Glycobiology vol. 11 no. 7 pp. 549–556, 2001

Complete analysis of the glycosylation and disulfide bond pattern of human β-hexosaminidase B by MALDI-MS

Christina Gabriele Schuette[1], Judith Weisgerber, and Konrad Sandhoff[2]

Kekulé-Institut für Organische Chemie und Biochemie, Universität Bonn,Gerhard-Domagk-Str. 1, D-53121 Bonn, Germany

Received on November 22, 2000; revised on January 16, 2001; accepted on March 12, 2001

β-hexosaminidase B is an enzyme that is involved in the degradation of glycolipids and glycans in the lysosome. Mutation in the HEXB gene lead to Sandhoff disease, a glycolipid storage disorder characterized by severe neurodegeneration. So far, little structural information on the protein is available. Here, the complete analysis of the disulfide bond pattern of the protein is described for the first time. Additionally, the structures of the N-glycans are analyzed for the native human protein and for recombinant protein expressed in SF21 cells.

For the analysis of the disulfide bond structure, the protein was proteolytically digested and the resulting peptides were analyzed by MALDI-MS. The analysis revealed three disulfide bonds (C91–C137; C309–C360; C534–C551) and a free cysteine (C487). The analysis of the N-glycosylation was performed by tryptic digestion of the protein, isolation of glycopeptides by lectin chromatography and mass measurement before and after enzymatic deglycosylation. Carbohydrate structures were calculated from the mass difference between glycosylated and deglycosylated peptide. For β-hexosaminidase B from human placenta, four N-glycans were identified and analyzed, whereas the recombinant protein expressed in SF21 cells carried only three glycans. In both cases the glycosylation belongs to the mannose-core- or high-mannose-type, and some carbohydrate structures are fucosylated.

Key words: hexosaminidase B/Sandhoff disease/disulfide bonds/MALDI-MS/N-glycans

Introduction

The human lysosomal β-hexosaminidases are dimeric proteins consisting of two structurally related subunits, α and β. Both the heterodimer hexosaminidase A (HexA) and the β homodimer hexosaminidase B (HexB) are involved in the lysosomal degradation of glycolipids, glycoproteins, and proteoglycans through the release of terminal, β-glycosidically linked N-acetylglucosamines and N-acetylgalactosamines. Mutations in the HEXB gene coding for the β subunit lead to a loss of function of both enzymes and cause Sandhoff disease, a severe lysosomal storage disorder characterized by progressive neurodegeneration (reviewed in Gravel *et al.*, 1995).

Both subunits are synthesized as glycosylated precursor proteins that dimerize in the Golgi apparatus (Proia *et al.*, 1984) and reach the lysosome by the mannose-6-phosphate pathway (Sonderfeld-Fresko and Proia, 1989). The β subunit contains five potential N-glycosylation sites. In a study using site-directed mutagenesis, it has been demonstrated that four of these are utilized (Sonderfeld-Fresko and Proia, 1989). The glycans connected to three N-glycosylation sites in the mature lysosomal HexB have been analyzed by nuclear magnetic resonance (NMR) spectroscopy (O'Dowd *et al.*, 1988).

In the lysosome, the proteins are subjected to proteolytic processing. In the β subunit, a part of the N-terminal sequence is removed along with two internal peptides, leading to a mature enzyme consisting of three peptides βp, βa, and βb, which remain connected by disulfide bond linkages (Mahuran *et al.*, 1988). Because the protein contains seven cysteine residues, a maximum of three disulfide bonds can be formed. The only one of these bonds which has been localized so far is the linkage C309–C360, which connects the peptides βa and βb (Sagherian *et al.*, 1993). The connectivity of the remaining five cysteines is so far unknown. The importance of the disulfide bond connectivity for the enyzmatic function is indicated by the fact that a missense mutation that leads to the severe infantile form of Sandhoff disease causes an exchange of cysteine 534 to tyrosine (Kuroki *et al.*, 1995) and therefore results in the loss of one disulfide bond.

Here, we describe the complete analysis of the disulfide bond structure of HexB. Additionally the glycosylation is analyzed completely by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). In respect to the glycosylation, the mature HexB isolated from human placenta (plHexB) is compared to the recombinant protein expressed in SF21 cells using the baculovirus expression system (bcHexB).

Results

Identification of two disulfide bonds and a free cysteine

Pure bcHexB was subjected to tryptic digestion with and without prior reduction and alkylation. Figure 1 shows the sequence of the protein after cleavage of the signal sequence but without proteolytic processing in the lysosome (proform), as it is expressed in SF21 cells. The putative tryptic peptides are indicated. In the MALDI-MS spectra (Figure 2), several tryptic peptides can be identified. Two signals were observed

[1]Present address: Max-Planck-Institut for Biophysical Chemistry, Dept. 190 Am Fassberg 11, D-37077 Göttingen, Germany.
[2]To whom correspondence should be addressed

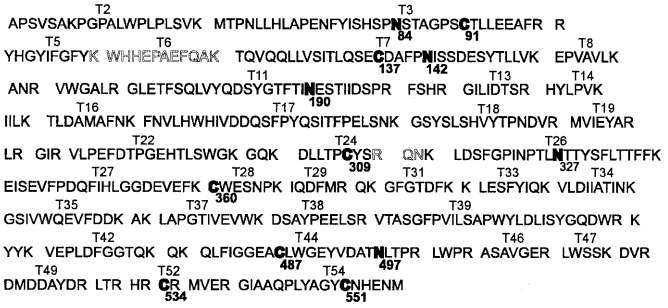

Fig. 1. Amino acid sequence of the proform of HexB (Swiss Prot Access no. P07686). Tryptic cleavage sites are indicated by gaps, tryptic peptides are named T and numbered from 1 to 54. Amino acids that are removed by processing in the mature enzyme are shown as open characters. Cysteines and potential N-glycosylation sites are marked with boldface.

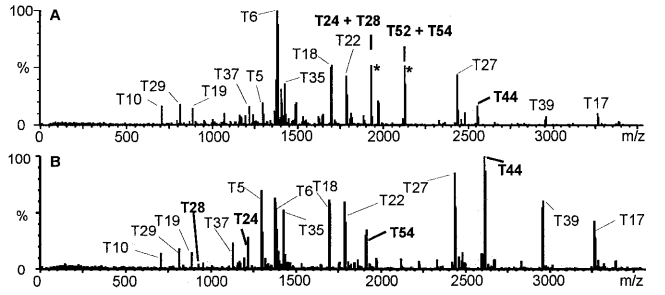

Fig. 2. MALDI mass spectra of the tryptic digest of bcHexB before (**A**) and after (**B**) reduction and IAA alkylation. Tryptic peptides are annotated. Cysteine-containing peptides are marked with boldface, signals that disappear on reduction are marked by an asterix. Spectra were acquired in reflectron mode using α-cyano-4-hydroxycinnamic acid as the matrix. Sequence coverage is 50%, due to nontryptic cleavage (see Figure 3).

that occur in the digest of the untreated protein (Figure 2A, asterix) and disappear on reduction (Figure 2B). The signal at 1927.9 Da corresponds to the expected mass of the disulfide bonded species T24–T28 (disulfide bond between C309 and C360), and the signal at 2127.0 Da corresponds to T52–T54 (disulfide bond between C534 and C551). The corresponding free and iodacetamide (IAA)-modified peptides T24, T28, and T54 are detected in the reduced and alkylated digest (Figure 2B, bold; Table I). Due to its low molecular weight, the dipeptide T52 could not be detected. These results imply the existence of a disulfide bond between C309 and C360, which has been described previously (Sagherian *et al.*, 1993), and of a second disulfide bond between C534 and C551.

The peptide T44, which contains C487 and the putative N-glycosylation site N497, was detected as a free and unglycosylated peptide in both spectra (Figure 2, bold). This demonstrates that N497 is not glycosylated, which is in agreement with previous results (Sonderfeld-Fresko and Proia, 1989). The detection as a free peptide also implies that C487 is not involved in a disulfide bond. To confirm this assumption, HexB was subjected to alkylation with IAA or vinylpyridine (VP) without prior reduction and subsequently digested by trypsin. In the resulting MALDI-MS spectra, T44 is the only peptide that carries a modification (Table II), demonstrating that C487 is present as a free cysteine in native HexB.

Table I. Theoretical and measured masses of tryptic peptides of recombinant HexB that are relevant for the disulfide bond arrangement

Tryptic peptide	Expected mass	Measured mass
Digest without reduction		
T24 + T28	1927.9	1927.8
T52 + T54	2126.9	2127.0
T44	2553.2	2553.2
Digest after reduction and alkylation (IAA)		
T28	920.4	920.4
T24	1124.5	1124.5
T54	1908.8	1908.8
T44	2610.3	2610.3

Monoisotopic masses of the $M+H^+$ species are measured in the reflectron mode. Mass accuracy is in the range of 500 p.p.m. Ions below 500 Da are suppressed by the matrix suppression lens.

Table II. Theoretical and measured masses of tryptic peptide T44 of HexB without alkylation, and after alkylation of the protein with either IAA of VP without prior reduction.

	Without alkylation	Alkylation with IAA	Alkylation with VP
Expected mass	2553.2	2610.3	2658.3
Measured mass	2553.2	2610.3	2658.4

Monoisotopic masses of the $M+H^+$ species are measured in the reflectron mode. Mass accuracy is in the range of 500 p.p.m.

To confirm the existence of the disulfide bonds between C309 and C360 and between C534 and C551, the peptide species were isolated from the tryptic digest by reversed phase high-performance liquid chromatograph (RP-HPLC) and masses were measured without treatment, after reduction, and after subsequent alkylation. In both cases the disulfide-bound species disappeared on reduction, giving rise to the free peptides that underwent a mass shift on alkylation (Table III), thus proving the existence of these two disulfide bonds.

The results presented so far demonstrate the existence of two disulfide bonds (C309–C360 and C534–C551) and of a free cysteine at C487. The experiments were repeated with plHexB and gave identical results, except for a mass shift in T24 due to the loss of R312 by processing in the mature enzyme (Mahuran *et al.*, 1988) (data not shown).

Identification of nontryptic peptides

The MALDI-MS of the tryptic digest of pure HexB showed several masses that could not be assigned to tryptic peptides but were present in all preparations measured. This was suprising because of the high purity of the protein as demonstrated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (data not shown). For identification, several of these peptides were isolated by RP-HPLC and sequenced by MALDI–post source decay (PSD). Their sequences correspond to HexB sequences, with at least one terminus produced by a nontryptic cleavage. During the course of the investigation, 30 such peptides were identified (Figure 3, Table IV). For

Table III. Masses of disulfide bound species isolated by RP-HPLC

	Expected mass	Measured mass	Species
Species T24 + T28			
untreated	1927.9	1927.7	T24 + T28
after reduction	863.4	863.4	T28
	1067.5	1067.5	T 24
after reduction + alkylation	920.4	920.2	T28 + IAA
	1124.5	1124.4	T24 + IAA
Species T52 + T54			
untreated	2126.8	2126.7	T52 + T54
after reduction	1851.8	1851.8	T54
	277.1	—	T52
after reduction + alkylation	1956.9	1956.8	T54 + VP
	382.2	—	T52 + VP

Masses are measured after isolation without prior treatment, after reduction with after dithiothreitole and after alkylation with either IAA or VP. Monoisotopic masses of the M+H$^+$ species are measured in the reflectron mode. Mass accuracy is in the range of 500 p.p.m. Ions below 500 Da are suppressed by the matrix suppression lens.

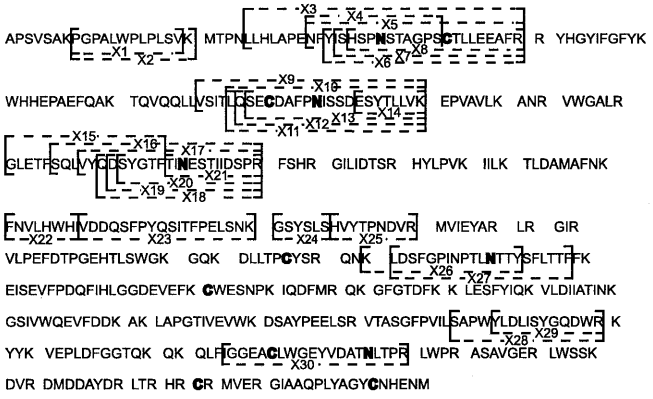

Fig. 3. Nontryptic peptides of HexB: The amino acid sequence of pro-HexB is shown. Tryptic cleavage sites are indicated as gaps, potential N-glycosylation sites and cysteines are marked with boldface. Nontryptic peptides that were identified are marked by brackets and named X with a running number from 1 to 30.

reference, they were named X with a number indicating their position in the sequence. Their generation may be due to a contamination of trypsin by chymotrypsin as proposed by O'Dowd *et al.* (1988). But because all attempts to suppress this activity by using different preparations of trypsin, shorter incubation times, or less protease were unsuccessful, no final explanation for their existence can be given. Figure 4 shows the sequence and PSD spectra of three of these peptides (X7, X13, and X16), which contain N-glycosylation sites and are therefore of interest for the analysis of the glycoforms.

Identification of a third disulfide bond

As described above, two disulfide bonds and the free cysteine were identified. The existence of a disulfide bond between the remaining two cysteines (C91 and C137) is implied by the fact

Table IV. Masses and sequences of nontryptic peptides of HexB that were identified during the analysis.

Peptide	Mass	Sequence	Peptide	Mass	Sequence
X1	1060.6	P52–L61	X16	2735.3	S176–R199
X2	1374.8	P52–K64	X17	2407.1	V179–R199
X3	3415.7	L69–R99	X18	2145.0	Q181–R199
X4	1579.7	N76–S90	X19	2016.9	D182–R199
X5	2642.2	N76–R99	X20	1901.9	S183–R199
X6	2381.1	Y78–R99	X21	1346.7	T188–R199
X7	2218.0	I79–R99	X22	952.5	F231–H237
X8	2017.9	H81–R99	X23	2328.1	I230–K257
X9	2860.4	V129–K154	X24	613.3	G258–S263
X10	2332.0	L133–V153	X25	1100.5	H264–R272
X11	2460.1	L133–K154	X26	1781.9	K315–Y330
X12	2347.1	Q134–K154	X27	2350.1	L316–F336
X13	2219.0	S135–K154	X28	1969.9	S446–R461
X14	952.5	E147–K154	X29	1528.7	Y450–R461
X15	1954.9	G171–F187	X30	2166.4	I482–R501

Sequences were determined by PSD or peptides were identified by their mass and glycosylation. Monoisotopic masses of the M+H$^+$ species are measured in the reflectron mode. Mass accuracy is in the range of 500 p.p.m..

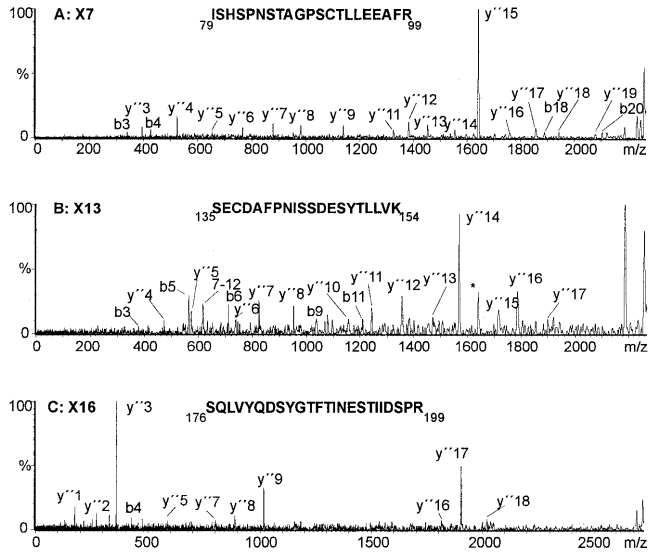

Fig. 4. MALDI-PSD spectra of three nontryptic peptides. (**A**) Peptide X7 (amino acid 79–99), containing cysteine C91 and N-glycosylation site N84. (**B**) Peptide X13 (amino acid 135–154), containing cysteine C137 and N-glycosylation site N142. The signal marked by an asterisk is caused by a slight contamination by peptide X7 (corresponds to y″15 of X7). (**C**) Peptide X16 (amino acid 176–199), containing N-glycosylation site N190. Spectra were acquired using α-cyano-4-hydroxycinnamic acid as the matrix.

that the N-terminal peptide is connected to the rest of the protein by a disulfide bond in the mature protein (Hubbes *et al.*, 1989). Considering the results presented above, such a bond can only exist between C91 and C137, but no direct proof for the existence of this bond has been presented so far.

Because both tryptic peptides which should be involved in this bond carry a N-glycosylation site (Figure 1) and several nontryptic cleavages were identified is this region (Figure 3), a heterogeneous mixture of disulfide bonded species had to be expected. To isolate this mixture from the tryptic digest, concanavalin A (Con A) chromatography was performed on the tryptic peptides produced from untreated plHexB. The eluate of the Con A column was separated by RP-HPLC. The RP-HPLC elution profile was compared to that of a sample treated identically but reduced prior to digestion. The elution profile of the untreated protein showed two signals that were not present in the profile of the reduced protein. In the fraction corresponding to one of these signals, a mixture of species with masses in the range of 7000–8500 Da was detected (data not shown). This corresponds to the mass range expected for glycosylated, disulfide bonded species containing C91 and C137. After deglycosylation by peptide N-glycanase F (PNGase F), the complexity of the mixture was reduced and the mass range was shifted to 4000–5000 Da. These species disappear on reduction, giving rise to a mixture of nontryptic peptides (Figure 5, Table V) that had been identified previously. It was possible to assign all masses detected in the mass range of 4000–5000 Da to disulfide-bonded species. Each of these was identified as a combination of nontryptic peptides, one of which contains C91 and the other contains C137 (Table V). One of the major disulfide-bonded species corresponds to the combination of X7 and X13. The PSD spectra that led to the identification of these two nontryptic peptides are shown in Figure 4. These results confirm the postulated existence of a disulfide bond between C91 and C137.

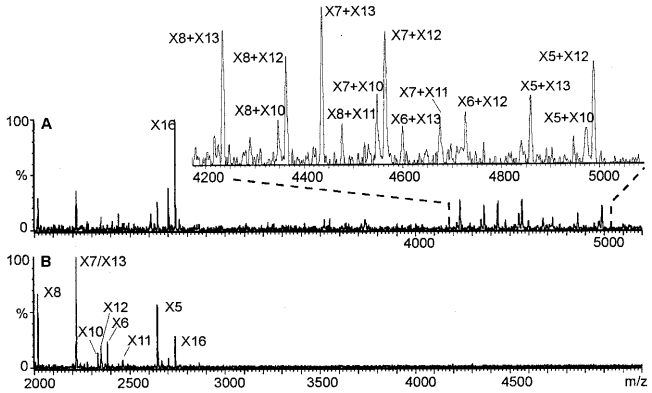

Fig. 5. MALDI mass spectra of RP-HPLC isolated disulfide-linked glycopeptides after deglycosylation by PNGase F (**A**) and after subsequent reduction (**B**). Disulfide bonded species are detected at masses between 4000 and 5000 Da (enlargement in panel **A**) and can be assigned to combinations of nontryptic peptides. Spectra were acquired in reflectron mode using α-cyano-4-hydroxycinnamic acid as the matrix.

Analysis of the glycosylation

Samples of plHexB and recombinant bcHexB were reduced and alkylated by IAA and subsequently digested by trypsin. The mixture of tryptic peptides was subjected to Con A chromatography and the glycopeptide-containing eluate of the Con A column was separated by RP-HPLC. The resulting fractions were analyzed by MALDI-MS. After deglycosylation by PNGase F the mass measurement was repeated. The resulting carbohydrate-free peptides were identified as tryptic peptides by their mass or sequenced by MALDI-PSD (see Figure 4). Glycosylation was calculated from the mass difference between glycosylated and deglycosylated species. The resulting mass spectra before and after deglycosylation are shown in Figure 6 (plHexB) and Figure 7 (bcHexB).

The groups of signals with a mass difference of 162 Da (hexose), which are observed in the MALDI-MS spectra of several of the isolated glycopeptides, are due to a heterogeneous glycosylation of the high-mannose type. After deglycosylation, one carbohydrate-free peptide is detected.

The resulting glycoforms that were deduced from the spectra are presented in Table VI and summarized in Table VII. The carbohydrate connectivites displayed in Table VI are putative. Because MS does not give information on connectivites, structures were drawn in analogy to known N-glycosylation structures (Kornfeld and Kornfeld, 1985). For fucose-containing structures, a terminal α1,3-connected mannose was

Table V. Masses of the HPLC-isolated mixture of disulfide bound glycopeptides measured after deglycosylation by PNGase F

Masses measured before reduction							
Peptide	X8+X10	X8+X11	X8+X12	X8+X13	X7+X10	X7+X11	X7+X12
M theor.	4348.7	4477.9	4363.7	4236.6	4548.9	4678.1	4565.0
M meas.	4348.3	4477.7	4364.0	4236.0	4548.0	4676.1	4564.4
Peptide	X7+X13	X6+X12	X6+X13	X5+X10	X5+X11	X5+X12	X5+X13
M theor.	4436.8	4728.1	4600.0	4973.4	5102.6	4989.4	4861.3
M meas.	4436.7	4727.8	4600.0	4973.1	5101.7	4989.0	4860.6
Masses measured after reduction							
Peptide	X5	X6	X7	X8	X10	X11	X12
M theor.	2642.2	2381.1	2218.0	2017.9	2331.0	2460.1	2347.1
M meas.	2642.2	2381.1	2218.2	2018.0	2330.2	2460.2	2347.1

Mass measurements were performed before and after reduction by dithiothreitole. M: mass of the singly charged ion ($M+H^+$). Masses above 3000 Da are average masses. Mass accuracy is in the range of 500 p.p.m.

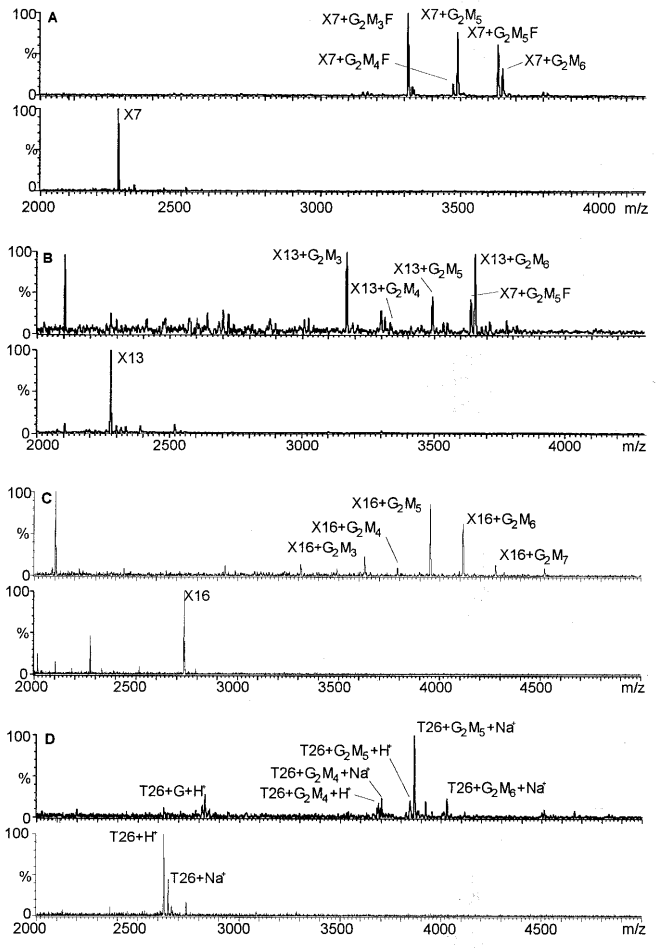

Fig. 6. MALDI mass spectra of glycopeptides derived from plHexB before (upper panel) and after (lower panel) treatment with PNGase F. Glycan structures are indicated. M = mannose, G = N-acetylglucoseamine, F = fucose. (**A**) Glycopeptide X7, containing N84; (**B**) glycopeptide X13, containing N142l (**C**) glycopeptide X16, containing N190; (**D**) glycopeptide T26, containing N327. Unless indicated differently, all ions are $M+H^+$ ions. Spectra were acquired in reflectron mode using α-cyano-4-hydroxycinnamic acid as the matrix.

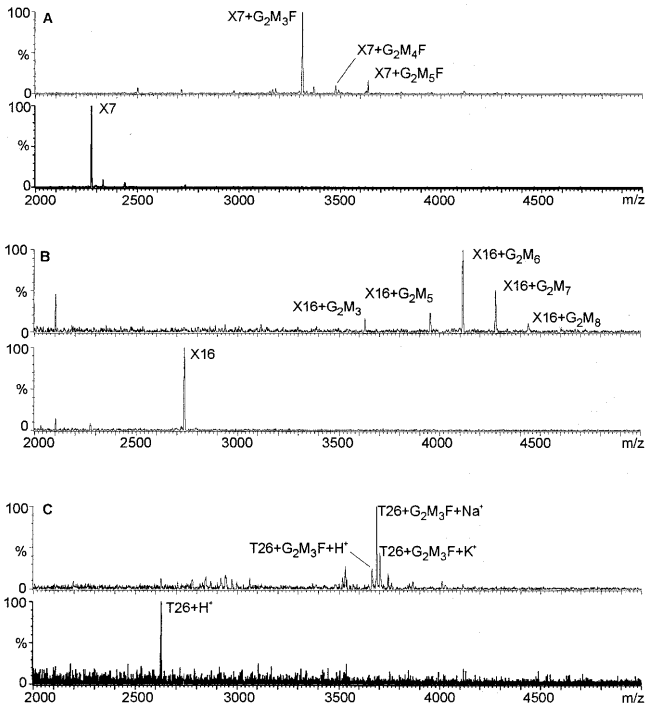

Fig. 7. MALDI mass spectra of glycopeptides derived from bcHexB before (upper panel) and after (lower panel) treatment with PNGase F. Glycan structures are indicated. M = mannose, G = N-acetylglucoseamine, F = fucose. (**A**) Glycopeptide X7, containing N84; (**B**) glycopeptide X16, containing N190; (**C**) glycopeptide T26, containing N327. Unless indicated differently, all ions are $M+H^+$ ions. Spectra were acquired in reflectron mode using α-cyano-4-hydroxycinnamic acid as the matrix.

assumed, which is in accordance with the known specificities of fucosyltransferases in man and insect cells (Staudacher *et al.*, 1999).

The analysis revealed several differences between the protein purified from human tissue (plHexB) and the recombinant protein (bcHexB). The glycan on N84 is only partially fucosylated in plHexB, whereas the fucosylation is homogeneous in bcHexB. In bcHexB, no glycan is detected on N142, but in plHexB a glycan of the high-mannose type is found attached to this site. Both proteins carry a high-mannose type glycan on N190, but glycosylation differs on N327. Here, bcHexB carries a homogeneous, fucosylated glycan of the type $GlcNAc_2Man_3Fuc$, whereas there are several glycans detected for plHexB. A minor signal (see Figure 6D) corresponds in mass to the glycosylation by a single GlcNAc. However, this signal disappears on PNGase F treatment, which contradicts the reported specificity of the glycanase (Gosselin *et al.*, 1993). But because this unusal glycosylation had been described previously at this site in a study using NMR spectroscopy for glycan analysis (O'Dowd *et al.*, 1988), we tentatively assign this structure. The discrepancy with the reported glycanase specificity should be kept in mind, though.

The major signal in the spectrum of this glycopeptide corresponds in mass to a nonfucosylated high-mannose type glycan (see Figure 6D). This glycan is described here for the first time.

Discussion

This report is the first to describe the complete analysis of post-translational modifications of HexB isolated from human tissue and of recombinant HexB expressed in SF21 cells using the baculovirus system. This allows an interesting comparison between the native human protein and the recombinant protein. Because both proteins are functional (Boose *et al.*, 1990), it was expected that the disulfide bond linkages are identical for both. This was confirmed by the MALDI analysis of the disulfide bonds. However, the N-glycosylation is not identical for the two different protein preparations. The most pronounced difference is the missing N-glycosylation of N142 in the recombinant protein. This may account for the reduced solubility of this protein in buffers of low ionic strength (unpublished observation). Because both proteins show identical

Table VI. Glycoforms of HexB from human placenta and recombinant HexB. Glycopeptide-containing HPLC fractions were subjected to MALDI-MS before and after deglycosylation by peptide-N-glycanase F. Connectivities displayed are hypothetical. Squares, G = GlcNAc; circles, M = Man; triangles, F = Fuc. Filled figures represent sugars present in all glycoforms detected. Monoisotopic masses of the M + H$^+$ species are measured in the reflectron mode, unless indicated differently.

	untreated		after deglycosylation	
plHexB	measured mass	Species	measured mass	Species
N84	3312.5	X7 + G_2M_3F	2275.5	X7
	3473.4	X7 + G_2M_4F		
	3490.5	X7 + G_2M_5		
	3636.7	X7 + G_2M_5F		
	3652.6	X7 + G_2M_6		
N142	3167.5	X13 + G_2M_3	2275.6	X13
	3330.0	X13 + G_2M_4		
	3491.5	X13 + G_2M_5		
	3653.5	X13 + G_2M_6		
N190	3627.5*	X16 + G_2M_3	2735.4	X16
	3789.5*	X16 + G_2M_4		
	3952.2*	X16 + G_2M_5		
	4114.3*	X16 + G_2M_6		
	4276.5*	X16 + G_2M_7		
N327	2827.7	T26 + G	2626.3	T26
	3678.8	T26 + G_2M_4		
	3840.7	T26 + G_2M_5		
	4024.9#	T26 + G_2M_6		
bcHexB				
N84	3312.6	X7 + G_2M_3F	2275.0	X7
	3474.2	X7 + G_2M_4F		
	3636.7	X7 + G_2M_5F		
N190	3628.4*	X16 + G_2M_3	2735.6	X16
	3952.2*	X16 + G_2M_5		
	4114.7*	X16 + G_2M_6		
	4276.7*	X16 + G_2M_7		
	4439.1*	X16 + G_2M_8		
N327	3662.1	T26 + G_2M_3F	2625.4	T26

*average mass
#sodium adduct

Table VII. Glycans found on the different glycosylation sites in plHexB and bcHexB

Glycosylation site	Glycan plHexB	Glycan bcHexB
N84	$GlcNAc_2Man_{3-5}Fuc$, $GlcNAc_2Man_{5-6}$	$GlcNAc_2Man_{3-5}Fuc$
N142	$GlcNAc_2Man_{3-6}$	—
N190	$GlcNAc_2Man_{3-7}$	$GlcNAc_2Man_{3-8}$
N327	GlcNAc, $GlcNAc_2Man_{4-6}$	$GlcNAc_2Man_3Fuc$
N495	—	—

Glycan structures are deduced from the mass differences between the glycosylated and the deglycosylated peptide.

specific activities against the synthetic substrate 4-methylumbelliferyl-2-acetamido-2-deoxy-β-D-glucopyranoside (MUG) (Schuette, 1999), the difference in glycosylation does not affect the active site or have a major influence on the structure of the protein.

The availability of the complete disulfide bond structure should make structure predictions by molecular modeling more reliable in the future and makes it possible to compare HexB to homologous proteins like the chitobiase (ChB) from *Serratia marcencens*. As the human hexosaminidases, this protein protein belongs to family 20 of glycosyl hydrolases. It is the only protein of this family for which a crystal structure exsists and shows a high homology to HexA and HexB in its central region, which contains the active site. A homology model for the active domain of HexA has been proposed on the basis of this structure (Tews *et al.*, 1996a). Like HexB, ChB contains three disulfide bonds that are arranged sequentially along the sequence. The catalytic domain of both proteins contains one disulfide bond (C505–C578 in ChB [Tews *et al.*, 1996b] and C309–C360 in HexB). Although the positions of the cysteines are not exactly conserved, they show only a slight shift in their position and span a homologous sequence of similar length containing the catalytically active glutamate in both proteins (Glu540 in ChB [Tews *et al.*, 1996a], Glu 355 in HexB [Pennybacker *et al.*, 1997; Liessem *et al.*, 1995]).

The position of the remaining two disulfide bonds does not seem to be equivalent in ChB and HexB, though. Although ChB contains two disulfide bonds which span short sequences N-terminal to the catalytic domain, HexB contains just one bond in this region that spans a much longer sequence containing the processing site between the βp- and the βb-peptide. The disulfide bond at the very C-terminus of HexB does not find an equivalent in ChB either, because ChB does not contain any disulfide bonds C-terminal of the catalytic domain. Whether these two bonds stabilize regions of HexB that perform tasks specific for the human lysosomal hexosaminidases, for example the formation of dimers, can only be speculated on at the moment. A mutation that exchanges Cys534 to Tyr and therefore abolishes the C-terminal disulfide bond leads to a severe infantile form of Sandhoff disease with a complete loss of HexB activity (Kuroki *et al.*, 1995). This demonstrates the importance of this disulfide bond for the function of human HexB.

The glycosylation of the protein purified from human placenta had been analyzed previously by O'Dowd *et al.* (1988) using NMR spectroscopy. This method allows an exact analysis of connectivities between sugars; the exact number of identical sugars with identical connectivites is hard to determine, though. This is not a problem with MALDI-MS. It is therefore not suprising that we detected a higher heterogeneity in the carbohydrate structures than was reported by O'Dowd *et al.* (1988). One major difference between our results and the glycan structures published by O'Dowd *et al.* (1988) is the glycosylation on N327. In their study they reported only the unusal glycosylation by a single GlcNAc residue. This result was confirmed by our data, but we found an additional major glycan of the high-mannose type in this position. Moreover, the glycan structure of N84 was not determined by O'Dowd *et al.* (1988), because the βp-peptide is lost in the purification procedure they used. This glycan was analyzed for the first time in the study presented here and was found to be a partially fucosylated high-mannose type glycan.

The combination of proteinchemical methods and MALDI-MS which has been used to determine the disulfide bond linkages and the glycosylation of HexB is extremely sensitive and especially suited for the analysis of complex mixtures. Although it has been used for similar purpose for some time (Yazdanparast *et al.*,

1987; Carr and Roberts, 1986), the special properties of the lysosomal protein HexB gave rise to some new problems that have been solved in the work presented here. The extreme stability of HexB against proteolytic enzymes made it necessary to use high concentrations of proteases. This led to an unspecific cleavage in addition to the expected specific proteolysis. Although this phenomenon had been described previously for HexB (O'Dowd *et al.*, 1988), it had never been analyzed in detail. We successfully identified the products of the unspecific cleavage by MALDI-PSD sequencing and were therefore able to determine the localization of the disulfide bonds and the glycan structures. These results may facililtate the analysis of posttranslational modification in lysosomal proteins by MALDI-MS in the future.

Materials and methods

Purification of HexB from human placenta

Homogenization, ammonium sulfate precipitation, chromatography on Con A Sepharose, and chromatography on Octyl-sepharose were performed as described for the purification of acid sphingomyelinase (Lansmann *et al.*, 1996).

Cation exchange chromatography: The flowthrough of the Octylsepharose chromatography was extensively dialyzed against ion exchange buffer (10 mM phosphate buffer, pH 6.0) and applied to a POROS HS column (4.6 × 10 mm, PerSeptive) at a flow rate of 10 ml/min. The column was washed with 15 column volume (CV) of ion exchange buffer. Elution was performed using a gradient of 0–1000 mM NaCl in 10 CV of ion exchange buffer and keeping up 1000 mM NaCl for an additional 3 CV. Fractions were tested for HexB activity using the synthetic substrate MUG as described (Pennybacker *et al.*, 1996). Fractions containing HexB were collected and pooled.

Desalting by gel filtration: The column (HiPrep Desalting 26/10, Pharmacia) was equilibrated in 5 CV of affinity buffer (10 mM phosphate, pH 6.0, 100 mM NaCl) at 15 ml/min. Fifteen milliliters of eluate of the cation exchange chromatography were applied to the column and eluted in 1.7 CV of affinity buffer.

Preparation of affinity gel: Five grams of bromcyan-activated Sepharose 4 B (Sigma) were swollen in 1 mM HCl and washed with 1 L of 1 mM HCl. After removing the washing solution by suction filtration the gel was added quickly to a solution of 25 mg of 2-acetamido-N-(ε-aminocaproyl)-2-deoxy-β-D-glucopyranosylamine (Sigma) in coupling buffer (100 mM $NaHCO_3$/NaOH, pH 10.0). The mixture was shaken at 4°C over night. After removing the coupling solution by suction filtration the gel was washed several times with changes of coupling buffer and washing buffer (100 mM citrate buffer, pH 4.0) and was finally stored in affinity buffer.

Affinity chromatography: A column containing 20 ml of affinity gel (18 mm × 80 mm) was equilibrated in affinity buffer and loaded with desalted eluate of the cation exchange chromatography (4 U/ml gel) at 2 ml/min. The column was washed with 2.5 CV of affinity buffer and elution was performed using 2.5 CV of 125 µM N-acetyl-β-D-glucono-1,5-lactone (Toronto Research Chemicals) in affinity buffer.

Purification of recombinant HexB

HexB was expressed in SF21 cells as described for acid sphingomyelinase (Bartelsen *et al.*, 1998) using the recombinant baculovirus described previously (Boose *et al.*, 1990). The medium was collected by centrifugation at 20,000 × *g*. The pH of the medium was set to 4.0 by addition of 1:10 volume of 200 mM citrate buffer, pH 4.0.

Cation exchange chromatography: After filtration (0.4 µm) the medium was applied to a column (28 mm × 105 mm, CV = 65 ml) containing SP Sepharose fast flow (Pharmacia) at a flow rate of 20 ml/min. The column was washed with 3 CV of 20 mM citrate buffer, pH 4.0. Elution was performed at a flow rate of 10 ml/min using 1 CV of 10 mM phosphate buffer, pH 6.0, a gradient of 0–500 mM NaCl in this buffer in 5 CV and an additional 3 CV of 500 mM NaCl in this buffer. The fractions containing HexB were collected and subjected to further purification. Desalting and affinity chromatography were performed as described for HexB from human placenta.

Tryptic digestion

Lyophilized modified trypsin was resuspended at a concentration of 1 µg/µl as described by the supplier. Trypsin was added at a concentration of 1:20 (w/w) to a solution of 50–100 µM pure HexB in 25 mM NH_4CO_3. The mixture was incubated at 37°C overnight, and the digestion was stopped by freezing or by acidification with trifluoracetic acid (TFA).

Reduction

To a solution of protein or peptide in 25 mM NH_4CO_3 dithiothreitole was added to a concentration of 10 mM. The solution was incubated at 95°C for 15 min.

Alkylation with IAA

To a solution of protein or peptide in 25 mM NH_4CO_3 IAA was added to a concentration of 22 mM. The solution was incubated in the dark at room temperature for 15 min.

Isolation of glycopeptides

A column containing 100 µl of Con A Sepharose (Sigma) was equilibrated in loading buffer (50 mM phosphate buffer, pH 7.0, 200 mM NaCl) and the tryptic digest of 200 µg of HexB was added. The column was washed with 2 ml of loading buffer. Elution was performed using 1 ml of 10% α-D-methylglucoside in loading buffer. The eluate was then subjected to RP-HPLC as described below.

RP-HPLC

RP-HPLC was performed on a Smart System (Pharmacia) using a Vertex column (2 × 250 mm) filled with Nucleosil C18–120 µm material at a flow rate of 100 µl/min. TFA (0.1%) in water was used as solvent A, 0.06% TFA in 70% acetonitrile and 30% isopropanol as solvent B.

Tryptic peptides: The column was equilibrated in 5% solvent B. After injection of the sample the column was washed for 15 min with 5% solvent B. Elution was achieved by a gradient of 5% solvent B to 88% solvent B in 55 min.

Glycopeptides: The column was equilibrated in 5% solvent B. After injection of the sample the column was washed for 15 min with 5% solvent B. Elution was achieved by a gradient of 5% solvent B to 70% solvent B in 65 min.

Deglycosylation by PNGase F

The glycopeptide was isolated by RP-HPLC. The solvent was removed in a vacuum-concentration centrifuge and the peptide was resuspended in 20 µl of 25 mM NH_4CO_3. Two hundred units of PNGase F (New England Biolabs) were added and the mixture was incubated at 37°C for 2 h. Digestion was stopped by freezing or by acidification with TFA.

MALDI-MS

Matrix solutions: α-Cyano-4-hydroxycinnamic acid (Sigma): 10 mg/ml in 50% acetonitrile, 50% water containing 0.1% TFA. Sinnapinic acid (Sigma): 10 mg/ml in 40% acetonitrile, 60% water containing 0.1% TFA.

Sample preparation: For RP-HPLC fraction, 1 µl of sample was mixed with 1 µl of matrix solution on the target and dried at room temperature. For digestion mixtures, 1 µl of sample was mixed with 0.5 µl of 5% TFA in water and 1 µl of matrix solution on the target and dried at room temperature.

Measurement: Measurements were performed on a TOF Spec E (Micromass) at an acceleration voltage of 20 kV. An extraction voltage of 19.5 kV and a focus voltage of 15.5 kV were used. For linear measurements (above 4 kDa) pulse voltage was set to 1200 V, for reflectron measurements a pulse voltage of 2200 V was used. Measurements were performed at threshold laser energy. Calibration was achieved using commercially available peptides (Sigma).

PSD was performed as described by the supplier.

Acknowledgments

We thank Dr. Richard L. Proia for supplying the recombinant baculovirus containing HexB cDNA. This work was supported by the Deutsche Forschungs Geimeinschaft Grant SFB 284 and by a PhD scholarship of the Boehringer Ingelheim Fonds für medizinische Grundlagenforschung for C.G. Schuette.

Abbreviations

bcHexB, protein expressed in SF21 cells using the baculovirus expression system; ChB, chitobiase; Con A, concanavalin A; CV, column volume; Fuc, fucose; GlcNAc, N-acetyl-glucosamine; HexA, β-hexosaminidase A; HexB, β-hexosaminidase B; IAA, iodacetamide; MALDI, matrix-assisted laser desorption/ionization; Man, mannose; MS, mass spectrometry; MUG: 4-methylumbelliferyl-2-acetamido-2-deoxy-β-D-glucopyranoside; NMR, nuclear magnetic resonance; plHexB, protein purified from human placenta; PNGase F, peptide N-glycanase F; PSD, post source decay; RP-HPLC, reversed phase high-performance liquid chromatography; TFA, trifluoracetic acid; VP, vinylpyridine.

References

Bartelsen, O., Lansmann, S., Nettersheim, M., Lemm, T., Ferlinz, K., and Sandhoff, K. (1998) Expression of recombinant human acid sphingomyelinase in insect Sf21 cells: purification, processing and enzymatic characterization. *J. Biotechnol.*, **63**, 29–40.

Boose, J.A., Tifft, C.J., Proia, R.L., and Myerowitz, R. (1990) Synthesis of a human lysosomal enzyme, β-hexosaminidase B, using the baculovirus expression system. *Prot. Exp. Purif.*, **1**, 111–120.

Carr, S.A., and Roberts, G.D. (1986) Carbohydrate mapping by mass spectrometry: a novel method for identifying attachment sites of Asn-linked sugars in glycoproteins. *Anal. Biochem.*, **157**, 396–406.

Gosselin, S., Payle, K., and Viswanatha, T. (1993) Substrate specificity of peptide:N-glycosidase from Flavobacterium meningosepticum. *Glycobiology*, **3**, 419–421.

Gravel, R.A., Clarke, J.T.R., Kaback, M.M., Mahuran, D.J., Sandhoff, K., and Suzuki, K. (1995) The GM2 gangliosidoses. In Scriver, C.R., Beaudet, A.L., Sly, W.S., and Valle, D., eds., *The Metabolic and Molecular Bases of Inherited Disease*. McGraw-Hill, New York, pp. 2839–2879.

Hubbes, M., Callahan, J., Gravel, R.A., and Mahuran, D.J. (1989) The amino-terminal sequences in the pro-α and -β polypeptides of human lysosomal β hexosaminidase A and B are retained in the mature isozymes. *FEBS Lett.*, **249**, 316–320.

Kornfeld, R., and Kornfeld, S. (1985) Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.*, **54**, 631–664.

Kuroki, Y., Itho, K., Nadaoka, Y., Tanaka, T., and Sakuraba, H. (1995) A novel missense mutation (C522Y) is present in the β-hexosaminidase β-subunit gene of a Japanese patient with infantile Sandhoff disease. *Biochem. Biophys. Res. Commun.*, **212**, 564–571.

Lansmann, S., Ferlinz, K., Hurwitz, R., Bartelsen, O., Glombitza, G.J., and Sandhoff, K. (1996) Purification of acid sphingomyelinase from human placenta: characterization and N-terminal sequence. *FEBS Lett.*, **399**, 227–231.

Liessem, B., Glombitza, G.J., Knoll, F., Lehmann, J., Kellermann, J., Lottspeich, F., and Sandhoff, K. (1995) Photoaffinity labeling of human lysosomal β-hexosaminidase B. Identification of Glu-355 at the substrate binding site. *J. Biol. Chem.*, **270**, 23693–23699.

Mahuran, D.J., Neote, K., Klavins, M.H., Leung, A., and Gravel, R.A. (1988) Proteolytic processing of pro-α and pro-β precursors from human β-hexosaminidase. *J. Biol. Chem.*, **263**, 4612–4618.

O'Dowd, B.F., Cumming, D.A., Gravel, R.A., and Mahuran, D. (1988) Oligosaccharide structure and amino acid sequence of the major glycopeptides of mature human β hexosaminidase. *Biochemistry*, **27**, 5216–5226.

Pennybacker, M., Liessem, B., Moczall, H., Tifft, C.J., Sandhoff, K., and Proia, R.L. (1996) Identification of domains in human β-hexosaminidase that determine substrate specificity. *J. Biol. Chem.*, **271**, 17377–17382.

Pennybacker, M., Schuette, C.G., Liessem, B., Hepbildikler, S.T., Kopetka, J.A., Ellis, M.R., Myerowitz, R., Sandhoff, K., and Proia, R.L. (1997) Evidence for the involvement of Glu-355 in the catalytic action of human β-hexosaminidase B. *J. Biol. Chem.*, **272**, 8002–8006.

Proia, R.L., d´Azzo, A., and Neufeld, E.F. (1984) Association of α- and β-subunits during the biosynthesis of β-hexosaminidase in cultured human fibroblasts. *J. Biol. Chem.*, **259**, 3350–3354.

Sagherian, C., Poroszlay, S., Vavougios, G., and Mahuran, D.J. (1993) Proteolytic processing of the proβ chain of β-hexosaminidase occurs at basic residues contained within an exposed disulfide loop structure. *Biochem. Cell Biol.*, **71**, 340–347.

Schuette, C.G. (1999) Structural analysis of β-hexosaminidase B and GM2-activator protein by proteinchemical methods and MALDI-mass spectrometry. Ph.D. diss., University of Bonn.

Sonderfeld-Fresko, S., and Proia, R.L. (1989) Analysis of the glycosylation and phosphorylation of the lysosomal enzyme, β-hexosaminidase B, by site-directed mutagenesis. *J. Biol. Chem.*, **254**, 7692–7697.

Staudacher, E., Altmann, F., Wilson, I.B.H., and März, L. (1999) Fucose in N-glycans: from plant to man. *Biochim. Biophys. Acta*, **1472**, 216–236.

Tews, I., Perrakis, A., Oppenheim, A., Dauter, Z., Wilson, K., and Vorgias, C.E. (1996a) Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. *Nat. Struc. Biol.*, **3**, 638–648.

Tews, I., Vincentelli, R., and Vorgias, C.E. (1996b) N-Acetylglucosaminidase (chitobiase) from Serratia marcescens: gene sequence, and protein production and purification in Escherichia coli. *Gene*, **170**, 63–67.

Yazdanparast, R., Andrews, P.C., Smith, D.L., and Dixon, J.E. (1987) A new approach for detection and assignment of disulfide bonds in peptides. *J. Biol. Chem.*, **262**, 2507–2513S.