cDNA encoding the glycosyl-phosphatidylinositol-specific phospholipase C of Trypanosoma brucei

(glycolipid anchor/variant surface glycoprotein)

Dale Hereld, Gerald W. Hart, and Paul T. Englund

Department of Biological Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205

Communicated by M. Daniel Lane, August 24, 1988 (received for review July 22, 1988)

ABSTRACT VSG lipase of Trypanosoma brucei specifically cleaves the glycosyl-phosphatidylinositol membrane anchor of the trypanosome variant surface glycoprotein (VSG), releasing this protein from the plasma membrane. It also cleaves similar membrane anchors on some mammalian proteins. VSG lipase may play a role in processes such as parasite differentiation or antigenic variation. We describe here the cloning and sequencing of a cDNA encoding VSG lipase from T. brucei.

The variant surface glycoprotein (VSG) of Trypanosoma brucei is anchored to the plasma membrane by a covalently attached glycosyl-phosphatidylinositol. Ferguson et al. have recently reported the complete structure of this glycolipid anchor (1). It contains dimyristoyl phosphatidylinositol glycosidically linked to nonacylated glucosamine. Attached to the glucosamine is an oligosaccharide composed of mannose and galactose, and a phosphodiester links ethanolamine to one of the mannosyl residues. Finally, an amide bond joins ethanolamine and the α-carboxyl of the protein's C-terminal amino acid residue.

Glycosyl-phosphatidylinositol anchors of similar structure are also found on cell-surface proteins in other protozoa and in higher eukaryotes including humans (for reviews, see refs. 2-4). These proteins include Thy-1, alkaline phosphatase, acetylcholinesterase, and decay accelerating factor of mammals and several surface antigens and membrane-bound enzymes in protozoa. A glycosyl-phosphatidylinositol related in structure to the membrane anchors has been postulated to mediate some of the actions of insulin (5).

The function of glycosyl-phosphatidylinositol membrane anchors is not yet clear, but in some cases they may facilitate specific and regulated release of proteins from cell surfaces by phospholipases or other enzymes that cleave the glycolipid. This possibility is especially appealing in the case of the trypanosome VSG; this protein is known to be lost from the cell when the parasite enters its tsetse fly vector (6, 7) and possibly also following endocytosis (4, 8) or during antigenic variation. VSGs of different trypanosome variants differ dramatically in amino acid sequence; therefore, if they were anchored by a hydrophobic peptide sequence it might be difficult to achieve specific release by proteolysis. However, the presence of a common glycolipid anchor on all VSGs would provide a uniform cleavage site for specific release by a phospholipase.

In 1983, Cardoso de Almeida and Turner (9) discovered a membrane-bound enzyme activity in T. brucei that cleaves the VSG glycolipid. When purified to homogeneity, this enzyme consists of a single polypeptide of about 37-40 kDa (10-12). It is a Ca2+-independent phospholipase C which efficiently cleaves VSG membrane anchors, but it has little or no activity against free phosphatidylinositol. This enzyme was the first example of a class of enzymes called glycosyl-phosphatidylinositol-specific phospholipases. More recently, there have been reports of a similar enzyme in rat hepatocytes that may be involved in insulin action (13) and a glycosyl-phosphatidylinositol-specific phospholipase D, of unknown function, from mammalian serum (14, 15).

The trypanosome enzyme, which we designate VSG lipase, is apparently quiescent on living bloodstream parasites, yet upon cell lysis under nonnematuring conditions it cleaves all of the VSG from the membranes within several minutes. Its dormancy in living cells could be due to regulation of the enzyme activity; the enzyme could be specifically activated for release of VSG during appropriate stages of the trypanosome life cycle. As part of a study of the biological role and mode of regulation of VSG lipase, we have cloned and sequenced its cDNA.* It encodes a polypeptide of 40,760 Da, a value consistent with the protein's electrophoretic behavior. Despite the fact that VSG lipase appears to be a membrane protein, the deduced sequence does not include strongly hydrophobic domains or an N-terminal signal sequence. There are no significant similarities between the VSG lipase sequence and those of other known proteins including several recently reported mammalian phosphatidylinositol-specific phospholipase C.

MATERIALS AND METHODS

Purification of VSG Lipase. VSG lipase was purified from T. brucei (IL.Tat 1.3) by two different methods. To generate rabbit antiserum and the mouse monoclonal antibody, it was purified by phenyl-Sepharose and carboxymethyl-Sephadex chromatography as described (method 1; see ref. 11).

For amino acid sequencing, VSG lipase was affinity purified by a procedure similar to that of Bulow and Overath (method 2; see ref. 10). A VSG lipase-specific murine monoclonal antibody, 2A6-6, which binds the native active enzyme, was generated by standard methods and a screening assay like that used by Bulow and Overath (10). Antibody was adsorbed from hybridoma culture supernatants onto protein A-Sepharose (Pharmacia) and covalently coupled to the resin by the method of Schneider et al. (16). An n-octyl glucoside extract of 5.4 × 1010 trypanosomes (40 ml; ref. 11) was passed through a 0.1-ml column of the resin overnight at 4°C. The column was then washed at 4°C at 0.5 ml/min with 2.5 ml of each of the following solutions in succession: (i) TEN (50 mM Tris-HCl/5 mM EDTA/150 mM NaCl, pH 7.5) containing 1% Nonidet P-40 (NP-40) and 0.5 M NaCl; (ii) TEN containing 1% NP-40; (iii) TEN containing 0.1% NP-40. Finally, the enzyme was eluted from the column with 2.5 ml of 50 mM NaP/0.1% NP-40, pH 12.0, at the same flow rate.

Abbreviations: NP-40, Nonidet P-40; VSG, variant surface glycoprotein.

*The sequence reported in this paper is being deposited in the EMBL/GenBank data base (accession no. J04124).
Fractions (0.25 ml) were immediately neutralized with 0.1 ml of 1 M Tris-HCl (pH 7.5); those containing VSG lipase activity (as assayed as described in ref. 11) were pooled. Greater than 80% of the activity loaded onto the column was recovered in ~1.5 ml of the eluting buffer, and NaDODSO4/PAGE analysis of this material revealed a homogeneous 36-kDa protein (see Fig. 3A). Its specific activity, measured as described in ref. 11, was ~10^4 units/mg. This preparation is designated affinity-purified VSG lipase.

**Production of Antiserum.** Rabbit antiserum was obtained by immunization with VSG lipase (15 μg, prepared by method 1) excised from a NaDODSO4/polyacrylamide gel. Gel pieces were equilibrated in 0.15 M NaCl, sonicated, and injected as an emulsion with Freund’s complete adjuvant by standard procedures. VSG lipase-specific antiserum was obtained after the rabbit was given two comparable booster injections with Freund’s incomplete adjuvant.

**Screening of cDNA Library.** About 10^8 clones of an amplified Agt11 cDNA library prepared from RNA of *T. brucei rhodesiense* (WRA/Tat 1.1) bloodstream trypanosomes (generously provided by J. Donelson, University of Iowa) were screened with VSG lipase-specific antiserum (diluted 1:2000) (17). Alkaline phosphatase conjugated to goat anti-rabbit immunoglobulin was used to detect bound antibodies. Several clones, each containing a 0.95-kilobase (kb) insert, were obtained. One, designated DH1, was used as a hybridization probe (18) to screen a second *T. brucei rhodesiense* (WRA/Tat 1.1) Agt11 cDNA library (also provided by J. Donelson) by standard methods (19). Of 10^6 recombinants screened, 6 cross-hybridizing clones were detected. A 1.4-kb cDNA, designated DH6, was used for DNA sequencing.

**cDNA Sequencing.** cDNA inserts were subcloned into the plasmid vector Bluescript (Stratagene, San Diego, CA) by established methods (19). Plasmid constructs were amplified in *Escherichia coli* strain DH5a (Bethesda Research Laboratories) and sequenced with Sequenase (United States Biochemicals, Cleveland, OH) following the manufacturer’s protocol. To completely sequence both DNA strands, deletions were created from each end of the 1.4-kb cDNA using exonuclease III and mung bean nuclease as directed by the supplier (Stratagene).

**RESULTS**

**Isolation and Characterization of cDNAs.** We first isolated a phage clone from an Agt11 cDNA library by immunoscreening with a VSG lipase-specific antiserum. This clone contained a 0.95-kb cDNA insert designated DH1. The probing of western blots of infected cell extracts, using the same antiserum, indicates that this clone encodes a 145-kDa immunoreactive isopropyl thiogalactoside-inducible protein (data not shown).

We then obtained a 1444-base-pair (bp) cDNA, designated DH6, by screening a second library with radiolabeled DH1 sequence as a hybridization probe. Based on sequencing, the 3′ end of DH6 overlaps ~800 bp of the DH1 sequence and contributes an additional 600 bp of sequence at the 5′ end. Both strands of DH6 were completely sequenced (Fig. 1). The sequence of the “sense” strand is presented in Fig. 2.

DH6 contains an open reading frame encoding a polypeptide of 358 amino acid residues, beginning with a methionine codon at nucleotide 240. This polypeptide sequence is also shown in Fig. 2. To prove that this sequence is that of VSG lipase, we determined amino acid sequences of tryptic peptides derived from affinity-purified enzyme (prepared by method 2).

**Amino Acid Sequencing of VSG Lipase Peptides.** Initial attempts to obtain the VSG lipase N-terminal amino acid sequence were unsuccessful, possibly because the protein has a blocked N terminus. Therefore, we digested affinity-purified enzyme with trypsin and separated the resulting tryptic peptides by reverse-phase HPLC (Fig. 3B). The amino acid sequences of three peptides, designated T1, T2, and T3, are shown in Table 1. All three of these peptides are contained in the amino acid sequence predicted by the cDNA (the peptide sequences are underlined in Fig. 2). Therefore, the DH6 cDNA must encode VSG lipase.

**Properties of the cDNA Sequence.** Translation almost certainly begins at the methionine codon at nucleotide 240 because it is the only initiation codon present between in-frame termination codons at nucleotide 258 and 600 (Fig. 2). The sequence that encodes peptide T3 (beginning at nucleotide 258) is comprised of a 239-bp 5′ untranslated sequence, 1077 bp encoding VSG lipase, and a 128-bp 3′ untranslated sequence. An additional ~150 bp of 3′ untranslated sequence was obtained from cDNA DH1 (data not shown).

DH6 represents most of the 5′ end of the VSG lipase mRNA because the sequence at its extreme 5′ end closely resembles part of the miniexon sequence, a 39-mer which is present at the 5′ end of all known mRNAs of trypanosomes and related organisms (for review, see ref. 22). The 13-nucleotide miniexon-like sequence of DH6 is identical to the 3′ end of the *T. brucei brucei* miniexon sequence in all but one position (20). The substitution of a thymine for an adenine at nucleotide 5 (Fig. 2) could imply that the *T. brucei rhodesiense* miniexon differs from that of *T. brucei brucei*; however, all trypanosomatid miniexons that have been sequenced have adenine at this position (23). Alternatively, it could reflect a reverse transcriptase error during cDNA synthesis or the use of a miniexon gene that differs from the majority of the roughly 200 miniexon genes present in the parasite genome (22).

It is possible that neither DH1 nor DH6 cDNAs represent the complete 3′ end of VSG lipase mRNA. Since neither sequence has more than 8 adenines at its 3′ end, it is uncertain whether these are encoded in the gene or are added posttranscriptionally. When available, the sequence of the VSG lipase gene may clarify this point.

**Analysis of the Deduced Amino Acid Sequence.** The predicted amino sequence of VSG lipase (358 residues) given in Fig. 2 has a calculated molecular mass of 40,760 Da, in agreement with reports of 37–40 kDa based on NaDODSO4/PAGE (10–12). Since VSG lipase has been assumed to be a membrane-associated protein (see Discussion), we performed a hydrophathy analysis of the amino acid sequence (24) to identify domains that might interact with the lipid bilayer. The results of this analysis (Fig. 4) are discussed below.

**DISCUSSION**

Clone DH6 encodes the entire VSG lipase sequence. The original clone corresponding to this sequence (DH1) was isolated by immunoscreening with a rabbit antiserum. Although this clone could have encoded a protein that contaminated or cross-reacted with VSG lipase, protein sequencing...

Fig. 2. cDNA and amino acid sequences of VSG lipase. The sequence of the sense strand of DH6 and the deduced amino acid sequence are shown. Asterisks overline nucleotides that match the minexon sequence determined by Dorfman and Donelson (20). Open circle indicates stop codon. Residues confirmed by amino acid sequencing are underlined. Nucleotide numbering is at left and amino acid residue numbering is at right. All sequences in the coding region were confirmed on both strands except for the nucleotide at position 462, which was ambiguous on one of the strands.

Fig. 3. Tryptic peptides of VSG lipase. Affinity-purified VSG lipase (60 μg in 1.45 ml) was concentrated to 1.0 ml and adjusted to pH 8.5 with NaOH. NaDODSO₄ was added to a final concentration of 0.1% and the mixture was incubated for 2 min at 100°C. The protein was treated sequentially with dithiothreitol (5 μM in 5 μl, 75 min, 3°C), iodoacetamide (30 μM in 10 μl, 135 min, 25°C), and with dithiothreitol again (5 μM in 5 μl, 30 min, 25°C). The protein was then precipitated with chloroform and methanol (21). (A) Coomassie-stained NaDODSO₄/polyacrylamide gel of the alkylated affinity-purified VSG lipase (2 μg). (B) HPLC of alkylated VSG lipase (≈58 μg) digested with L-1-tosylamido-2-phenylethyl chloromethyl ketone-treated trypsin (2 μg) in 0.2 ml of 50 mM NH₄HCO₃ (pH 8) at 37°C for 18 hr. Trifluoroacetic acid (0.01 vol) was added and the reaction products were fractionated by reverse-phase HPLC using a large-pore C₁₈ column (Alltech, no. 71079). After a 5-min wash with 0.1% trifluoroacetic acid, tryptic peptides were eluted with a linear gradient of acetonitrile (0–60%) in 0.1% trifluoroacetic acid over 60 min (flow rate, 1 ml/min). Background absorbances (230 nm), present in a mock digest from which VSG lipase was omitted, were automatically subtracted from the profile shown. Peptides T1, T2, and T3 were analyzed for amino acid sequence (see Table 1).
Table 1. Amino acid sequences of VSG lipase tryptic peptides

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence determined</th>
<th>Yield, % of DH6</th>
<th>Corresponding nucleotides</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>MYVTQAGTPR</td>
<td>29</td>
<td>987–1019</td>
</tr>
<tr>
<td>T2</td>
<td>DLEDVXIGVPPX</td>
<td>12</td>
<td>948–986</td>
</tr>
<tr>
<td>T3</td>
<td>WSPQWSMSDTR</td>
<td>ND</td>
<td>238–290</td>
</tr>
</tbody>
</table>

HPLC fractions (0.5 ml) containing peptides T1, T2, and T3 were lyophilized and analyzed with a gas-phase automated amino acid sequencer (Applied Biosystems, Foster City, CA). Amino acids are denoted by the single-letter code. Yield refers to the derivatized amino acid detected in the first cycle expressed as a percentage of the theoretical yield of 1.6 nmol (58 μg) of VSG lipase used for sequencing. Losses could be due to incomplete trypsinization or could have occurred during HPLC or concentration. ND, not determined.

The molecular mass of the homogeneous VSG lipase polypeptide is 37–40 kDa (10–12). This value is similar to that deduced from the cDNA sequence (40,760 Da), indicating no posttranslational processing events grossly alter the protein’s apparent molecular mass. Although we cannot rule out the possibility that cleavage of the peptide was compensated by addition of some moiety, experiments with lactics (9) have failed to detect glycosylation at the four potential N-linked glycosylation sites (residues 82, 101, 315, and 533).

VSG lipase contains no significant sequence similarities with other proteins presently listed in the protein sequence data base. In particular, it has no similarities with the sequences of four distinct mammalian phosphorylidyinositol-specific phospholipase C isozymes (25–29).

Although the intracellular localization of VSG lipase is not yet known, several physical properties suggest that this enzyme is a membrane-associated protein. In cell extracts, the activity is present exclusively in the particulate fraction and has been solubilized only by detergents (8–12). The homogeneous isolated enzyme associates with liposomes (10), and the activity partitions into the hydrophobic phase during Triton X-114 phase separation (12).

Despite VSG lipase’s apparent association with membranes, its deduced amino acid sequence does not resemble that of a typical membrane protein. In contrast to many eukaryotic membrane proteins, it does not have a cleavable N-terminal signal sequence. As shown in Fig. 2, only five amino acids separate the N-terminal methionine residue and the first residue of T3, a peptide isolated from the mature protein. Since T3 was generated by tryptic cleavage, the neighboring Lys-6 is probably also present in the protein. Therefore, at most only five residues could have been cleaved from the N terminus of the primary translation product to form the mature protein.

Not all membrane proteins have N-terminal signal sequences. For example, some mammalian plasma membrane proteins synthesized in the endoplasmic reticulum, such as asialoglycoprotein receptor, transferrin receptor, and HLA-DR invariant chain have uncleaved internal signal sequences (reviewed by Wickner and Lodish, ref. 30). These proteins are synthesized with their C-terminal domains on the luminal side of the endoplasmic reticulum membrane and N-terminal domains on the cytoplasmic side; these domains are joined by a transmembrane hydrophobic sequence. If VSG lipase was processed in this way, its C-terminal domain would be topologically on the same side of the membrane as the VSG substrate. However, the Kyte–Doolittle plot (Fig. 4) does not offer support for this model. There are no internal hydrophobic sequences that are strong candidates for transmembrane domains. One of the more prominent hydrophobic sequences is between residues 57 and 75. However, this sequence contains an arginine and several serines and threonines, and therefore its hydropathic index is less than that found for most transmembrane domains (24). An alternative possibility is that VSG lipase has multiple membrane-spanning amphipathic α-helices, but analysis of the sequence revealed no obvious candidates for such a structure.

It is possible that VSG lipase is held to the membrane by a covalently attached lipid. However it lacks the N-terminal signal sequence and its C-terminal hydrophobic tail is present in the precursors of proteins anchored by glycosyl-phosphatidylinositol (4). Alternatively, VSG lipase might be modified by a fatty acid. If it were anchored by a myristoyl group attached to its N terminus, like p60src, or by a palmitoyl linked to an internal cysteine residue, like p21ras, it is unlikely that it would localize on the same side of the membrane as its VSG substrate (31).

Further investigation will be necessary to determine the mode of interaction of the VSG lipase polypeptide with the membrane. These studies should contribute to an understanding of the biological function of this enzyme.

We thank John Donelson for providing cDNA libraries, James Ntamb and Chelly Coutu for help with DNA cloning and sequencing, Don Kaiser for assistance with monoclonal antibody production, P. Shih, B. Tschudin, and Paul B. Rubenstein for advice with the peptide sequencing, and Russell Doolittle for help with computer analysis of the protein sequence. We also thank Tamara Doering, Wayne Masterson, Jessica Krakow, Nancy Shaper, and Joel Shaper for valuable discussions. We appreciate comments on the manuscript by John Donelson, Mervyn Turner, Kojo Mensa-Wilmot, Peter Devesot, and David Sabatini. Viu Klein and Shirley Skiles provided critical support and assistance. This research was supported by grants from the National Institutes of Health (AI21334) and from the MacArthur Foundation. D.H. was supported by Medical Scientist Training Program Grant T32GM07309.
