MOLECULAR AND CELLULAR BIOLOGY, Dec. 1990, p. 6785–6790
0270-7306/90/126785-06$02.00/0

Vol. 10, No. 12

Structural Isoforms of a Membrane Transport Protein from *Leishmania enriettii*

SEAN P. STACK,[1] DAVID A. STEIN,[1]† AND SCOTT M. LANDFEAR[1,2]*

Department of Microbiology and Immunology[1] and Vollum Institute for Advanced Biomedical Research,[2] The Oregon Health Sciences University, Portland, Oregon 97201

Received 20 July 1990/Accepted 6 September 1990

A membrane transport protein of the glucose transporter superfamily from *Leishmania enriettii* is encoded by a family of tandemly repeated genes. The first gene in this tandem repeat codes for a structural isoform that contains a unique amino-terminal hydrophilic domain, probably located in the cytoplasm; the remainder of the protein is identical to the polypeptide encoded by the internal genes in the tandem repeat. The unique isoform is represented by a distinct stable RNA.

The glucose transporter superfamily encompasses multiple structurally related membrane transport proteins in organisms as diverse as *Escherichia coli* and humans (1). All of these proteins possess significant amino acid sequence similarity to one another, contain 12 putative transmembrane α helices, lack cleavable signal sequences, and have cytoplasmically oriented hydrophilic domains at both their amino and carboxy termini (1, 4, 8). Various members of this superfamily transport different carbohydrate ligands such as glucose, arabinose, xylose, or galactose. In mammals, multiple structurally distinct isoforms of the glucose transporter are encoded by different unlinked genes, and these isoforms have different tissue distributions and unique kinetic properties (7).

Trypanosomatid protozoans of the genus *Leishmania* live in two very different environments during their life cycle (19). Promastigotes are extracellular, flagellated organisms that colonize the alimentary tract of the sandfly vector, while amastigotes are nonmotile intracellular parasites that live in the lysosomal vesicles of the vertebrate host macrophages. Previously, we have cloned a gene, designated Pro-1, that is expressed primarily in the insect or promastigote stage of the parasite life cycle and which encodes a member of the glucose transporter superfamily (3); the substrate specificity of this transporter is currently unknown. This gene is arranged in a single tandem repeat containing approximately seven to nine copies of the repeat unit (18). Here, we present sequence from the first (5′-proximal) unit of the tandem repeat. This first copy encodes a unique isoform of the transporter that contains a distinct hydrophilic amino-terminal domain; however, the remainder of the protein-coding region of this first copy is identical to a previously sequenced (3) internal unit of the repeat. This first copy is transcribed to form a distinct stable mRNA and does not appear to be a pseudogene. These data suggest that at least two isoforms of this transporter are expressed in *Leishmania* parasites; these isoforms might have distinct functional properties or subcellular locations.

Isolation of the 5′-flanking sequence from the Pro-1 tandem repeat. In a previous report (18), we identified a 2.75-kb *Pst*I fragment of genomic DNA that included the 5′ end of the Pro-1 tandem repeat. We have cloned this 5′-flanking sequence by screening a size-selected *Pst*I genomic library (18) with a probe specific for the 5′ end of the Pro-1 gene (probe A in reference 18). Figure 1 shows a map of this cloned fragment, designated p2.75, and its position within the Pro-1 tandem repeat. As expected for a 5′-flanking fragment, this insert contains sequence encoding the Pro-1 transporter protein at its right-hand terminus and sequence unrelated to the Pro-1 tandem repeat at its left-hand terminus (see Fig. 3).

To demonstrate that the p2.75 fragment was linked to the Pro-1 tandem repeat, we first digested genomic DNA to completion with *Sal*I, an enzyme that does not cut within the Pro-1 tandem repeat but has a single site within the p2.75 insert (Fig. 1). We then partially digested this *Sal*I-cut genomic DNA with *Cla*I, an enzyme that cuts once within each unit of the Pro-1 tandem repeat (Fig. 1) and, upon partial digestion, produces a ladder of Pro-1-containing fragments (18) separated by 3.6 kb, the size of the repeat unit. If the p2.75 fragment is at the 5′ end of the tandem repeat and is not an unlinked orphon copy, then this *Cla*I partial-*Sal*I complete digest should produce a 2.3-kb terminal fragment (Fig. 1A) plus a ladder of bands increasing in size by 3.6-kb units above this terminal fragment, resulting from the partial digestion with *Cla*I. To detect this pattern, we probed a Southern blot of these genomic double digests (Fig. 2) with a sequence that is specific for the p2.75 insert: the *Pst*I-*Hin*dIII fragment of p2.75 designated S1 probe in Fig. 1. This Southern blot reveals the predicted ladder of bands originating at 2.3 kb and increasing in size by units of approximately 3.6 kb (Fig. 2, lane 3; measured band sizes: a, 2.4 kb; b, 6.2 kb; c, 10 kb; d, 15 kb). We conclude that the p2.75 fragment does represent the 5′-terminal *Pst*I fragment of the Pro-1 tandem repeat.

Sequence of the 5′-terminal fragment reveals a new isoform of the Pro-1 transporter. Figure 3 shows the DNA sequence of the p2.75 insert and the deduced amino acid sequence of the protein-coding region. Nucleotides 1882 to 2796 (corresponding to the white boxed region in Fig. 1) are identical, with the exception of a single translationally silent nucleotide change (number 2652, T-to-C transition; Fig. 3), to the corresponding region of the previously sequenced (3) Pro-1 gene, which represents an internal unit of the repeat. However, the sequence from nucleotides 1 to 1881 is completely different from that reported for the internal copy of Pro-1. Remarkably, this point of divergence between the p2.75 and Pro-1 sequences (between nucleotides 1881 and 1882) occurs

* Corresponding author.
† Present address: Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331.

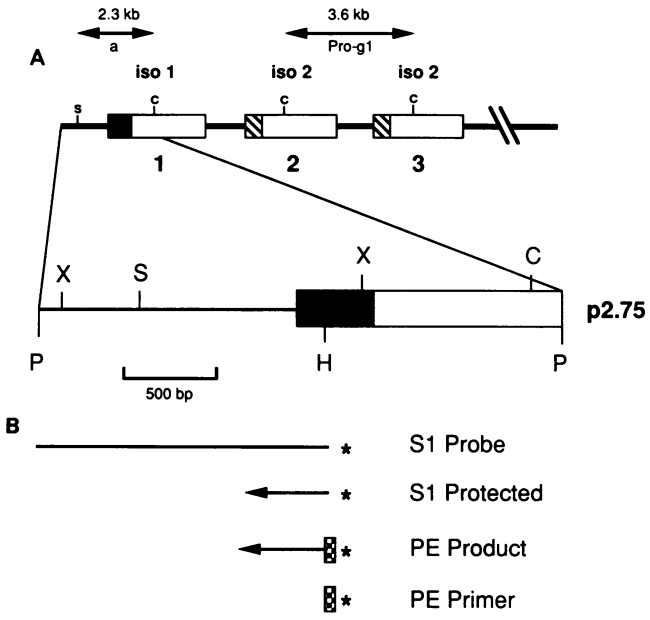

FIG. 1. (A) Chromosomal arrangement of Pro-1 transporter genes. The first three copies in the tandem repeat are indicated with rectangles designating the protein-coding regions of each unit. The shaded box represents the unique amino-terminal domain of the first copy (iso 1 for isoform 1), and the cross-hatched boxes represent the amino-terminal domains of the internal copies of the gene (iso 2 for isoform 2). The double-headed arrows indicate the distances, in kilobases, between restriction sites. The arrow marked a is the restriction fragment a referred to in Fig. 2; the arrow marked Pro-g1 indicates the previously reported Pro-g1 genomic clone (3). The expanded region below depicts the p2.75 subclone that is derived from the first copy of the gene. Restriction sites: P, *Pst*I; X, *Xho*I; S, *Sal*I; H, *Hin*dIII; C, *Cla*I. (B) The restriction fragment used as a probe for S1 mapping (S1 probe) and the synthetic oligonucleotide used for the primer extension experiment (PE primer); the star designates the 5′ end label. Both fragments are labeled at the unique *Hin*dIII site. The arrows indicate the major S1-protected fragment (S1 Protected) and primer extension product (PE Product).

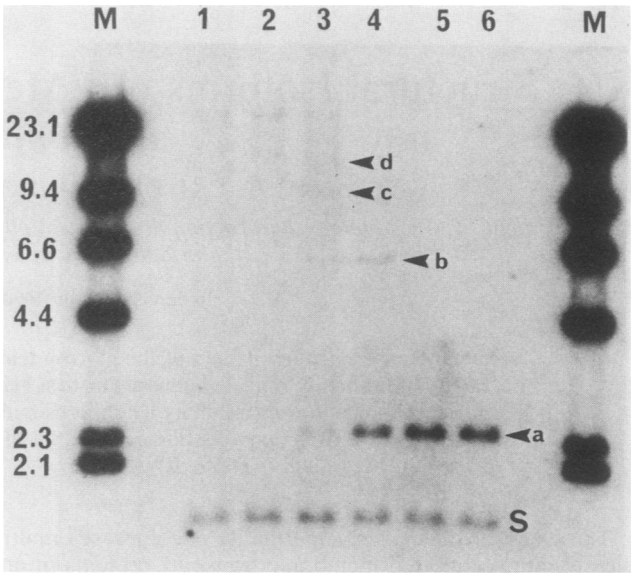

FIG. 2. Linkage of p2.75 insert to the Pro-1 tandem repeat in *Leishmania* genomic DNA, demonstrated by Southern blot hybridization. Genomic DNA (2 μg per lane) was first digested to completion with *Sal*I and then digested for 1 h at 37°C with increasing amounts of *Cla*I: 0.1 U (lane 1), 0.2 U (lane 2), 0.5 U (lane 3), 1 U (lane 4), 2 U (lane 5), and 12 U (lane 6). The digests were then separated on a 0.7% agarose gel, blotted onto a nylon filter, and hybridized with a radiolabeled probe that is specific for the p2.75 insert (a uniformly labeled [6] *Hin*dIII-*Pst*I restriction fragment designated S1 probe in Fig. 1). The band marked a is the 2.3-kb *Sal*I-*Cla*I fragment indicated by the double-headed arrow in Fig. 1. Bands marked b, c, and d are the fragments generated by partial *Cla*I digestion that are separated from each other by approximately 3.6 kb in size. The band marked S is a *Sal*I fragment of genomic DNA that overlaps with the 5′ end of the p2.75 insert and hence hybridizes with the probe. Lanes M are *Hin*dIII digests of λ DNA; numbers at the left indicate the sizes, in kilobase pairs, of these molecular weight markers.

within the deduced protein-coding regions of both the p2.75 insert and the Pro-1 sequence. The p2.75 sequence codes for 132 amino acids upstream of this point of divergence, beginning with a methionine at position 1486 (Fig. 3). By contrast, the previously determined Pro-1 sequence contains only 48 amino acids between this point of divergence and its deduced initiator methionine (Fig. 3, italicized amino acids). Consequently, the first copy of the gene, represented by the p2.75 insert, contains an amino-terminal 132 amino acids that is completely different from the first 48 amino acids of the previously sequenced internal copy of the transporter gene (see stippled and cross-hatched boxes in Fig. 1, representing the two different amino-terminal sequences). These results imply that the first copy of the tandem repeat encodes a different structural isoform of the transporter, designated isoform 1, compared with the internal copies of the gene, designated isoform 2 (Fig. 1).

It is possible that the internal copies of the tandem repeat code for polypeptides that also differ somewhat from each other, and hence there may be more than two isoforms represented by this gene family. However, Southern blots of genomic DNA digested with several restriction enzymes

FIG. 3. Sequence of the p2.75 insert derived from the first copy of the Pro-1 gene in the tandem repeat. Numbers at the right indicate nucleotide (top row) and amino acid (bottom row) positions. The deduced amino acid sequence, representing isoform 1, is shown under the DNA sequence. The boxed amino acids represent the predicted transmembrane α helices, as determined by the algorithm of Eisenberg et al. (5); each helix is numbered beneath the box. The initiating methionine is the only in-frame methionine between the beginning of the mRNA and the first transmembrane helix, and the five bases preceding this methionine fit the eucaryotic translation initiation consensus sequence CCACC (9). The 20 bases that are underlined (nucleotides 1622 to 1641) are complementary to the oligonucleotide used in the primer extension experiment (Fig. 1 and 5). The two AG sequences underlined (nucleotides 1211 to 1212 and 1221 to 1222) are potential splice acceptor sequences that are near the 5′ end of the p2.75 exon, as determined by S1 mapping (Fig. 5), and could serve as acceptor sites for the 39-nucleotide spliced leader (12). The amino acids in italics are from the previously determined (3) sequence of the Pro-1 polypeptide; hence, these represent the amino-terminal domain of isoform 2 (cross-hatched boxes in Fig. 1). The remainder of the deduced isoform 2 sequence is identical to that of isoform 1, as shown here. A single translationally silent nucleotide difference between the two sequences also occurs at nucleotide 2652 of p2.75, indicated by the italicized T above this sequence. Sequencing methods have been reported previously (3).

```
AGGTTCTGGCAGTCGGCGTTAACTGCAACCCCATGGCCGAAGCCTCCGCCGTGTTGGCCCATCTGCACACCCTTACCACAATGCCGCTCATCGTCTACACAAACTCG    214
GGCGAGTCCTACGACCCAGTCACTATGACGTGGCGCCCTATATCGGTGAGCGATGGCGCTACTCTCTGTTTGTCAGCATTAGCGCCGGAGTGGGCTTCCCACGGTGC    321
TCGGCTTATTGGCGGGTGCTGCCGCACCGGACCATCTGACGTTGCCGGGGCCGCCGCAGCGCTCAGTAATGCCGGCTTTATCGTGTAGCGGTCCAACTCCCAGAAAG'   428
AGGTGGCTGGCCAACAAGAAAATAAGAGGGTGGTCAGAGGGCATAGAAAGTCGGCGTTCCGCGCCTCACCACGCACGGTGAGAACCGACCTTTCCTTCCCTTCGCAG    535
CAAAGCGAGACGGGTTTATATATTCAAGTGCAACGTGCTTGCGCGTCGTCGACGATTACCGCTCTTTTTTCGCCTTTCATGATTGGCCAGCGCACCCTTCCTGAGCC    642
CCGCGTCACTTAAGTCGAGATTTCTTATCCTTCTTCATCCGACTCAGAGGTCACCCATGAGGGTGTGCCTGAGTGTTGAACACGCGCAGGATAGTTCTGCTGAACAC    749
GAGAAGAAACCAAATTCGGAGCCCATCACGTTTTTTTGCTTCCTTTAGCGTCTTCGGTGGGGTCCTCAGCGTCTCTTTGCCCTCCTCCCCTTTCTTTCGGGAACAAC    856
ACACACACACACACACACACACACATGTGTAAAGACGCAAGGCGAGTTGGGTTGCCGACTTGGTTGCCGGCTACCCCACTCCGACTATTACATCCCTTCCCCTTCCGTG    963
CACGGAGGAAGGGTCGAAAGTGGGGAGCGGAGATGAACCACAGCCTCCTTTCTCATCACTCGACCGCTAACACTATCATTCTGAGTGACTAATGCACGCGCGTGCAT    1070
CGGCGTCTTCCATCCGAAGCATTCCATGCGCACCATCACATCTCAGTAAGGAAGTTGCCGTTCTTTTTCGAAATCCAGCTCAGTTCCCTCTCCGCCTTTCTCGCTCC    1177
TGGGCCCACCTCCTATGGATACATCACCGCATCAGCGGGCATCAGGTCACCACCTCCCGCTTCTGCTATGATACCTCGCCCATACTCGACATAAACGTCAAAGGGCG    1284
CCTAAACCATTAATTCCCCCGCACACGCTAATACACACTAACACTAACAGGCGAGCCTAATCGACGTAAAGAACGAAAACGAACCACATCCAAATCTCACATAAATC    1391

AACAAAATCACCGCCACTTCCTGTCCAGAAAGACAGCTCCATCCCTCCATCATTCTAAAGCCTAACATCGTCTTTTCCTGCGTATGCAAGCAGC ATG AGC TAC    1494
                                                                                                  Met Ser Tyr      3

TAC CCC CCC AAA AGC CAG GAT CAG GGG CAG CTC CCA CTG AAA ACG TTC AGC TCA CCG CCT CGG CGA ACA GGA ACC ACG TCC    1575
Tyr Pro Pro Lys Ser Gln Asp Gln Gly Gln Leu Pro Leu Lys Thr Phe Ser Ser Pro Pro Arg Arg Thr Gly Thr Thr Ser      30

CAC GCA GCC CAT AAC GAC TGC GTC GCG GAA AGC GAA ACT TTG CCA ACA ACC CCA CCG CTA CCA AGC TTC CTC AGG GGC AAC    1656
His Ala Ala His Asn Asp Cys Val Ala Glu Ser Glu Thr Leu Pro Thr Thr Pro Pro Leu Pro Ser Phe Leu Arg Gly Asn      57

GAC GTG CAG TTG CCT AAG ACA GCT TCT GTT GCT CAC TCC TTA ACG ACA TCG CCG CCA TCA GTC AAC AAC CTT TCG CCT GGC    1737
Asp Val Gln Leu Pro Lys Thr Ala Ser Val Ala His Ser Leu Thr Thr Ser Pro Pro Ser Val Asn Asn Leu Ser Pro Gly      84

GCG GGC CCG CAT ACT CAC CAC AGA ATT GCC AAT CCG ATC AAC CCT GCA AGC ACC GAA GAC GAT ACC ACC TTT TCC ACC ACA    1818
Ala Gly Pro His Thr His His Arg Ile Ala Asn Pro Ile Asn Pro Ala Ser Thr Glu Asp Asp Thr Thr Phe Ser Thr Thr     111
Met Ser Asp Arg Val Glu Val Asn Glu Arg Arg Ser Asp Ser Val Ser Glu Lys Glu Pro Ala Arg Asp Asp Ala Arg Lys      27

TCT GCG TCA CAA GAT CCA CCT CGA GAA TCA TCA CTT TTC TCC TCA CTC AAC ATT CGC GCT ATT CTT GTG CAA GCG ATC GGT    1899
Ser Ala Ser Gln Asp Pro Pro Arg Glu Ser Ser Leu Phe Ser Ser Leu Asn Ile Arg Ala Ile Leu Val Gln Ala Ile Gly     138
Asp Val Thr Asp Asp Gln Glu Asp Ala Pro Pro Phe Met Thr Ala Asn Asn Ala Arg Val Met                            48

GGC AGC CTG AAC GGC TAC TCG ATC GGC TTT GTC GGC GTG TAC TCA ACG CTG TTT GGC TAC AGT ACG AAC TGC GCG AGC TTC    1980
Gly Ser Leu Asn Gly Tyr Ser Ile Gly Phe Val Gly Val Tyr Ser Thr Leu Phe Gly Tyr Ser Thr Asn Cys Ala Ser Phe     165
          1

CTC CAG GAG AAC AGC TGC ACG ACG GTG CCC AAC GCT GAC TGC AAG TGG TTT GTG AGC CCG ACT GGC AGC AGC TAC TGT GGC    2061
Leu Gln Glu Asn Ser Cys Thr Thr Val Pro Asn Ala Asp Cys Lys Trp Phe Val Ser Pro Thr Gly Ser Ser Tyr Cys Gly     192

TGG CCC GAG GTC ACG TGC CGC AAG GAG TAT GCT TAC TCC AGT CCT GCG GAG ATG CCA GGT GCG CTT GCC CGG TGC GAG GCA    2142
Trp Pro Glu Val Thr Cys Arg Lys Glu Tyr Ala Tyr Ser Ser Pro Ala Glu Met Pro Gly Ala Leu Ala Arg Cys Glu Ala     219

GAC TCG CGG TGC AGG TGG TCG TAC AGC GAC GAG GAG TGC CAG AAC CCG TCG GGC TAC TCA TCG TCG GAA AGC GGT ATC TTT    2223
Asp Ser Arg Cys Arg Trp Ser Tyr Ser Asp Glu Glu Cys Gln Asn Pro Ser Gly Tyr Ser Ser Ser Glu Ser Gly Ile Phe     246

GCT GGC TCG ATG ATT GCC GGC TGC CTG ATC GGC TCC GTC TTT GCT GGT CCG CTT GCG TCG AAG ATC GGT GCG AGG CTC TCG    2304
Ala Gly Ser Met Ile Ala Gly Cys Leu Ile Gly Ser Val Phe Ala Gly Pro Leu Ala Ser Lys Ile Gly Ala Arg Leu Ser     273
                    2

TTC CTG CTC GTT GGT CTC GTG GGT GTT GTG GCG TCG GTG ATG TAC CAC GCG TCG TGC GCG GCG GAC GAG TTT TGG GTA CTG    2385
Phe Leu Leu Val Gly Leu Val Gly Val Val Ala Ser Val Met Tyr His Ala Ser Cys Ala Ala Asp Glu Phe Trp Val Leu     300
                        3

ATC GTC GGC CGC TTC GTG ATT GGT CTG TTC TTG GGC GTG ATC TGC GTT GCG TGT CCT GTG TAC ACT GAT CAG AAC GCG CAC    2466
Ile Val Gly Arg Phe Val Ile Gly Leu Phe Leu Gly Val Ile Cys Val Ala Cys Pro Val Tyr Thr Asp Gln Asn Ala His     327
                      4

CCG AAG TGG AAG CGC ACG ATT GGC GTG ATG TTC CAG GTA TTC ACG ACG TTG GGC ATC TTC GTC GCT GCG CTG ATG GGC CTT    2547
Pro Lys Trp Lys Arg Thr Ile Gly Val Met Phe Gln Val Phe Thr Thr Leu Gly Ile Phe Val Ala Ala Leu Met Gly Leu     354
                                          5

GCG CTT GGC CAG AGC ATC CGG TTT GAT CAC GAC GGA GAC CAG AAG GTG ATG GCG CGC ATG CAG GGC CTG TGC GTG TTC TCG    2628
Ala Leu Gly Gln Ser Ile Arg Phe Asp His Asp Gly Asp Gln Lys Val Met Ala Arg Met Gln Gly Leu Cys Val Phe Ser     381
                    T                                                                                        1237

ACC CTG TTT AGT CTT CTG ACG GTC GTG CTT GGG ATT GTG ACG AGG GAA TCG CGC GCA AAG TTC GAC GGC GGT GAG GAG GGC    2709
Thr Leu Phe Ser Leu Leu Thr Val Val Leu Gly Ile Val Thr Arg Glu Ser Arg Ala Lys Phe Asp Gly Gly Glu Glu Gly     408
          6

GGC GCT GAG CTG AAC CCG AGC GAG TAC GGC TAC GTC GAG ATG ATC CCG CGA CTG CTG ATG GGC TGC GTG ATG GCC GGC ACG    2790
Gly Ala Glu Leu Asn Pro Ser Glu Tyr Gly Tyr Val Glu Met Ile Pro Arg Leu Leu Met Gly Cys Val Met Ala Gly Thr     435

CTG CAG                                                                                                      2796
Leu Gln                                                                                                       437
  7
```

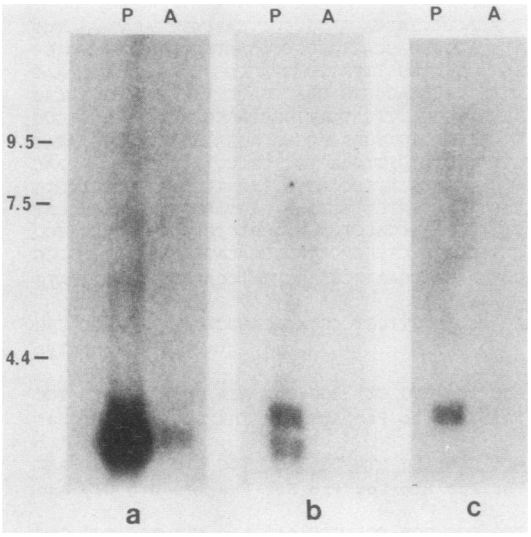

FIG. 4. Northern blots of *Leishmania* RNA, revealing a unique RNA encoded by the first unit of the Pro-1 tandem repeat. Total RNAs (5 μg) from promastigotes (lane P) and from amastigotes (lane A) were separated on a 1% agarose-formaldehyde gel (17) and blotted onto a nylon membrane. Filter a was hybridized to the entire Pro-g1 probe (Fig. 1), filter b was hybridized to the p2.75 probe, and filter c was hybridized to the *Hin*dIII-*Pst*I fragment of p2.75 that is unique to the first unit of the repeat (fragment designated S1 probe in Fig. 1). Numbers at the left indicate the sizes, in kilobases, of RNA molecular weight markers.

identify unique 5′ and 3′ restriction fragments and multiple identical internal fragments (18). Consequently, any sequence variation within the internal copies is not extensive enough to appear as restriction site polymorphisms in these restriction digests, whereas the unique amino-terminal sequence of p2.75 results in polymorphisms in a *Pst*I site (18), a *Hin*dIII site, and an *Xho*I site (Fig. 1).

Unique sequence of isoform 1 defines a new structural domain. The distinct amino-terminal sequences of isoforms 1 and 2 are both hydrophilic when analyzed by the graphic method of Kyte and Doolittle (10) (data not shown), as observed for the amino-terminal domains of all members of the glucose transporter superfamily. We emphasize that the point where the sequences of isoforms 1 and 2 begin to diverge is precisely at the beginning of the first predicted transmembrane segment (Fig. 3). Consequently, the divergent regions of these two isoforms are confined to the amino-terminal hydrophilic domains, which are believed to reside on the cytoplasmic side of the plasma membrane (4, 8), and the rest of the two sequences are identical over the region sequenced here.

Furthermore, the unique hydrophilic domain of isoform 1 (amino acids 1 to 132) is not significantly related in amino acid sequence to any other sequence reported in the NBRF protein data base (May 1990), as determined by using the FASTA algorithm (15) for detecting sequence similarity.

The first unit in the Pro-1 repeat encodes a unique stable RNA. The sequence of p2.75 reveals that this DNA has the potential to code for a polypeptide and suggests that it represents a functional gene and not a pseudogene. If this first copy of the repeat is transcribed and translated, there should exist a unique stable mRNA with the potential to code for isoform 1. Figure 4 shows Northern (RNA) blots of total promastigote and amastigote RNA probed with an internal unit of the genomic repeat (Pro-g1 probe; Fig. 4a), with the p2.75 insert (Fig. 4b), and with a probe (Fig. 4c) unique to the first unit of the repeat (*Pst*I-*Hin*dIII fragment of p2.75 also designated S1 probe in Fig. 1). The first two probes revealed the 3.3-kb Pro-1 mRNA previously reported (3) as well as another, less abundant RNA of about 3.7 kb. In contrast, the probe that is unique to p2.75 detected only the 3.7-kb RNA (Fig. 4c). This result indicates that p2.75 codes for a stable RNA, likely to be an mRNA, that is about 400 nucleotides larger than the mRNA templated from the internal copies of the repeat. Like the isoform 2 mRNA (3), this isoform 1 RNA encoded by p2.75 accumulates more abundantly in promastigotes than in amastigotes.

To confirm the Northern blot results and to map the 5′ map of the p2.75 mRNA, we performed S1 mapping experiments (11), using the 5′-end-labeled probe unique to the p2.75 insert, designated S1 probe in Fig. 1. In these experiments (Fig. 5A), promastigote RNA protected a major band of 430 nucleotides (Fig. 1; Fig. 5A, lane 3), confiriming the presence of a unique transcript from the p2.75 insert. Primer extension, using a synthetic icosanucleotide primer (designated PE primer in Fig. 1 and underlined in Fig. 3) that was end labeled on the same nucleotide as the S1 probe, produced a major primer extension product that was about 35 nucleotides longer than the S1 nuclease-protected fragment (Fig. 5B; compare lanes 3 and 4). This result is typical for comparison of S1 nuclease-protected and primer-extended fragments derived from the 5′ termini of *Leishmania* mRNAs (11) and is due to the presence of the 39-nucleotide (16) spliced leader or miniexon sequence (13) that is *trans*-spliced onto the 5′ ends of all known mRNAs from kinetoplastid protozoa (2). Furthermore, two potential splice acceptor sites (12) occur within the p2.75 sequence (underlined AGs in Fig. 3) at the position where spliced leader addition should occur, as determined by these mapping experiments. These observations provide further evidence that a unique functional mRNA is encoded by the terminal copy of the Pro-1 gene. The less intense bands above and below the major primer extension fragment could be due to mRNA precursors (14) and reverse transcriptase pause sites (17), respectively.

Possible biological significance of multiple transporter isoforms. The evidence presented above suggests that the first unit of the Pro-1 tandem repeat codes for a structural isoform of the Pro-1 transporter that is distinct from the protein products of the internal repeat units. To confirm this prediction, it will be necessary to raise an antibody that is specific for the isoform 1 polypeptide. Although multiple isoforms of mammalian glucose transporters have been observed, the structure of the *Leishmania* Pro-1 isoforms is unique. The differences between these *Leishmania* isoforms are confined exclusively to the amino-terminal hydrophilic domains, whereas the differences among the mammalian glucose transporters are distributed throughout the protein-coding sequences (7).

The predicted existence of different Pro-1 isoforms suggests that these transporters are likely to have different kinetic properties, as has been observed for the different mammalian glucose transporters (7). The probable orientation of the unique hydrophilic domains on the cytoplasmic side of the plasma membrane also raises the possibility that these two isoforms interact with different cytoplasmic polypeptides or even have different subcellular localizations. We are attempting to investigate these possibilities by developing both a functional assay and specific antibody probes for the two isoforms of the Pro-1 transporter.

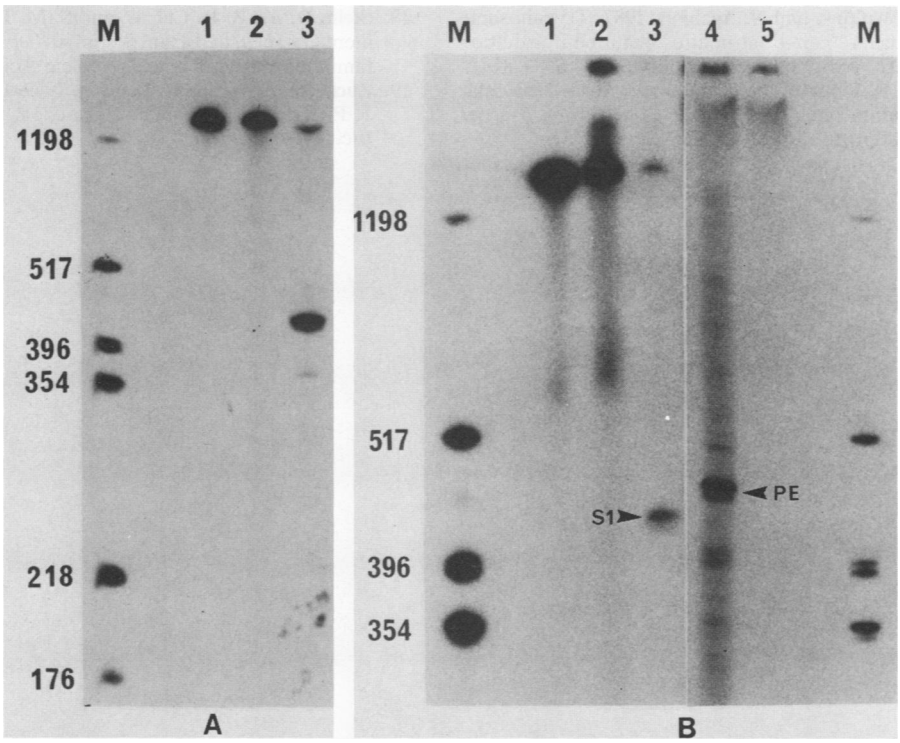

FIG. 5. RNA unique to the first unit of the Pro-1 tandem repeat, detected by S1 mapping and by primer extension. (A) S1 mapping experiments were performed as previously described (11), using an end-labeled probe unique to the first unit of the repeat (S1 probe in Fig. 1) hybridized in the presence (lane 3) or absence (lane 2) of 60 μg of total promastigote RNA. Lane 1 is the S1 probe prior to hybridization. (B) S1 mapping experiments (lanes 1 to 3) were performed as for panel A. Primer extension experiments (lanes 4 and 5) were performed by using an end-labeled icosanucleotide primer (Fig. 3) that contains the same 5′ terminus as does the S1 probe. Primer extensions were performed as detailed previously (11) in the presence (lane 4) or absence (lane 5) of 5 μg of polyadenylated RNA from promastigotes. Lanes M are end-labeled *Hin*fI fragments of SP64 DNA; numbers at the left indicate the sizes, in base pairs, of these molecular weight markers. All samples in panel B were run on the same gel, but lanes 1 to 3 were from a longer exposure than were lanes 4 and 5. For comparison, the arrowhead marked S1 indicates the most abundant fragment protected from S1 nuclease, and the arrowhead marked PE indicates the most abundant primer extension product.

Nucleotide sequence accession number. The GenBank accession number for the sequence reported here is M38214.

We thank Eric Barklis and Margaret Baron for thoughtful comments on the manuscript.

This work was supported by Public Health Service grant AI25920 from the National Institutes of Health. S.M.L. acknowledges a New Investigator Award in Molecular Parasitology from the Burroughs Wellcome Fund.

LITERATURE CITED

1. **Baldwin, S. A., and P. J. F. Henderson.** 1989. Homologies between sugar transporters from eukaryotes and prokaryotes. Annu. Rev. Physiol. **51**:459–471.
2. **Borst, P.** 1986. Discontinuous transcription and antigenic variation in trypanosomes. Annu. Rev. Biochem. **55**:701–732.
3. **Cairns, B. R., M. W. Collard, and S. M. Landfear.** 1989. Developmentally regulated gene from *Leishmania* encodes a putative membrane transport protein. Proc. Natl. Acad. Sci. USA **86**:7682–7686.
4. **Davies, A., K. Meeran, M. T. Cairns, and S. A. Baldwin.** 1987. Peptide-specific antibodies as probes of the orientation of the glucose transporter in the human erythrocyte membrane. J. Biol. Chem. **262**:9347–9352.
5. **Eisenberg, D., E. Schwarz, M. Komaromy, and R. Wall.** 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. J. Mol. Biol. **179**:125–142.
6. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabelling DNA restriction fragments to high specific activity. Anal. Biochem. **132**:6–13.
7. **Gould, G. W., and G. I. Bell.** 1990. Facilitative glucose transporters: an expanding family. Trends Biochem. Res. **15**:18–23.
8. **Haspel, H. C., M. G. Rosenfeld, and O. M. Rosen.** 1988. Characterization of antisera to a synthetic carboxyl-terminal peptide of the glucose transporter protein. J. Biol. Chem. **263**:398–403.
9. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **12**:857–872.
10. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
11. **Landfear, S. M., S. I. Miller, and D. F. Wirth.** 1986. Transcriptional mapping of *Leishmania enriettii* tubulin mRNAs. Mol. Biochem. Parasitol. **21**:235–245.
12. **Layden, R. E., and H. Eisen.** 1988. Alternate *trans*-splicing in *Trypanosoma equiperdum*: implications for splice site selection. Mol. Cell. Biol. **8**:1352–1360.
13. **Miller, S. I., S. M. Landfear, and D. F. Wirth.** 1986. Cloning and characterization of a *Leishmania* gene coding for a spliced RNA leader sequence. Nucleic Acids Res. **18**:7341–7360.
14. **Muhich, M. L., and J. C. Boothroyd.** 1988. Polycistronic transcripts in trypanosomes and their accumulation during heat shock: evidence for a precursor role in mRNA synthesis. Mol. Cell. Biol. **8**:3837–3846.
15. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85**:2444–2448.

16. **Perry, K., K. P. Watkins, and N. Agabian.** 1987. Trypanosome mRNAs have unusual "cap 4" structures acquired by addition of a spliced leader. Proc. Natl. Acad. Sci. USA **84**:8190–8194.
17. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, vol. 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Stein, D. A., B. R. Cairns, and S. M. Landfear.** 1990. Developmentally regulated transporter in *Leishmania* is encoded by a family of clustered genes. Nucleic Acids Res. **18**:1549–1557.
19. **Zuckerman, A., and R. Lainson.** 1979. Leishmania, p. 57–133. *In* J. P. Krier (ed.), Parasitic protozoa, vol. 1. Academic Press, Inc., New York.