The Activator Protein for Glucosylceramide β-Glucosidase from Guinea Pig Liver

IMPROVED ISOLATION METHOD AND COMPLETE AMINO ACID SEQUENCE*

(Received for publication, May 2, 1988)

Akira Sano†, Norman S. Radin‡, Linda L. Johnson§, and George E. Tarr††

From the † Mental Health Research Institute and the ‡ Department of Biological Chemistry, the University of Michigan, Ann Arbor, Michigan 48109

β-Glucosidase activator (SAP-2) is a family of heat-stable, acidic glycoproteins which stimulate enzymatic hydrolysis of glucosylceramide. In this study, we improved the purification method and found that SAP-2 is highly heterogeneous. A hot water extract of frozen guinea pig liver was fractionated by ammonium sulfate sedimentation, then chromatographed with DEAE-Sephadex, Sephadex G-75, and concanavalin A-Sepharose. A fraction binding to concanavalin A-Sepharose was purified further with a C4 high performance liquid chromatography reverse phase column. This yielded several peaks, the main one of which was studied. The specificity of the purified SAP-2 was 35 units/mg (1 unit produces 50% stimulation of a basal glucosidase preparation). N-terminal amino acid sequence showed that this preparation is a mixture of polypeptides differing in the presence or absence of one or two of the end amino acids. The complete amino acid sequence of the 81 residues in SAP-2 has been determined. Comparison of the sequence of the guinea pig SAP-2 with the sequence of human sphingomyelinase activator revealed 58% homology and quite similar hydrophobicity profiles. Both proteins possess a highly hydrophilic region around Asn-22, which is glycosylated, and 6 cysteine residues, in oxidized form, located in the same positions. Comparison with the published nucleotide sequence for the precursor form of the human activator protein for sulfatide sulfatase (SAP-1) suggested that this activator also has a possibly glycosylated Asn and 6 Cys residues at similar positions, although the remainder of the molecule is somewhat different. Examination of another region of the precursor’s nucleotide sequence, assuming a few changes in the identifications, revealed the presence of the sphingomyelinase activator. It appears that two or more activators are derived from a single precursor protein. Marked homologies were seen also with a lung surfactant protein and a sulfated glycoprotein from Sertoli cells.

The glycosphingolipid, GlcCer, is catabolized by a β-glucosidase in a wide variety of organisms and cells to form ceramide and glucose. The enzyme activity in vitro is greatly stimulated by taurocholate, acidic lipids, or—in the presence of a small amount of acidic lipid—by an activator protein that has been called Factor P (1), HSF (2), SAP-2 (3), glucosidase (4), cohydrolyase sphingolipid I (5), AP (6), and sphingolipid activator protein A (7). For brevity here we call it SAP-2. It has been said to stimulate also sphingomyelinase and galactosylceramide β-galactosidase. The protein has been isolated from bovine and human spleen and from human brain (4, 5, 8, 9). It showed heterogeneity in PAGE, and concanavalin A chromatography yielded binding and non-binding forms of the protein (5, 6). The different members of the family were found to have similar stimulating activities, stability on PAGE with Stains All, and immunoreactivity with polyclonal antibodies prepared against the mixture (5). SAP-2 accumulates greatly in the spleen of patients with Gaucher disease (genetic defect in glucosidase activity), and its accumulation in liver could be induced by injecting entulified GlcCer into mice (10). SAP-2 apparently acts by combining with the enzyme and acidic lipid to form an activated complex, rather than by simply solubilizing the substrate (11). The SAP-2s from normal and Gaucher human spleen were found to differ slightly in molecular weight and electrophoretic properties, but all preparations seemed to work on enzymes from different sources.

An activator preparation acting on sphingomyelinase and, to some extent, on GlcCer β-glucosidase, has been isolated from a Gaucher spleen and its complete amino acid sequence has been described (12). In this paper, we describe an improved method for isolating several forms of glucosidase activators from a normal tissue, guinea pig liver, describe the primary structure of one form of the activator, and point to the similarities and differences in the structures of the two proteins. Previously published data on a precursor protein of a different activator are reinterpreted and some striking homologies with other proteins are noted.

EXPERIMENTAL PROCEDURES AND RESULTS*

DISCUSSION

Comparison of Hydrolase Activators—Comparison in Fig. 5 of the sequences of SAP-2 and of Aα, sphingomyelinase activator from human Gaucher spleen (12) shows that there is 58% homology and that the residues important for protein

---

*This study was supported by Grant NS 03192 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Recipient of the Senator Jacob Javits Neuroscience Investigator Award from the National Institutes of Health. To whom correspondence and reprint requests should be addressed.

The abbreviations used are: GlcCer, glucosylceramide or glucocerebroside; BSA, bovine serum albumin; Buffer A, 10 mM sodium phosphate, pH 7.0, 0.02% NaN3; HPLC, high performance liquid chromatography; PAGE, polyacrylamide gel electrophoresis; SAP 2, β-glucosidase-activating protein; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid; ConA, concanavalin A.

---

² Portions of this paper (including "Experimental Procedures," "Results," Tables I-III, and Figs. 1-4) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.
structure (Cys-5, -8, -36, -47, -72, and -78 and glycosylated Asn-22) are identical. The finally deduced sequence for the Aᵦ protein was based on peptides which failed to overlap in three different points; nevertheless, the alignment is in good agreement with ours. One point of difference between the two is in the C-terminal region which, in SAP-2, ends with an 81st residue and in Aᵦ ends in an 80th residue. The differences can be attributed to evolutionary differentiation and, possibly, also to the existence of more than one form of SAP-2 (7).

The hydropathy profiles of the two activator proteins, calculated by the method of Kyte and Doolittle (24), were surprisingly similar despite the 42% residue substitutions (Fig. 6A). The first 30 and the last 30 residues were very similar while a significant difference in the overall hydropathy characteristics appeared in the central region. Despite this difference, the central regions showed remarkable similarity in the 81st residue and in Ala ends in an 80th residue. The differences can be attributed to evolutionary differentiation and, possibly, also to the existence of more than one form of SAP-2 (7).

Another intriguing observation comes from reinterpretation of the nucleotide sequence for this precursor protein (25): if we omit adenylic acid 1019 from the data and then reencode the triplets starting with nucleotide 1018, we obtain the sequence for the human activator Aᵦ up to amino acid residue 67 (nucleotide 1111). If we further suggest that nucleotides 1112 and 1113 should be replaced by three nucleotides coding for Pro, the remaining 12 amino acid residues of activator Aᵦ are also in agreement with the nucleotide sequence. Thus, it seems highly likely that the large precursor of the activator protein, thus augmenting its hydrophilicity.

Analysis of the nucleotide sequence of the cDNA corresponding to another human activator protein (SAP-1) has been reported (25). This protein stimulates the activity of other sphingolipid hydrolases. The sequence codes for two different sphingolipid hydrolase activators. This conclusion is supported by the observations that human chromosome 10 codes for both activators (26, 27). If we further examine the precursor DNA sequence and assume guanylic acid 1192 is the beginning of a third activator protein (X), it is then possible to see an Asn-22 followed by Ser-Thr (again a presumed glycosylated residue) and 5 Cys residues in the same positions as in SAP-2 and Aᵦ. Only the 6th Cys residue seems to be missing. All four proteins have a Tyr-54. The hydropathy profiles for X and Aᵦ (not shown) are strikingly similar despite the low degree of homology (Fig. 6B). The former lacks the strong hydrophobic region around Asn-22, although it does exhibit the somewhat hydrophilic region in the center of the molecule.
Subsequent to this reinterpretation, we have learned (29) that the precursor protein for SAP-1 does indeed contain within it a very close copy of the SAP-2 protein, as well as two more very similar proteins (presumably other sphingolipid hydrolase activators).

**Structure-Function Correlation**—The basic amino acids in guinea pig SAP-2 are located in the first 27 N-terminal residues; in the human activator, A4, they are localized in the first 41 residues (Fig. 5). The acidic amino acids, on the other hand, are rather diffusely distributed. Thus, especially at the acidic pH characteristic of lysosomes, the glucosidase activators possess a distinct positively charged region. Both proteins also possess a hydrophobic region due to the first 19 residues (Figs. 3 and 6A). Thus, this region is a good candidate for the binding of an acidic lipid in the active ternary complex with the enzyme (11, 30).

The hydrophilic region of SAP-2 apparently is not involved in the binding to β-glucosidase since deglycosylation in this region did not destroy the stimulating activity (18).

SAP-1, unlike the glucosidase activator, has been shown to form a solubilizing complex with the substrates of the sphingolipid hydrolases that it activates and apparently does not interact positively with acidic lipids. These differences in mechanism of action are consistent with the above hypothesis since SAP-1 lacks a basic amino acid sequence possibly. Possibly the hydrophobic regions due to residues 3-18 and 24-29 are responsible for the substrate binding (Fig. 6B).

Dewji et al. (25) have pointed to the marked similarity in sequence pattern and homology between human SAP-1 and a region of a sulfated glycoprotein from rat Sertoli cells (31). The sequence homology between the two proteins was 76%. In view of the close similarity, we suggest that the Sertoli protein functions as an activator (or inhibitor) for some of the hydrolases in the sperms acrosome, a region needed for penetration of the egg cell zona pellucida. A more recent paper on the Sertoli protein (32) has shown that it possesses four domains, very much like the sphingolipid activator precursor protein (29).

In addition, it may be noted that a lipid-binding lung surfactant protein (33) also has the same pattern as the hydrolase activators, particularly the positions of the 6 Cys residues (oxidized), the Tyr following the second pair of Cys residues, and the 2 Pro residues near the third and fifth Cys groups. However, the lung protein lacks the glycosylation site, a feature which helps explain its extreme lipophilic nature. The overall homology between human A4 and surfactant protein amounts to 20 amino acids out of 80. More recent work on this protein has shown that it too occurs as part of a precursor protein with more than three similar domains, like the SAP proteins.

Acknowledgments—We are indebted to Dr. Fulvio Perini and Laura L. Looest of the Department of Pharmacology for assistance in sequence analysis. The Protein Identification Resource, Georgetown University Medical Center, kindly used their database and computer to find the marked similarity between SAP-2 and surfactant proteins, as well as additional intriguing relationships. Inez Mason contributed valuable laboratory assistance.

Note added in Proof—Furst et al. (34) have also recently noted the presence of SAP-1 and SAP-2, as well as a third protein ("component C") in the large precursor protein.

REFERENCE


35. J. A. Whitsett, personal communication.
EXPERIMENTAL PROCEDURES

Materials—Hamster guinea pig were decapitated and the livers were frozen rapidly on dry ice, then stored at −80°C before use. Commercially available frozen guinea pig livers from Dr. Fred Bohnsack (Chicago, IL) were also used. Liver from splotch was chosen as the SAP-2 source because the splotch protein appears to be higher than 1.5 by two successive isolation from the large spleen of large animals makes them difficult to homogenize. Guinea pig, rather than a large animal, was used as the liver source because of the relatively shorter time needed to remove liver. This reduces the danger of pentostemum changes.

DEAE-Sepharose and Sephadex G-75 were from Pharmacia Fine Chemicals, Uppsala, Sweden. Ammonium sulfate was purchased from Boehringer Mannheim. Other materials were from Sigma Chemical.

SAP-2 Assay—Samples were assayed for stimulatory activity toward a partially purified preparation of glucosidase by incubation with methylumbelliferyl glucoside in acetate buffer, pH 4.5, with Triton X-100 as previously described (14). Protein Determination and Amin Acid Analysis—The biochemical and method (15) was used for protein measurement with BSA as standard. Amino acid composition analysis was done with Pharmacia-LKB automated amino acid analyzers.

Isolation of SAP-2—All steps were carried out at room temperature except where indicated. In the run described here, thirty-four liters (219 g) were crashed in an aluminum foil with a wooden mallet and then about 5 liters of boiling water for 10 min. This was done to minimize the danger of artificial enzymatic degradation of SAP-2. The mixture was sonicated with a 20-kHz Turrckson ultrasonic bath (vibrating bath). Nodes at bottom, while still hot, then the homogenate was cooled in ice cold water and centrifuged 20 min at 5000 g. The resultant pellet was rehomogenized with 1.5 liters of water and the supernatant portion, after centrifugation as above, was pooled with the first extract.

Ammonium sulfate was added slowly, with stirring, to produce 45% saturation and, after standing an additional hour, the mixture was centrifuged 20 min as above. This step was repeated with the supernatant portion and 80% saturated ammonium sulfate and the resultant pellet was dissolved in 60 ml of Buffer A and dialyzed against the same buffer.

The sample was pumped at 33.6 ml/h into a column of DEAE-Sepharose (2.6 × 37 cm) previously equilibrated with Buffer A and followed by 75 ml of Buffer A and 1200 ml of the same buffer containing a linear gradient of 0-0.7 M NaCl. A yellow-brown protein material eluted before SAP-2, with overlapping. Fractions (15 ml) of glucosidase-stimulating activity, eluting at 0-4 M NaCl, were pooled and adjusted to 80% saturation with ammonium sulfate for 1 h. The fraction was centrifuged 15 min at 10,000 g and the pellet was dissolved in 8.2 ml of Buffer A containing 0.1 M NaCl.

This sample was now pumped at 10 ml/h through a Sephadex G-75 column (2.4 × 34 cm) packed and eluted with the same buffer. The step figure not shown produced a relatively broad peak for SAP-2, a feature noted before with human and bovine SAP-2 indicating of heterogeneity or molecular weight. The SAP-2 fractions, eluting between 91 and 218 ml, were pooled and concentrated by ammonium sulfate precipitation as before.

The pellet was dissolved in 1.5 ml of Buffer A, 0.1 M NaCl, then pumped at 10 ml/h into a column of ConA-Sepharose (1.3 × 4.5 cm) previously equilibrated with the same buffer and followed by 26 ml of the buffer and 60 ml of the buffer containing a linear gradient of 0-100 mM methyl mannoside. Three peaks with SAP-2 activity were seen, the first. Two passing through the column in the starting buffer and the third eluting with 10 mM methyl mannoside (Fig. 1). Using a shorter ConA column, we had previously found only two peaks with human SAP-2-D. The SAP-2 fractions which eluted after the gradient started were pooled, dialyzed against water, and lyophilized.

The dry SAP-2 sample was now dissolved in 0.1% TFA (100 mg) and injected into a 0.2 cm loop in an HPLC system. Separation was accomplished with a high-performance C4 reverse phase column (Alltech MicroBond 300 C4, 4.6 × 150 mm, 5 μm, particle, mounted in 200 nm effective length, solvent was a linear gradient of 5-60% acetonitrile in 0.1% TFA. This step produced several overlapping peaks of active material which almost coincided with the UV absorbance curve (Fig. 2). The last part of the main peak was used for further analysis.

Fractions were evaporated to dryness under vacuum at a centrifuge SpeedVac, Savant Instruments, Henrietta, NY) and dissolved in 0.05% methanol for analysis. Methanol has been used earlier for handling a hydrophilic protein. The main active peak fractions were pooled in a polypropylene centrifuge tube and lyophilized for other analyses. Peptide fragments were purified similarly. The SAP-2 activity was stable during C4 chromatography and eluting steps despite the low pH and high concentration of organic solvent.

Evaluation of the Purification Steps—As noted with previous purification methods, SAP-2 specific activity could not be determined with the enzyme stimulation assay in the initial extract because of the presence of inhibitory material, even after ammonium sulfate fractionation. Approximate assays indicated that more than 85% of the SAP-2 precipitated between 45 and 80% of saturation with ammonium sulfate. The inhibitory material eluted at the amino acid exchange column just after the SAP-2, suggesting it is a more acidic substance than SAP-2. Recent work has shown that much of the inhibition is due to the presence of ribonuclease in the extract (10).

The SAP-2 purified to the final stage (Table I) had a specific activity of 35 units/μg, similar to the one reported for bovine SAP-2 but somewhat lower than that of human spleen SAP-2 (21).

Table I Purification of glycoprotein from guinea pig liver (370 g)

![Table I Purification of glycoprotein from guinea pig liver (370 g)](image)

**SOS-PAGE** of our preparation revealed three bands of 6.5, 8.5, and 10 kDa, visible with silver staining (11b). In our previous study of the SAP-2 from human spleen (10), the ConA-binding material yielded only two bands. With native PAGE, the concentration of SAP-2 in guinea pig liver can be estimated from the data in Table I as being >18 μg/g. This value includes the protein portion that does not bind to concanavalin A. Previously published concentration for total SAP-2 isolated from normal human and bovine spleen, corrected for losses, are similar (45). However, the reported value for human brain, 14 mg/g, corrected for yield to 236 μg/g, is much higher.

Concentrations in mouse liver, measured by enzyme-linked immunosorbent assay (ELISA), were 3.9-4.5 μg/g.

The major novel features of our isolation method are the use of procedures to minimize pentosteum change, the use of ammonium sulfate as an initial concentration and purification step, and the use of an HPLC reverse phase column.

Deglycosylation—This was done as described previously (16) with protein N-glycosylase F in EDTA and o-phosphatidylcholine, incubated for 10 h. The incubation mixture was purified with the Cu column to remove the enzyme and salts. All of the UV-absorbing material was then pooled and used in the next step.

Reduction and Alkylation—SAP-2 was reduced with dithiothreitol in guinea pig liver (370 g) and alkylated with vinylpyridine or succinic anhydride (16). The reduced, alkylated SAP-2 was separated from the reaction with the Cu column, the solvent was removed by centrifugal vacuum evaporation and the residues were dissolved in 0.05% triethylamine, pooled, lyophilized.

CLEAVAGE METHODS—Cleavage of SAP-2 with CNBr was performed in 70% formic acid (14) or 70% TFA (21) for 18 h at room temperature in the dark. The reaction mixture was lyophilized; the residue was dissolved in 0.1% TFA, and the products were separated by HPLC as above.

Digestion of 500 μg of reduced, alkylated SAP-2 with 3 μg of trypsin was carried out in 0.1 ml of 0.1 M ammonium bicarbonate, pH 7.9, at 37°C for 2.5 h. The digested sample was added to pH -2 with TFA and then subjected to HPLC.

CNBr fragmentation of 140 μg of the large peptide derived from tryptic digestion (17) was performed in 70% TFA, after which the solution was lyophilized and the residue fractionated by HPLC as above. The large peptide from this step was further digested with staphylococcal V8 protein (110 μg) in 0.1 ml of 0.1 M ammonium acetate, pH 8.0, at 37°C for 18 h. The digest was diluted with TFA and applied to HPLC (Fig. 5A)

Peptide (114 μg) was also digested with 3 μg of trypsin-free chymotrypsin in 0.1 ml of 0.1 M ammonium bicarbonate, pH 7.9, at 37°C for 7 min. The sample was then diluted with TFA and subjected to HPLC (Fig. 5B).

**Activator Protein for β-Glucosidase**

**SUPPLEMENTARY MATERIAL TO THE ACTIVATOR PROTEIN FOR GLUCOSYLCERAMIDE β-GLUCOSIDASE FROM GUINEA PIG LIVER**

**Isolation Method and Complete Amino Acid Sequence**

Anna Soo, Norman B. Radin, Linda L. Johnson, and George E. Tair

**Fig. 2.** HPLC separation of SAP-2 with a C4 reverse phase column. Line A shows the absorbance at 250 nm. The fractions (1 ml) were assayed for enzyme activity (line B).
Sequence Determination—Automatic Edman degradation was carried out in a model 470A gas phase sequencer and 112A PTH analyzer (Applied Biosystems). Manual sequence analyses were performed using the partition method or the film method with Polybrene (18).

RESULTS

Amino Acid Sequence—The first few cycles of N-terminal analysis showed that SAP-2 consisted of three polypeptides that differed by the absence or presence of the first and second amino acids. The major peptide possessed the sequence Glu-Ser-Val-Thr-Cys, another peptide began with Ser-Val-Thr-Cys, and the third began with Val-Thr-Cys. The same sequence and mixture were observed when mixed forms of SAP-2 were assayed (when the ConA-Sepharose step was omitted). Human A2A activator from Gaucher spleen was also found to contain a similar mixture of shortened forms (21).

Sequencing of guinea pig liver SAP-2 revealed the presence of 81 amino acids, making the calculated Mw 8743. The peptides used to determine the sequence are indicated in Fig. 4. Except for position 80, these peptides overlap by at least three residues. Identifications of all residues were confirmed by compositional amino acid analysis (Table II) and duplicate sequence analyses of the peptides.

The data show the number of residues in each peptide fragment, analytically determined and (in parentheses) calculated from sequencing data. The four peptides in the right side of the table were all derived from the tryptic peptide shown in the Typl column.

The CNBr produced Gly from Metl, which overlapped Gly in the chromatogram. This peptide was contaminated with a peptide beginning with Ser-C1.

The amino acid compositional analysis of SAP-2 (Table III) agrees well with the results of sequencing, indicating that the sequence is complete. A low content of aromatic amino acids was noted. Trp and Phe were absent. The absence of Trp was established fluorimetrically as well as by sequencing.

TABLE II

<table>
<thead>
<tr>
<th>Residue</th>
<th>Amino Acid</th>
<th>Calculated Value</th>
<th>Analytical Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ala</td>
<td>4.1</td>
<td>4</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>Ser</td>
<td>1.2</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>3</td>
<td>Thr</td>
<td>6.1</td>
<td>6</td>
<td>0.0</td>
</tr>
<tr>
<td>4</td>
<td>Pro</td>
<td>2.1</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td>5</td>
<td>Gln</td>
<td>3.5</td>
<td>3</td>
<td>0.0</td>
</tr>
<tr>
<td>6</td>
<td>Gly</td>
<td>2.3</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>His</td>
<td>0.8</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>8</td>
<td>Met</td>
<td>2.1</td>
<td>Val</td>
<td>0.2</td>
</tr>
</tbody>
</table>

A Male equivalents, assuming Mw = 8743.

B Sum of free acid and amide acid.

C Average of three values.

The six Cys residues were shown to be in disulfide linkages. Alkylation of native and reduced SAP-2 with 2-mercaptoethanol, followed by amino acid analysis, showed the presence of six pyroglutamyl-Cys residues in the latter and none in the former. An attempt at alkylating native SAP-2 did not change the elution from the C4 column, but alkylation after reduction resulted in conversion to an earlier-eluting peak.

Study of the sugars in SAP-2 showed that it contained an N-linked glycosylated chain (19). The chain appears to be attached to Asn-22 since the sequence Asn-X-Thr/Ser pattern common to all N-glycosidically linked sequences (X is any amino acid except Pro) (22). This observation is also in accord with our finding that sequencing SAP-2 produced gag at position 22 after deglycosylation with N-glycosidase F. This occurs because the enzyme cleaves the amide linkage (23). Moreover N-terminal sequencing of native SAP-2, after cleavage with CNBr, showed a blank at position 22, evidently because glycosylated Asn is normally lost in the procedure.

Sequencing the C-terminal region of SAP-2 required two approaches. Digestion of peptide Typl line 1, Fig. 4, with chymotrypsin yielded many peptides (Fig. 3B), several of which were identified by compositional analysis as being part of the C-terminal region. However, the sequencing data of these peptides were reliable only up to Ser-60. Digestion of intact, unordered SAP-2 with CNBr in 70% TFA yielded four peptides were sequenced without separation. TFA was used because of its superior cleavage of Met-Thr/Thr linkages (29). The N-terminal sequence and the two expected sequences following Met-16 and Met-66 were readily identified. In addition, the C-terminal dipeptide, Ser-Gly, could be uniquely identified from this mixture. These data yield a calculated content of three Gly residues, in good agreement with the analytical data for intact SAP-2 (Table III).