Characterization of a mutation in a family with saposin B deficiency: A glycosylation site defect

(sphingolipid activator protein/SAP-1/metachromatic leukodystrophy/arylsulfatase A)

KEITH A. KRETZ*, GEOFFREY S. CARSON*, SATOSHI MORIMOTO*†, YASUO KISHIMOTO*,
ARVAN L. FLUHARTY‡, AND JOHN S. O'BRIEN*§

*Department of Neurosciences and Center for Molecular Genetics, University of California, San Diego, School of Medicine, M-034J, La Jolla, CA 92093; and

‡University of California, Los Angeles, Mental Retardation Research Center Group at Lanterman Developmental Center, Pomona, CA 91766

Communicated by Dan L. Lindsley, January 19, 1990

ABSTRACT Saposins are small, heat-stable glycoproteins required for the hydrolysis of sphingolipids by specific lysosomal hydrolases. Saposins A, B, C, and D are derived by proteolytic processing from a single precursor protein named prosaposin. Saposin B, previously known as SAP-1 and sulfatide activator, stimulates the hydrolysis of a wide variety of substrates including cerebroside sulfate, GM1 ganglioside, and globotriaosylceramide by arylsulfatase A, acid β -galactosidase, and α -galactosidase, respectively. Human saposin B deficiency, transmitted as an autosomal recessive trait, results in tissue accumulation of cerebroside sulfate and a clinical picture resembling metachromatic leukodystrophy (activator-deficient metachromatic leukodystrophy). We have examined transformed lymphoblasts from the initially reported saposin B-deficient patient and found normal amounts of saposins A, C, and D. After preparing first-strand cDNA from lymphoblast total RNA, we used the polymerase chain reaction to amplify the prosaposin cDNA. The patient's mRNA differed from the normal sequence by only one C $\rightarrow$ T transition in the 23rd codon of saposin B, resulting in a threonine to isoleucine amino acid substitution. An affected male sibling has the same mutation as the proband and their heterozygous mother carries both the normal and mutant sequences, providing additional evidence that this base change is the disease-causing mutation. This base change results in the replacement of a polar amino acid (threonine) with a nonpolar amino acid (isoleucine) and, more importantly, eliminates the glycosylation signal in this activator protein. One explanation for the deficiency of saposin B in this disease is that the mutation may increase the degradation of saposin B by exposing a potential proteolytic cleavage site (arginine) two amino acids to the amino-terminal side of the glycosylation site when the carbohydrate side chain is absent.

The lysosomal hydrolysis of sphingolipids is catalyzed by the sequential action of acid hydrolases. Several small heat-stable glycoproteins called sphingolipid activator proteins have been discovered that act as natural nonspecific detergents, or stimulate a specific hydrolase, or both. The complete nucleotide sequence of a cDNA encoding prosaposin, the precursor of saposins A, B, C, and D, has been elucidated (1, 2). Prosaposin is a 524-amino acid glycoprotein and examination of its amino acid sequence reveals four saposin domains. Each domain is $\approx$ 80 amino acid residues long; has nearly identical placement of cysteine residues, glycosylation sites, and helical regions; and is flanked by potential proteolytic cleavage sites (lysine or arginine). Proteolytic cleavage of prosaposin at or near these dibasic amino acids was predicted to give rise to four saposin proteins (1). Each of

these four saposin proteins has now been isolated and their activating properties have been determined (3–14).

Saposins A and C specifically activate hydrolysis of glucocerebroside by β -glucosylceramidase (D-glucosyl-N-acyl-sphingosine glucosylhydrolase; EC 3.2.1.45) and of galactocerebroside by galactosylceramidase (D-galactosyl-N-acyl-sphingosine galactohydrolase; EC 3.2.1.46) (3, 4). Saposin D specifically activates the hydrolysis of sphingomyelin by sphingomyelin phosphodiesterase (sphingomyelin cholinephosphohydrolase; EC 3.1.4.12) (5). Saposins A, C, and D appear to exert their activities by binding to the respective enzymes, raising the maximal velocity of hydrolysis and lowering the Michaelis constant (5, 6) (S.M. and Y.K., unpublished data).

Saposin B, previously designated by several different terms (7, 10, 14, 15), stimulates the hydrolysis of galactocerebroside sulfate by arylsulfatase A (aryl-sulfate sulfohydrolase; EC 3.1.6.1) (7–9), GM1 ganglioside by acid β -galactosidase (β -D-galactoside galactohydrolase; EC 3.2.1.23) (10, 11), and globotriaosylceramide by α -galactosidase A (α -D-galactoside galactohydrolase; EC 3.2.1.22) (12, 13). This activator protein may have even broader substrate specificity since it also is an activator of glycerolipid hydrolysis (14). Saposin B activates by a mechanism different from saposins A, C, and D; it interacts with lipid substrates solubilizing them for enzymatic hydrolysis. The physiological significance of saposin B is underscored by the discovery of its absence in a variant form of metachromatic leukodystrophy (activator-deficient metachromatic leukodystrophy) (16–18). In this report, we present evidence for a single base change as the molecular defect in activator-deficient metachromatic leukodystrophy found in two siblings of consanguineous parents and propose that this mutation gives rise to a glycosylation site defect. These results were previously presented independently in preliminary form by our group and by Wenger *et al.* (19, 20).

MATERIALS AND METHODS

Quantitation of Saposins. A HPLC method was developed to quantitate the levels of saposins. Transformed lymphoblasts from proband YF with saposin B deficiency and a normal control were grown in suspension culture and collected by precipitation. After washing in phosphate-buffered saline the cell pellets were lyophilized, resuspended, homogenized, boiled, and centrifuged. Supernatant proteins were fractionated by HPLC sequentially on two columns, a hydrophobic Vydac C4 column (The Separations Group, Hesperia, CA) using an acetonitrile gradient followed by an

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.

Abbreviation: PCR, polymerase chain reaction.

†Present address: Faculty of Pharmaceutical Sciences, Kyushu University, Fukuoka 812, Japan.

§To whom reprint requests should be addressed.

Fig. 1. Structure of prosaposin cDNA and location of PCR and sequencing primers. Open box represents the prosaposin open reading frame and lines represent untranslated sequence (from refs. 1 and 2). Hatched areas represent the four saposin regions, as indicated (1, 2). The PCR primers are labeled P5 and P3 and their sequences are as follows: P5, ACGTACTCTAGACGCGCTATGTACGCCCTCTT; P3, ATCGATGAGCTCCACTGATGTCCCAAGCCACCA. The underlined portions of the PCR primers are restriction sites engineered in the primers (*Xba* I for P5 and *Sac* I for P3). The positions of the sequencing primers (S1–S5) are also shown.

anion-exchange Aquapore AX-300 column (Western Analytical Products, Temecula, CA) using a salt gradient. On the first column, saposins A, C, and D were collected as clustered peaks and on the second column, they were separated as individual peaks, which were quantified. Details of this method will be given in a separate report (S.M. and Y.K., unpublished data). Saposin B had previously been shown to be nearly absent in cultured cells from patient YF by a quantitative immunologic method (17). We also could not detect saposin B in lymphoblasts from patient YF after SDS/PAGE and immunoblotting with monospecific anti-saposin B antibodies.

Polymerase Chain Reaction (PCR) Amplification of Prosaposin cDNA. Total RNA was isolated from transformed lymphoblasts of the index patient by using the RNA isolation kit (Invitrogen, San Diego, CA) according to the manufacturer's instructions. First-strand cDNA was then prepared by using the Red Module (Invitrogen) and oligo(dT) as primer according to the manufacturer's instructions. PCR was performed as described by Saiki *et al.* (21). Frozen cell culture stocks of skin fibroblasts from EF, the affected brother of YF, and her mother were used as the source of the RNA for analysis of these samples.

Sequencing. Initially, prosaposin cDNA was cloned into pBS-II (Stratagene) and multiple clones were sequenced using the Sequenase kit (United States Biochemical) according to the manufacturer's instructions. Direct sequencing of PCR products was performed according to Kretz *et al.* (22). The PCR products were separated from the primers on a 1% NuSieve (FMC) low-melt agarose gel. The bands were excised and melted at 68°C. The sequencing primer was annealed to the PCR product by heating at 95°C for 10 min and quickly cooling to 37°C. Sequencing was then carried out with the Sequenase kit according to the manufacturer's instructions except that the labeling reaction was carried out at 37°C.¶

RESULTS AND DISCUSSION

It was reported earlier that the proband YF had a severe deficiency of saposin B in fibroblasts (<5% of normal) (16, 17). Since saposin B is generated by proteolytic processing of a precursor that also gives rise to three other activator proteins (1), it was important to determine the levels of the other three activator proteins in this patient. HPLC analysis of normal and saposin B-deficient lymphoblasts showed no deficiency of saposin A, C, or D in the proband YF (Table 1), suggesting that the defect was localized to the region encoding saposin B.

PCR amplification was performed to generate enough copies of the prosaposin cDNA for nucleotide sequencing. Primers were constructed so that the sequence of the entire

prosaposin open reading frame and some 3' flanking sequence totaling 2170 base pairs could be analyzed (Fig. 1). Initially, the product from patient YF was cloned into the phagemid pBS II and several clones were sequenced. A single C → T transition was found in the 23rd codon of saposin B (Fig. 2), a single base change resulting in the replacement of a threonine residue by an isoleucine residue (Fig. 3). No other base changes were found after sequencing the entire prosaposin open reading frame. These results are in accord with those recently reported by Wenger *et al.* (20).

To provide additional evidence that this base substitution is the disease-causing mutation, we amplified prosaposin cDNA from fibroblast RNA from the affected brother of the proband (EF), her mother, and several controls, sequencing the PCR products directly. As expected, EF has the same base change as YF, and her mother has two bands of approximately equal intensity at this position, representing the normal and mutant alleles (Fig. 2). Samples from three normal subjects each gave the normal sequence at this position.

We propose that the single base change is a point mutation that generates proteolytically sensitive saposin B. After removal of its signal peptide, prosaposin is glycosylated at five glycosylation sites [two in saposin A, one each in saposins B, C, and D (1, 3, 5, 23)]. Prosaposin is proteolyzed to generate saposins A, B, C, and D, with cleavage occurring at or near basic dipeptides at their polypeptide boundaries (1, 2, 24). We propose that the point mutation in the F family abolishes the glycosylation site in saposin B (Asn-Xaa-Thr; Fig. 3) and exposes an arginine residue two residues amino terminal from this site due to the absence of the carbohydrate side chain that normally protects this site from cleavage. Previously, we analyzed the structure of saposin B by helical wheel depictions of its helical structure and analysis of hydrophathy plots. We concluded that the region in question is hydrophilic and exposed at a turn of the polypeptide chain. Since normal processing of prosaposin occurs by proteolysis at dibasic amino acid residues, proteolytic enzymes are present that could accomplish cleavage of the mutant protein at the exposed arginine. This proposal is consonant with pulse-chase experiments in fibroblasts from YF in which prosaposin is normally synthesized and excreted in the presence of ammonium chloride but mature saposin B is not detected (25). The availability of a rapid PCR-based screening

Table 1. Quantitation of saposins A, C, and D

	Control	Patient YF
Saposin A	2.2	2.8
Saposin C	1.0	2.8
Saposin D	21.1	29.1

Results are expressed as $\mu\text{g/g}$ dry weight. Transformed lymphoblasts from a control and the saposin B-deficient patient YF were analyzed for saposin A, C, and D content by HPLC.

¶The sequence reported in this paper has been deposited in the GenBank data base (accession no. M32221).

FIG. 2. Sequence of a portion of the saposin B coding region. Direct sequence data of PCR products from the index case (YF), affected brother (EF), their mother (M), and a control (C) using the S2 sequencing primer are shown. Note the C $\rightarrow$ T transition in the two patient samples and the double band in the mother.

method with direct sequencing to detect the defect will be useful in diagnosing additional patients with this disease.

Examination of the primary structures of saposins A, C, and D revealed that each of the glycosylation sites (two sites in saposin A, and one site each in saposins C and D) has either arginine or lysine, or both, residues within three amino acids of the glycosylation signal. Similarly, inspection of the amino acid sequence of the rat sulfated glycoprotein (SGP-1) (26), known to be 76% homologous to the protein sequence of the human SAP precursor, showed that three of its four glycosylation sites have closely associated arginine or lysine residues. These observations are consistent with the highly conserved structures of these proteins and, moreover, suggest that mutational events destroying other glycosylation sites might also result in the absence of gene product due to the unmasking of proteolytic sites.

Glycosylation of proteins has been proposed as a means to retard proteolysis by altering conformation and by protecting sensitive sites (27, 28). Several proteins (including fibroblast fibronectin, acetylcholine receptor of embryonic muscle, and myoblast fusion protein) were found to be degraded more rapidly when cell cultures were treated with tunicamycin, a specific inhibitor of protein glycosylation. The increased cellular degradation of the nonglycosylated proteins was partially arrested by proteinase inhibitors *in vivo* and it was found that nonglycosylated proteins were degraded more rapidly by specific proteases *in vitro*. This led to the proposal (27, 29, 30) that specific cleavage sites are unmasked by the absence of carbohydrate side chains. To our knowledge the mutation reported here, in which abolition of a glycosylation site could lead to accelerated proteolysis and human disease, has not been reported elsewhere.

	19		*		27
Normal	CGG	ACC	AAC	TCC	ACC TTT GTC CAG GCC
	arg	thr	asn	ser	thr phe val gln ala
Mutant	CGG	ACC	AAC	TCC	ATC TTT GTC CAG GCC
	arg	thr	asn	ser	ile phe val gln ala
		↑			

FIG. 3. Result of the mutation found in saposin B-deficient patient YF. The DNA sequence and the deduced protein sequence of both the normal and patient samples are shown. Boldface type indicates the differences between the normal and the patient sequences. The asterisk denotes the position of the potential glycosylation site in the normal sequence. The arrow denotes the potential site of proteolytic cleavage exposed when the asparagine at residue 21 is not glycosylated.

To further inquire whether mutations of the type reported here might be more generalized, we undertook a brief survey of other human proteins by using the Protein Identification Resource (National Biomedical Research Foundation, Washington, version 22) computer data base. The search looked, first, for the presence of glycosylation signals (Asp-Xaa-Ser or Thr) and, second, for arginine or lysine residues within five amino acids on either side of the signal. The results of the search are as follows: (i) fibronectin contains 11 potential glycosylation sites, 3 with proximal proteolytic cleavage sites; (ii) the acetylcholine receptor α -chain precursor contains a single glycosylation site with a lysine two residues away; (iii) 6 lysosomal proteins, excluding prosaposin, contain a total of 24 potential glycosylation sites, 10 with nearby lysine or arginine residues; (iv) 13 unselected glycoproteins exhibited 47 glycosylation sites, 24 with associated lysine and/or arginine residues. Thus, 46% of the 83 glycosylation signals found in this survey have closely associated potential proteolytic sites. These results hint that the molecular defect reported here may be a member of a possible large class of mutations.

We would like to thank Dr. Brian Martin for synthesizing the oligonucleotide primers used in this investigation. This investigation was supported in part by National Institutes of Health Grants NS-08682 (J.S.O.), HD-18983 (J.S.O.), NS-13559 (Y.K.), and by the United Leukodystrophy Foundation (K.A.K.).

- O'Brien, J. S., Kretz, K. A., Dewji, N. N., Wenger, D. A., Esch, F. & Fluharty, A. L. (1988) *Science* **241**, 1098–1101.
- Nakano, T., Sandhoff, K., Stümper, J., Christomanou, H. & Suzuki, K. (1989) *J. Biochem.* **105**, 152–154.
- Morimoto, S., Martin, B. M., Yamamoto, Y., Kretz, K. A., O'Brien, J. S. & Kishimoto, Y. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3389–3393.
- Wenger, D. A., Sattler, M. & Roth, S. (1982) *Biochim. Biophys. Acta* **712**, 639–649.
- Morimoto, S., Martin, B., Kishimoto, Y. & O'Brien, J. S. (1988) *Biochem. Biophys. Res. Commun.* **156**, 403–410.
- Radin, N. S. (1984) in *The Molecular Basis of Lysosomal Storage Disorders*, eds. Brady, R. O. & Barranger, J. A. (Academic, New York), pp. 93–112.
- Fischer, G. & Jatzkewitz, H. (1975) *Hoppe-Seyler's Z. Physiol. Chem.* **356**, 605–613.
- Mehl, E. & Jatzkewitz, H. (1964) *Hoppe-Seyler's Z. Physiol. Chem.* **339**, 260–276.
- Fischer, G. & Jatzkewitz, H. (1978) *Biochim. Biophys. Acta* **528**, 69–76.
- Wenger, D. A. & Inui, K. (1984) in *Molecular Basis of Lysosomal Storage Disorders*, eds. Brady, R. O. & Barranger, J. A. (Academic, New York), pp. 1–18.
- Li, S.-C., Wan, C.-C., Mazzotta, M. Y. & Li, Y.-T. (1974) *Carbohydr. Res.* **34**, 189–193.

12. Li, S.-C., Kihara, H., Serizawa, S., Li, Y.-T., Fluharty, A. L., Mayes, J. S. & Shapiro, L. J. (1985) *J. Biol. Chem.* **260**, 1867–1871.
13. Gärtner, S., Conzelmann, E. & Sandhoff, K. (1983) *J. Biol. Chem.* **258**, 12378–12385.
14. Li, S.-C., Sonnino, S., Tettamanti, G. & Li, Y.-T. (1988) *J. Biol. Chem.* **263**, 6588–6591.
15. Li, S.-C. & Li, Y.-T. (1976) *J. Biol. Chem.* **251**, 1159–1163.
16. Stevens, R. L., Fluharty, A. L., Kihara, H., Kaback, M. M., Shapiro, L. J., Marsh, B., Sandhoff, K. & Fischer, G. (1981) *Am. J. Hum. Genet.* **33**, 900–906.
17. Inui, K., Emmett, M. & Wenger, D. A. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3074–3077.
18. Wenger, D. A., DeGala, G., Williams, C., Taylor, H. A., Stevenson, R. E., Pruitt, J. R., Miller, J., Garen, P. D. & Balentine, J. D. (1989) *Am. J. Med. Genet.* **33**, 255–265.
19. Kretz, K., Ginns, E., Carson, G., Morimoto, S., Kishimoto, Y., Fluharty, A. & O'Brien, J. (1989) *Am. J. Hum. Genet.* **45**, Suppl., A202 (abstr.).
20. Wenger, D. A., Zhang, X.-L., Rafi, M. & DeGala, G. (1989) *Am. J. Hum. Genet.* **45**, Suppl., A13 (abstr.).
21. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239**, 487–491.
22. Kretz, K. A., Carson, G. S. & O'Brien, J. S. (1989) *Nucleic Acids Res.* **17**, 5864.
23. Sano, A. & Radin, N. S. (1988) *Biochem. Biophys. Res. Commun.* **154**, 1197–1203.
24. Fürst, W., Machleidt, W. & Sandhoff, K. (1988) *Biol. Chem. Hoppe-Seyler* **369**, 317–328.
25. Wenger, D. A., Zhang, X.-L., d'Amato, T. A., Dewji, N. N. & O'Brien, J. S. (1989) in *Lipid Storage Disorders: Biological and Medical Aspects*, eds. Salvayer, R., Gatt, S. & Douste-Blazy, L. (Plenum, New York), pp. 337–345.
26. Collard, M. W. & Griwold, M. D. (1987) *Biochemistry* **26**, 3297–3303.
27. Olden, K., Bernard, B. A., Humphries, M. J., Yeo, T.-K., Yeo, K.-T., White, S. L., Newton, S. A., Bauer, H. C. & Parent, J. B. (1985) *Trends Biochem. Sci.* **10**, 78–82.
28. Olden, K., Parent, J. B. & White, S. L. (1982) *Biochim. Biophys. Acta* **650**, 209–232.
29. Beeley, J. G. (1976) *Biochem. J.* **159**, 335–345.
30. Beeley, J. G. (1977) *Biochem. Biophys. Res. Commun.* **76**, 1051–1055.