

Aspartylglycosaminuria in the Finnish population: Identification of two point mutations in the heavy chain of glycoasparaginase

[lysosomal storage disease/glycoproteinosis/polymerase chain reaction/RNA/ N^4 -(β -*N*-acetylglucosaminy)-*L*-asparaginase]

ILKKA MONONEN*^{†‡}, NORA HEISTERKAMP[§], VESA KAARTINEN^{†¶}, JULIAN C. WILLIAMS*, JOHN R. YATES III^{||}, PATRICK R. GRIFFIN^{||}, LEROY E. HOOD^{||}, AND JOHN GROFFEN[§]

*Division of Medical Genetics and [§]Section of Molecular Diagnosis, Department of Pathology, Childrens Hospital of Los Angeles, Los Angeles, CA 90027; [†]Department of Clinical Chemistry, Kuopio University Central Hospital, SF-70210 Kuopio, Finland; [¶]Department of Chemistry, University of Kuopio, SF-70210 Kuopio, Finland; and ^{||}Division of Biology, California Institute of Technology, Pasadena, CA 91125

Contributed by Leroy E. Hood, December 26, 1990

ABSTRACT Aspartylglycosaminuria is an inherited lysosomal storage disease caused by deficiency of glycoasparaginase (EC 3.5.1.26) and occurs with higher frequency among Finns than other populations. We have purified human glycoasparaginase and determined about 90% of the amino acid sequence of its light subunit and >70% of that of its heavy subunit by Edman degradation and mass spectrometry. Additional sequence data were obtained from the cloning and subsequent nucleotide analysis of a cDNA corresponding to the normal human glycoasparaginase gene. The enzyme is encoded by a single mRNA as a single polypeptide that is posttranslationally processed to generate the subunits and is glycosylated. After preparing first-strand cDNA from leukocyte and fibroblast total RNA, we used the polymerase chain reaction to amplify the glycoasparaginase cDNA of eight Finnish aspartylglycosaminuria patients. We demonstrate that the Finnish patients' mRNA sequence differed from the normal sequence by two single-base changes six nucleotides apart from one another in the heavy chain of glycoasparaginase. The first change resulted in the replacement of arginine by glutamine (R161Q), whereas the second change resulted in a cysteine to serine substitution (C163S). Both mutations resulted in novel restriction endonuclease sites and were present in all eight Finnish aspartylglycosaminuria patients originating from different pedigrees, but they were absent from Finnish and non-Finnish controls and a non-Finnish case of aspartylglycosaminuria. These results indicate molecular homogeneity in aspartylglycosaminuria alleles in the Finnish population.

Glycoasparaginase [N^4 -(β -*N*-acetylglucosaminy)-*L*-asparaginase, EC 3.5.1.26] is a lysosomal enzyme that catalyzes the hydrolysis of the linkage between *N*-acetylglucosamine and asparagine, the attachment point of *N*-glycosidic carbohydrate chains to the protein chain. The deficient activity of glycoasparaginase in humans causes an autosomal recessive, lysosomal storage disease, aspartylglycosaminuria (McKusick 20840), characterized by serious psychomotor retardation and accumulation of large amounts of aspartylglucosamine in body fluids and tissues (1). To study the biochemistry and cell biology of glycoasparaginase, we have isolated the enzyme from human leukocytes (2). We have shown that the native enzyme is an 88-kDa glycoprotein that has a heterotetrameric structure composed of two 25-kDa heavy chains and two 19-kDa light chains that by themselves are enzymatically inactive. It has been suggested that human glycoasparaginase is encoded as a single polypeptide that is posttranslationally processed to generate the subunits (3). At least two different cDNA clones for human glycoasparaginase have been reported (3, 4), but as of yet, the expression

of an active glycoasparaginase enzyme from these has not been described.

The genetic structure of the Finnish population is unique in that 20 or more monogenic disorders are more common in that population than in other Caucasian populations (5). Inborn errors commonly occurring in other populations, such as phenylketonuria, cystic fibrosis, and galactosemia, are almost nonexistent in Finland (6). Aspartylglycosaminuria is one of those "Finnish diseases," and its frequency in eastern and northern Finland has been estimated as 1:4000–1:7000 (7, 8). This is the highest frequency described for any glycoproteinosis in any population and is comparable to the incidence figures of the most common lysosomal storage diseases, Gaucher disease type I (9) and Tay–Sachs disease (10) among Ashkenazi Jews. Although isolated cases of this disorder presenting with mental retardation, dysplasia, cardiac valvular involvement, and angiokeratoma have been reported elsewhere in the world (1), a phenotype with mental retardation, speech disorder, motor clumsiness, coarse face, and osteochondrotic changes as the most prominent clinical findings is consistently found among the Finnish aspartylglycosaminuria patients (8). The molecular defect(s) underlying the disease is unknown. In this study, we present evidence for two single-base changes in the heavy chain of glycoasparaginase found in aspartylglycosaminuria patients of Finnish descent.** Both mutations were absent from controls of Finnish and non-Finnish origin as well as one aspartylglycosaminuria case of non-Finnish origin.

MATERIALS AND METHODS

Human Leukocyte Glycoasparaginase. Glycoasparaginase was purified 4600-fold from human leukocytes (2). The heavy (25-kDa) and light (19-kDa) subunits of the native 88-kDa enzyme were isolated by reverse-phase HPLC. Peptides produced from these subunits by cyanogen bromide cleavage and digestion with chymotrypsin, trypsin, or endoproteases Glu-C, Lys-C, and Asp-N, respectively, were isolated by reverse-phase HPLC and sequenced by automated Edman degradation (Applied Biosystems 477A protein sequencer and 120A phenylthiohydantoin analyzer) or mass spectrometry (Finnigan MAT TSQ70 triple quadrupole mass spectrometer) (11, 12).

Designation of Aspartylglycosaminuria Patients and Normal Controls. Total RNA was isolated from leukocytes of five and from fibroblast cell lines of another three Finnish aspartylglycosaminuria patients from different pedigrees. The pa-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

[‡]To whom reprint requests should be addressed at: Division of Medical Genetics, Childrens Hospital of Los Angeles, 4650 Sunset Boulevard, Los Angeles, CA 90027.

**The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M60808 and M60809).

tients had typical clinical manifestations of aspartylglycosaminuria. They excreted large amounts of aspartylglycosamine in urine (13), and glycoasparaginase activity was completely absent in their leukocytes or fibroblasts. The five Finnish normal controls were selected so that their family history was negative with respect to aspartylglycosaminuria and the glycoasparaginase activity in their leukocytes was close to the upper level among controls in order to avoid heterozygosity (14).

Cell Lines. Three fibroblast cell lines of Finnish and one non-Finnish aspartylglycosaminuria patients (numbers GM00568, GM02056, GM02057, and GM03560, respectively) as well as the K562 human chronic myelogenous leukemia cell line (GM5372) were from the National Institute of General Medical Sciences Human Mutant Cell Repository, Coriell Institute for Medical Research (Camden, NJ).

Enzyme Assays. Glycoasparaginase activity was measured by HPLC (15).

PCR Amplification of Glycoasparaginase cDNA. An oligonucleotide corresponding to glycoasparaginase cDNA residues 1074–1095 (2) (Fig. 1, amplicer B) was annealed to total RNA isolated from peripheral blood leukocytes or cultured fibroblasts. First-strand cDNA was then prepared by using Moloney murine leukemia virus reverse transcriptase (Bethesda Research Laboratories). The second amplicer encompassed cDNA residues 107–128 (Fig. 1, amplicer A). Both amplicers contained a nucleotide substitution resulting in a *Sal* I site, which facilitated subsequent cloning. The nucleotide sequences of the amplicers were A, 5'-TGCCCTGGTCGTCGACACTTG-3' and B, 5'-AATACAGATGTCGACAGTAAAG-3'. PCR was performed as described by Kawasaki *et al.* (16) with minor modifications. The reaction mixture contained 2 μ g of total RNA, 50 pmol of each oligonucleotide amplicer, dNTPs (each at 0.2 mM), 20 mM Tris-HCl (pH 8.3), 2.5 mM MgCl₂, 50 mM KCl, nuclease-free bovine serum albumin at 100 μ g/ml, and 20 units of *Taq* polymerase, in a total volume of 100 μ l. The mixture was incubated 2 min at 94°C to denature the cDNA, annealed 90 sec at 55°C, and extended 2 min at 72°C. A new cycle of amplification was started by denaturing the DNA at 94°C. Thirty-five to 40 cycles of amplification were performed by using a programmable heat block (Perkin-Elmer/Cetus).

Sequencing of DNA. PCR-amplified DNAs were phenol extracted, ethanol precipitated, and digested with *Sal* I. The DNA was purified by electrophoresis on a low-melting-point agarose gel and phenol/sodium acetate extraction (17). DNA fragments were ligated into *Sal* I-digested M13mp18 or M13mp19 vector DNA (18). The nucleotide sequence of the single-stranded DNA was determined by the dideoxynucle-

otide chain-termination method according to Sanger *et al.* (19) by using oligonucleotide primers based on the cDNA sequence (see Fig. 1). The nucleotide sequences of the primers used were S1, 5'-TACTGGAACATACAACACA-3'; S2, 5'-GTTGTAATCCATAAGAC-3'; S3, 5'-ATGGTTGATCTTCTCC-3'; S4, 5'-CATTTGTAGATGTACCAGCAGCA-3'; and S5, 5'-CTACTAAAAGTGTGTGTG-3'. The complete DNA sequence of a normal gene was determined once, as well as the complete sequence of one patient, and the mutated area was resequenced once; all sequencing was performed in both directions, using Sequenase (United States Biochemical).

Mutant Screening by Allele-Specific Oligonucleotide Hybridization. The DNAs were subjected to electrophoresis on a 1.0% gel and blotted onto a nylon membrane (Nytran, Schleicher & Schuell). After prehybridization at 65°C, oligonucleotide hybridization with ³²P-end-labeled oligonucleotides was performed at 48°C overnight, followed by washing at gradually increased temperatures to distinguish between perfect matches and mismatches (20). The filters were exposed to x-ray film for 0.5–3 hr.

RESULTS AND DISCUSSION

Peptide Sequencing of Human Glycoasparaginase and Nucleotide Sequence Analysis of Its cDNA. Since at least two different cDNA clones for human glycoasparaginase have been reported (3, 4), but the expression of an active glycoasparaginase enzyme from these has not been described, we have studied the protein structure of glycoasparaginase isolated from human leukocytes (2). The purified enzyme had a specific activity of 2.2 units/mg of protein with *N*⁴-(β -*N*-acetylglucosaminyl)-L-asparagine as substrate. The 88-kDa enzyme migrated as a single band on native PAGE. On SDS/PAGE (under denaturing conditions) the tetrameric enzyme was dissociated to its heavy (25-kDa) and light (19-kDa) subunits. Our protein sequence data on the subunits of leukocyte glycoasparaginase indicated high similarity to that deduced from the nucleotide sequence of a cDNA clone isolated from a human placental library based on rat liver asparaginase amino acid sequence data (3). So oligonucleotide amplicers were synthesized according to that sequence for PCR analysis on human leukocyte mRNA. A pair of amplicers (Fig. 1) consistently allowed amplification of an \approx 1-kb cDNA product (see Fig. 5A, lane 2). Since both amplicers contained a nucleotide difference from the published sequence, resulting in generation of *Sal* I restriction sites, the PCR product of RNA of human cell line K562 could be cloned in the *Sal* I site of M13, and several clones were sequenced. The sequence from these K562 clones was similar to the nucleotide sequence reported by Fisher *et al.* (3) (data not shown). The amino acid sequence deduced from the K562 clones and the peptide sequences obtained from native glycoasparaginase are shown in Fig. 2. The peptide sequences that covered about 90% of the light chain and >70% of the heavy chain were similar to the amino acid sequence deduced from the nucleotide sequence. All of the sequences of the major peptides obtained from the subunits were found to be present within the cloned sequence. In the case of the heavy chain, the proteolytic and cyanogen bromide cleavages resulted in a number of peptides that either were present in small amounts, in mixtures of several peptides, or did not give a signal in peptide sequencing; less sequence data were obtained from the heavy chain than from the light chain. The protein sequence data unambiguously demonstrated that both our cDNA clone and that recently reported by Fisher *et al.* (3) correspond to glycoasparaginase. The evidence shows that the subunits of native, tetrameric human glycoasparaginase are encoded as a single polypeptide that is posttranslationally processed.

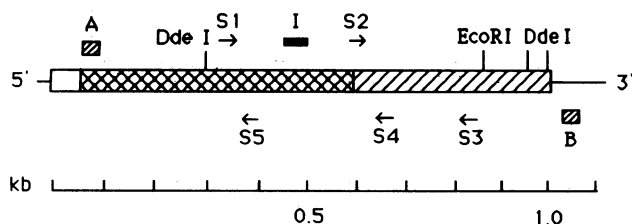


FIG. 1. Structure of glycoasparaginase cDNA and PCR amplicers. The boxed area represents the glycoasparaginase coding sequences; the lines represent 5' and 3' untranslated regions (from ref. 3). The open box indicates the signal peptide, the crosshatched box represents the heavy chain coding sequence, and the hatched box represents the light chain coding sequence of glycoasparaginase. The amplicers labeled A and B are shown above and below the cDNA. The black bar labeled I indicates the region shown in Fig. 3. The location of the *Eco*RI cleavage site and the three *Dde* I sites in the normal glycoasparaginase cDNA are shown. The direction and approximate location of the sequencing primers (S1–S5) are indicated with horizontal arrows. kb, Kilobases.

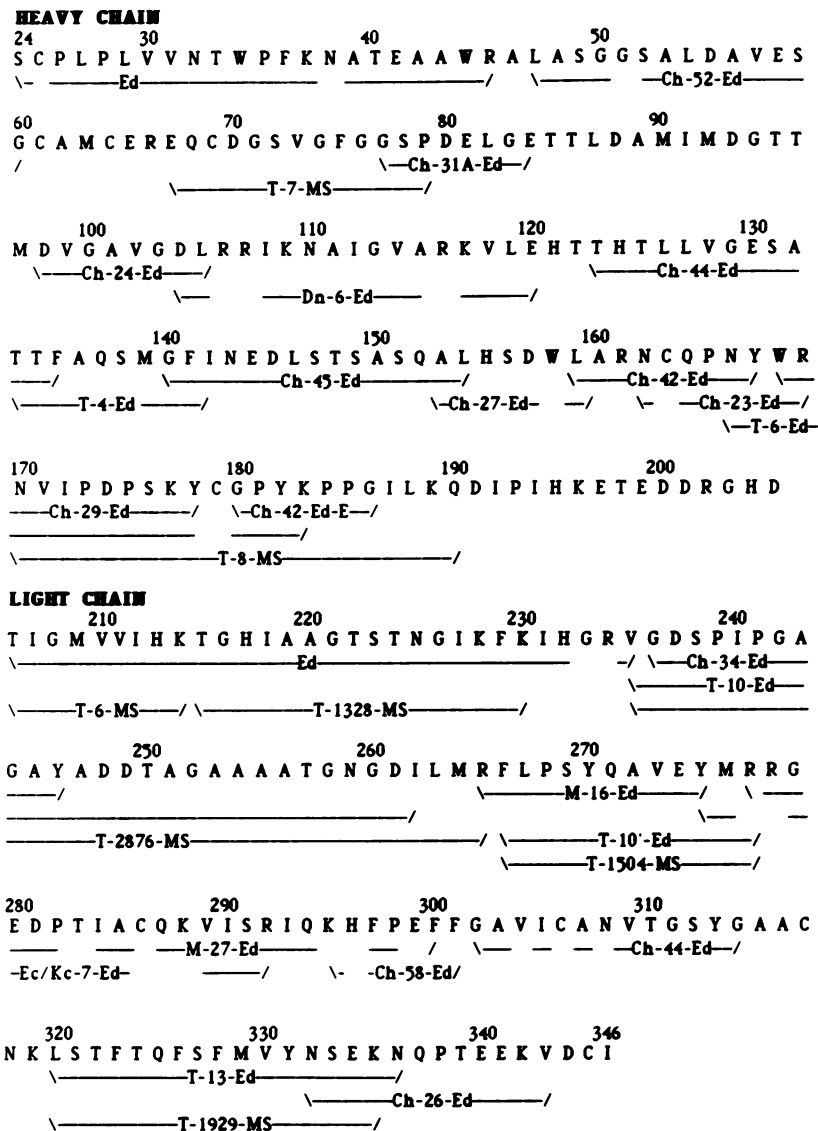


FIG. 2. Amino acid sequence data on human leukocyte glycoasparaginase obtained by Edman degradation and mass spectrometry compared to the amino acid sequence deduced from leukocyte and placental (3) cDNA. The amino acid sequence based on cDNA is shown on the top line using one-letter symbols, and the residues are numbered as in ref. 3. The peptides under the sequence are marked according to the method used in their generation and the technique used to sequence them. Proteolytic and chemical cleavage methods are designated as follows: Ch, chymotrypsin; Dn, endoproteinase Asp-N; Ec, endoproteinase Glu-C; Kc, endoproteinase Lys-C; T, trypsin; M, cyanogen bromide. The amino-terminal sequences of the subunits were obtained by sequencing native proteins. Peptides sequenced by Edman degradation are labeled Ed and those sequenced by mass spectrometry are labeled MS. Unidentified residues are designated by blank spaces. The slashes demarking peptide lengths represent peptide-terminal amino acids.

Nucleotide and Amino Acid Sequence Analysis of cDNAs of Finnish Aspartylglycosaminuria Patients. As an initial approach to identify mutations in aspartylglycosaminuria, the same pair of amplimers that enabled the amplification of virtually the complete coding sequence of the enzyme was used. PCR amplification was performed on cDNA synthesized from patient and control mRNA. The size of the amplification product of one Finnish aspartylglycosaminuria patient was the same as that of the normal cDNA product (see Fig. 5A, lane 1), excluding any major deletion in the amplified region. The product was subcloned and sequenced, and two single-base changes, G → A (mutation I) and G → C (mutation II), were found six bases apart from one another (Fig. 3), close to the carboxyl-terminal end of the heavy chain (Fig. 1). Mutation I created a new restriction enzyme site for *Dde* I, and mutation II created a new restriction enzyme site for *Eco*RI. Mutation I results in the replacement of an arginine residue by a glutamine residue (R161Q), and mutation II results in the replacement of a cysteine residue by a serine residue (C163S) (Fig. 4). The cDNA sequence of the light chain was found to be identical to that of the normal glycoasparaginase gene.

Hybridization of cDNAs with Allele-Specific Oligonucleotides. To provide additional evidence that these base substitutions are related to aspartylglycosaminuria and to obtain data regarding the frequency of these mutations, amplified

cDNAs were analyzed with allele-specific oligonucleotides. We amplified glycoasparaginase cDNA for Southern analysis from another six Finnish aspartylglycosaminuria patients from different pedigrees, one non-Finnish aspartylglycosaminuria patient, and three non-Finnish controls. Two 21-mer allele-specific oligonucleotides, one for each mutation, were synthesized as shown in Fig. 4 and used for hybridization (Fig. 5). The oligonucleotide for mutation I hybridized with all control and aspartylglycosaminuria DNAs after nonstringent posthybridization washing, but after more stringent washing at 62°C [3× standard saline citrate (SSC)], only signals from the samples from seven Finnish aspartylglycosaminuria patients remained (Fig. 5 B1 and B2). Similarly, the oligonucleotide specific for mutation II initially hybridized with all samples, but at 56°C (3× SSC), signals from normal controls and the non-Finnish aspartylglycosaminuria sample specifically disappeared (Fig. 5 C1 and C2). RNA isolated from a blood sample of an eighth Finnish aspartylglycosaminuria patient was similarly analyzed, and the patient was demonstrated to have the same two mutations as the seven patients described above (data not shown).

Digestion of cDNAs with *Eco*RI and *Dde* I Restriction Endonucleases. The presence of the *Eco*RI site in the Finnish aspartylglycosaminuria patients was further demonstrated by digestion of the PCR products with that particular enzyme. In addition to an *Eco*RI site that cleaves a 171-base-pair (bp)

than normal, indicating a deletion in some area of the gene. It was similar to controls rather than to the Finnish aspartylglycosaminuria patients at the *EcoRI* and *Dde I* sites (Fig. 5).

Significance of the Findings. Our extensive protein data on human leukocyte glycoasparaginase has enabled us to confirm the nucleotide sequence for the cDNA of the enzyme. We have also developed an approach to amplify almost the entire coding region of the enzyme for cloning, sequencing, and detection of mutations in aspartylglycosaminuria in that region in general.

We propose that the two single-base changes described above for the Finnish aspartylglycosaminuria patients are point mutations typical of aspartylglycosaminuria found in the Finnish population. Although it might have been possible that a mutation was present in noncoding or intron sequences in mRNA splicing, the current evidence demonstrates that the coding region of glycoasparaginase is affected in Finnish aspartylglycosaminuria patients. The replacement of an arginine in normal glycoasparaginase by a glutamine in the mutant enzyme results in the substitution of a basic amino acid for one containing an uncharged polar group. The replacement of cysteine by serine due to mutation II may abolish a disulfide bridge. Both amino acid changes will obviously modify the structure of the protein profoundly, but whether one or both are required to cause the deficiency in activity of glycoasparaginase remains to be established. For example, it is possible that one of the mutations corresponds to a neutral polymorphism while the other is causative for the gene defect. However, such polymorphism has not been found in the alleles of the five Finnish normal individuals studied to date. Characterization of the obvious deletion in the non-Finnish patient (Fig. 5, lane 7) along with other potential mutations causing the deficient activity of glycoasparaginase will provide important data for understanding of the structure and function of the enzyme.

The unique genetic structure of the Finnish population due to isolation by geographical, linguistic, and cultural barriers is well known, but very little is known about the molecular basis of the Finnish diseases. Gyrate atrophy of the choroid and retina caused by deficiency of ornithine δ -aminotransferase (EC 2.6.13) is one such Finnish genetic disease. Surprisingly, at least two different mutant alleles of ornithine δ -aminotransferase have been found in the Finnish population (21). In this study, we have shown that in eight Finnish aspartylglycosaminuria patients from different pedigrees two separate point mutations are found close to the carboxyl-terminal end of the heavy chain of glycoasparaginase. The presence of two mutations in affected patients, but not in normal population controls, is unusual, raising the interesting question of the mechanism of these events. The clinical finding that the phenotype of Finnish aspartylglycosaminuria patients is quite uniform (8), in addition to the data presented in this study, may indicate the presence of a single predominant genotype of aspartylglycosaminuria in the Finnish population. The findings that each of the mutations results in the generation of a new restriction site and that these mutations seem highly diagnostic for Finnish patients with aspar-

tylglycosaminuria should facilitate analysis of additional cases and may be useful in prenatal diagnosis in families with a history of this disease.

We thank Drs. M. Kaski, K. Raty, and E. Airaksinen for providing several of the Finnish aspartylglycosaminuria samples. We thank Drs. Joanne ten Hoeve, Q.-L. Hao, and Q.-S. Zhu for assistance; Dr. John Tomich and Bing Cai for the synthesis of oligonucleotides used in this study; and Ke-Cheng Chen for artwork.

1. Beaudet, A. L. & Thomas, G. H. (1989) in *The Metabolic Basis of Inherited Disease*, eds. Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D. (McGraw-Hill, New York), pp. 1603–1621.
2. Kaartinen, V., Williams, J. C., Tomich, J., Yates, J. R., III, Hood, L. E. & Mononen, I. (1991) *J. Biol. Chem.*, in press.
3. Fisher, K., Tollersrud, O. K. & Aronson, N. N. (1990) *FEBS Lett.* **269**, 440–444.
4. Peltonen, L. & Aula, P. (1987) *Am. J. Hum. Genet.* **41**, A233 (abstr.).
5. Norio, R. (1981) in *Biocultural Aspects of Disease*, eds. Rothchild, H. & Chapman, C. F. (Academic, London), pp. 359–415.
6. Nevanlinna, H. R. (1980) in *Population Structure and Genetic Disorders*, eds. Eriksson, A. W., Forsius, H. R., Nevanlinna, H. R., Workman, P. L. & Norio, R. K. (Academic, London), pp. 569–576.
7. Mononen, T., Mononen, I., Matilainen, R. & Airaksinen, E. (1991) *Hum. Genet.*, in press.
8. Aula, P., Autio, S., Raivio, K. O. & Rapola, J. (1982) in *Genetic Errors of Glycoprotein Metabolism*, eds. Durand, P. & O'Brien, J. (Springer, Berlin), pp. 123–152.
9. Barranger, J. & Ginns, I. (1989) in *The Metabolic Basis of Inherited Disease*, eds. Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D. (McGraw-Hill, New York), pp. 1677–1698.
10. Sandhoff, K., Conzelmann, E., Neufeld, E. F., Kaback, M. M. & Suzuki, K. (1989) in *The Metabolic Basis of Inherited Disease*, eds. Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D. (McGraw-Hill, New York), pp. 1807–1842.
11. Hunt, D. F., Yates, J. R., III, Shabanowitz, J., Winston, S. & Hauer, C. R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6233–6237.
12. Erickson, A. K., Payne, D. M., Martino, P. A., Rossomando, A. J., Shabanowitz, J., Weber, M. J., Hunt, D. F. & Sturgill, T. W. (1990) *J. Biol. Chem.* **265**, 19728–19735.
13. Mononen, T., Parviainen, M., Penttila, I. & Mononen, I. (1986) *Clin. Chem.* **32**, 501–502.
14. Mononen, I., Kaartinen, V. & Mononen, T. (1988) *Scand. J. Clin. Lab. Invest.* **48**, Suppl. 191, 7–11.
15. Kaartinen, V. & Mononen, I. (1990) *Anal. Biochem.* **190**, 98–101.
16. Kawasaki, E. S., Clark, S. S., Coyne, M. Y., Smith, S. D., Champlin, R., Witte, O. N. & McCormick, F. P. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5698–5702.
17. Heisterkamp, N., Groffen, J. & Stephenson, J. R. (1983) *J. Appl. Genet.* **2**, 57–68.
18. Norrander, J., Kempe, T. & Messing, J. (1983) *Gene* **26**, 101–106.
19. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
20. Valenzuela, D. M. & Groffen, J. (1986) *Nucleic Acids Res.* **14**, 843–852.
21. Mitchell, G. A., Brody, L. C., Sipila, I., Looney, J. E., Wong, C., Engelhardt, J. F., Patel, A. S., Steel, G., Obie, C., Kaiser-Kupfer, M. & Valle, D. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 197–201.