphotransferase in MLIII(C) has been enigmatic since the original description (17). The demonstration that GlcNAc-phosphotransferase is a multiple subunit enzyme encoded by 2 distinct genes suggested the MLIII(C) defect might be due to deficiency of 1 of the gene products. This study demonstrates that MLIII(C) results from a mutation in the  $\gamma$ -subunit gene and provides the first description to our knowledge of the molecular basis of a human mucopolidosis.

The mutation identified in this study results in a GlcNAc-phosphotransferase that is catalytically active but lacks the ability to recognize lysosomal enzymes as acceptors, strongly suggesting the  $\gamma$ -subunit functions in substrate recognition. The mechanism by which the  $\gamma$ -subunit functions is unclear, and the sequence demonstrates no homology with previously described proteins, providing no clues. How this mutation alters the structure of the GlcNAc-phosphotransferase complex is also unknown, but given that the mutant enzyme retains near normal activity with some substrates, it is unlikely the absence of the intact  $\gamma$  subunit interferes with assembly or stability of the complex.

The genetic relationships between the multiple variants of MLII and MLIII are complex and have been difficult to understand in the absence of information about the gene(s) involved (14–16, 35). This study demonstrates the MLIII(C) phenotype can result from a mutation in the GlcNAc-phosphotransferase  $\gamma$  subunit. Further study will be required to demonstrate



that the MLIII(C) phenotype is exclusively caused by mutations in the  $\gamma$ -subunit gene, although we believe this is likely. The mutant GlcNAc-phosphotransferase lacking an intact  $\gamma$  subunit still expresses near normal catalytic activity, strongly suggesting that the  $\gamma$  subunit does not contribute to the catalytic function of the enzyme. Because the catalytic activity is decreased or absent in MLII and MLIII(A) with all substrates, it is tempting to speculate that these diseases result from alterations in the  $\alpha$  and/or  $\beta$  subunits of the enzyme and that the  $\alpha$  and/or  $\beta$  subunits contain the catalytic portion of GlcNAc-phosphotransferase. This model also predicts that MLII and MLIII(A) would be genetically complemented by MLIII(C), a result that has been demonstrated previously (14).

### Acknowledgments

A. Raas-Rothschild was the recipient of a Chateaubriand fellowship of Ministère des affaires étrangères (France) and an R.M. Goodman fellowship (United States). These studies were supported in part by grants from the National Institutes of Health (NIH; HD31920 to W.M. Canfield and HG00313 and NSF/EPSCoR to B. Roe) and from the W.K. Warren Medical Research Institute and the Presbyterian Health Foundation (both to W.M. Canfield); by The Vaincre Les Maladies Lyssosomales Association (to A. Raas-Rothschild). We thank all of the Israeli physicians involved in the diagnosis of the patients, and Y. Zlotogara, J. Kaplan, and S. Lyonnet for their help. We thank R.M. Roberts, University of Missouri, Columbia, for providing purified uteroferrin. We thank K. Jackson and the Molecular Biology Resource Facility of the W.K. Warren Medical Research Institute for assistance with protein sequencing, and M. Arcade for secretarial expertise.

- Leroy, J.G., and DeMars, R.I. 1967. Mutant enzymatic and cytological phenotypes in cultured human fibroblasts. *Science*. **157**:804-806.
- Hickman, S., and Neufeld, E.F. 1972. A hypothesis for I-cell disease: defective hydrolases that do not enter lysosomes. *Biochem. Biophys. Res. Commun.* **4**:992-999.
- Hickman, S., Shapiro, L.J., and Neufeld, E.F. 1974. A recognition marker required for uptake of a lysosomal enzyme by cultured fibroblasts. *Biochem. Biophys. Res. Commun.* **1**:55-61.
- Kaplan, A., Achord, D.T., and Sly, W.S. 1977. Phosphohexosyl components of a lysosomal enzyme are recognized by pinocytosis receptors on human fibroblasts. *Proc. Natl. Acad. Sci. USA*. **74**:2026-2030.
- Kornfeld, S. 1987. Trafficking of lysosomal enzymes. *FASEB J.* **1**:462-468.
- Kornfeld, S., and Sly, W.S. 1999. I-cell disease and pseudo-hurler polydystrophy: Disorders of lysosomal enzyme phosphorylation and localization. In *Metabolic basis of inherited diseases*. McGraw-Hill Publishing Co. New York, NY.
- Couso, R., Lang, L., Roberts, R.M., and Kornfeld, S. 1986. Phosphorylation of the oligosaccharide of uteroferrin by UDP-GlcNAc: glycoprotein N-acetylglucosamine-1-phosphotransferase from rat liver, *Acanthamoeba castellanii*, and *Dictyostelium discoideum* requires alpha 1,2-linked mannose residues. *J. Biol. Chem.* **261**:6326-6331.
- Dahms, N.M., Lobel, P., and Kornfeld, S. 1989. Mannose 6-phosphate receptors and lysosomal enzyme targeting. *J. Biol. Chem.* **264**:12115-12118.
- Maroteaux, P., and Lamy, M. 1966. La pseudopolydystrophie de Hurler. *Presse Med.* **55**:2889-2892.
- Kelly, T.E., et al. 1975. Mucopolipidosis III (pseudo-Hurler polydystrophy): clinical and laboratory studies in a series of 12 patients. *Johns Hopkins Med. J.* **137**:156-175.
- Reitman, M.L., Varki, A., and Kornfeld, S. 1981. Fibroblasts from patients with I-cell disease and pseudo-Hurler polydystrophy are deficient in uridine 5'-diphosphate-N-acetylglucosamine: glycoprotein N-acetylglucosaminylphosphotransferase activity. *J. Clin. Invest.* **67**:1574-1579.
- Hasilik, A., Waheed, A., and von Figura, K. 1981. Enzymatic phosphorylation of lysosomal enzymes in the presence of UDP-N-acetylglucosamine. Absence of the activity in I-cell fibroblasts. *Biochem. Biophys. Res. Commun.* **98**:761-767.
- Varki, A., et al. 1982. Demonstration of the heterozygous state for I-cell disease and pseudo-Hurler polydystrophy by assay of N-acetylglucosaminylphosphotransferase in white blood cells and fibroblasts. *Am. J. Hum. Genet.* **34**:717-729.
- Mueller, O.T., Honey, N.K., and Little, L.E. 1983. Mucopolipidosis II and III. The genetic relationships between two disorders of lysosomal enzyme biosynthesis. *J. Clin. Invest.* **72**:1016-1023.
- Honey, N.K., Mueller, O.T., Little, L.E., Miller, A.L., and Shows, T.B. 1982. Mucopolipidosis III is genetically heterogeneous. *Proc. Natl. Acad. Sci. USA*. **79**:7420-7424.
- Little, L.E., Mueller, O.T., Honey, N.K., Shows, T.B., and Miller, A.L. 1986. Heterogeneity of N-acetylglucosamine 1-phosphotransferase within mucopolipidosis III. *J. Biol. Chem.* **261**:733-738.
- Varki, A.P., Reitman, M.L., and Kornfeld, S. 1981. Identification of a variant of mucopolipidosis III (pseudo-Hurler polydystrophy): a catalytically active N-acetylglucosaminylphosphotransferase that fails to phosphorylate lysosomal enzymes. *Proc. Natl. Acad. Sci. USA*. **78**:7773-7777.
- Bao, M., Booth, J.L., Elmendorf, B.J., and Canfield, W.M. 1996. Bovine UDP-N-acetylglucosamine: lysosomal enzyme N-acetylglucosamine-1-phosphotransferase. I. Purification and subunit structure. *J. Biol. Chem.* **271**:31437-31445.
- Chissoe, S.L., et al. 1995. Sequence and analysis of the human ABL gene, the BCR gene, and regions involved in the Philadelphia chromosomal translocation. *Genomics*. **27**:67-82.
- Gleeson, T.J., and Staden, R. 1991. An X windows and UNIX implementation of our sequence analysis package. *Comput. Appl. Biosci.* **7**:398.
- Dear, S., and Staden, R. 1991. A sequence assemble and editing program for efficient management of large projects. *Nucleic Acids Res.* **19**:3907-3911.
- Weissenbach, J., et al. 1992. A second-generation linkage map of the human genome. *Nature*. **359**:794-801.
- Lathrop, G.M., and Lalouel, J.M. 1984. Easy calculations of lod scores and genetic risks on small computers. *Am. J. Hum. Genet.* **36**:460-465.
- Lathrop, G.M., Lalouel, J.M., Julier, C., and Ott, J. 1985. Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. *Am. J. Hum. Genet.* **37**:482-498.
- Kruglyak, L., Daly, M., and Lander, E. 1995. Rapid multipoint linkage analysis of recessive traits in nuclear families, including homozygosity mapping. *Am. J. Hum. Genet.* **56**:519-527.
- Prober, J.M., et al. 1987. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science*. **238**:336-341.
- Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., and Sekiya, T. 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc. Natl. Acad. Sci. USA*. **86**:2766-2770.
- Reitman, M.L., Lang, L., and Kornfeld, S. 1984. UDP-N-acetylglucosamine: lysosomal enzyme N-acetylglucosamine-1-phosphotransferase. *Methods Enzymol.* **107**:163-172.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.
- Kozak, M. 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* **115**:887-903.
- Perlman, D., and Halvorson, H.O. 1983. A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. *J. Mol. Biol.* **167**:391-409.
- Henikoff, S., and Henikoff, J.G. 1994. Protein family classification based on searching a database of blocks. *Genomics*. **19**:97-107.
- Rezaie, A.R., and Esmon, C.T. 1992. The function of calcium in protein C activation by Thrombin and the Thrombin-Thrombomodulin complex can be distinguished by mutational analysis of protein C derivatives. *J. Biol. Chem.* **267**:26104-26109.
- Walter, M.A., Spillet, D.J., Thomas, P., Weissenbach, J., and Goodfellow, P.N. 1994. A method for constructing radiation hybrid maps of whole genomes. *Nat. Genet.* **7**:22-28.
- Shows, T.B., Mueller, O.T., Honey, N.K., Wright, C.E., and Miller, A.L. 1982. Genetic heterogeneity of I-cell disease is demonstrated by complementation of lysosomal enzyme processing mutants. *Am. J. Med. Genet.* **12**:343-353.