FL-160 Proteins of *Trypanosoma cruzi* Are Expressed from a Multigene Family and Contain Two Distinct Epitopes that Mimic Nervous Tissues

By Wesley C. Van Voorhis,* Lynn Barrett,* Raymond Koelling,‡ and Andrew G. Farr‡

From the *Department of Medicine, Division of Allergy and Infectious Diseases, and the ‡Department of Biological Structure, School of Medicine, University of Washington, Seattle, Washington 98195

Summary

The partial sequence of a gene encoding the COOH terminus of a protein of apparent molecular weight of 160 kD associated with the flagellum of trypomastigotes of *Trypanosoma cruzi* (FL-160 now renamed to FL-160-1) has been previously reported. The COOH terminus of FL-160-1 has an epitope, defined by 12 amino acids, which molecularly mimics a nervous tissue antigen of 48 kD found in myenteric plexus, sciatic nerve, and a subset of cells in the central nervous system. We now report that FL-160 is a family of highly related genes. The sequence has been determined for the entire open reading frame (ORF) of one of the members of the FL-160 gene family (FL-160-2) and three other partial ORFs. Sequence analysis reveals the various members of the FL-160 gene family to be ~80% homologous in the predicted amino acid sequence, but all retain the 12-amino acid molecular mimicry epitope on the COOH terminus. Comparison of the sequence of FL-160-2 to other sequences demonstrates amino acid homology to bacterial sialidase (27%), members of the SA85 gene family (25–30%) and the shed acute-phase antigen/neuraminidase/trans-sialidase gene family (25–30%). Quantitative hybridization at high stringency suggests 750 copies of FL-160 are present in the DNA of each parasite. Reverse transcription and sequence analysis demonstrates that at least five of the members of the FL-160 gene family are transcribed. The NH_2 terminus of one of the FL-160 gene products was expressed and antibodies prepared. Antibodies directed to either the COOH or the NH_2 terminus of FL-160 bind a 160-kD *T. cruzi* protein. Both antibodies bind the surface membrane in the flagellar pocket of the trypomastigote. Antibodies to the NH_2 terminus bind epineurium and scattered linear densities in sciatic nerve in a pattern distinct from the pattern with antibodies to the COOH terminus. Thus, there are at least two distinct molecular mimicry epitopes on the FL-160 molecule and both mimic epitopes found in nervous tissues. FL-160 may be involved in the generation of autoimmunity to nervous tissues by molecular mimicry, observed in chronic Chagas' disease.

T he World Health Organization estimates that 16–18 million people are infected with *Trypanosoma cruzi*, the protozoan kinetoplastid parasite that causes Chagas' disease or American trypanosomiasis (1). Most persons infected with *T. cruzi* are asymptomatic. 10–30% of those infected with *T. cruzi* will have manifestations of Chagas' disease in their lifetime (1). The manifestations include cardiomyopathy, neuropathies, and dilatation of colon or esophagus due to destruction of the myenteric plexus (1). In sites of tissue destruction, parasites are almost never found but inflammatory infiltrates of lymphocytes and macrophages are frequent (2). Autoantibodies and lymphocytes directed to cardiac and nervous tissue are frequently found in those with Chagas' disease and experimental animal models of Chagas' disease (for

a review see reference 3). This has led to the hypothesis that autoimmunity is responsible for the pathophysiology of Chagas' disease. Many workers have demonstrated that autoreactive Abs and lymphocytes are also directed to *T. cruzi* antigens (for a review see reference 3). Autoimmunity in Chagas' disease has been hypothesized to develop because of a loss of tolerance for self-antigens combined with molecular mimicry of host tissues (particularly nervous and cardiac tissues) by *T. cruzi* (3).

FL-160 is a candidate for molecular mimicry of host nervous tissues by *T. cruzi*. FL-160 has been defined by expression of the 3' fragment of a gene from *T. cruzi*, representing the COOH terminus of the FL-160 protein (4–5). Abs to the COOH terminus FL-160 recombinant protein recognized

a 160-kD surface protein associated with the flagellum of *T. cruzi* trypomastigotes, the motile form of *T. cruzi* found in mammals. These Abs directed to FL-160 also bound mammalian myenteric plexus and peripheral nerve axons (4). The FL-160 cross-reactive epitope for binding of anti-FL-160 Abs to mammalian nervous tissues was mapped to a single 12-amino acid peptide encoded within the 3' end of the FL-160 gene (5). It was also shown that humans with Chagas' disease make Abs to this cross-reactive epitope and that these Abs bind human nervous tissue (5). Thus, the molecular mimicry epitope of FL-160 may have a role in the development of autoimmunity in Chagas' disease.

We now report that FL-160 is a family of genes in *T. cruzi*. We have cloned partial sequences of three members of this gene family and a complete sequence of one. All of the members cloned to date encode the 12-amino acid molecular mimicry epitope. The NH₂ terminus of FL-160 has amino acid homology to bacterial sialidases. Quantitative hybridization suggests that 750 copies of the FL-160 genes are found in each parasite. At least five copies of the FL-160 genes appear to be transcribed. Abs directed to the COOH and NH₂ termini FL-160 both react with a 160-kD protein on the surface of *T. cruzi* trypomastigotes present on the surface membrane in the flagellar pocket of trypomastigotes. Abs to the NH₂ terminus of FL-160 appear to cross-react with a different epitope of sciatic nerve than the Abs directed to the COOH terminus. Thus, FL-160 has two molecular mimicry epitopes in common with mammalian nervous tissues.

Materials and Methods

Parasites. A subclone of *T. cruzi* CL strain (6) was used for all the experiments described in this paper. Trypomastigotes were grown in coculture with murine BALB/c 3T3 fibroblasts in DMEM with 10% newborn calf serum, and epimastigotes were grown in axenic cultures in liver-infusion/tryptone medium with 10% FCS (4).

Cloning and Sequencing of FL-160 Genes. The cloning and sequencing of FL-160 was previously described (4, 5). FL-160 will be renamed FL-160-1 to distinguish it from other members of the multigene family. FL-160-2 was selected from a library of genomic DNA cut with BamHI and ApaI, size selected for 3–8 kB fragments, and ligated into BamHI and ApaI cut Bluescript™ SK-plasmid (pBS¹; Stratagene, La Jolla, CA). FL-160-3 and 4 were selected from a library of randomly sheared *T. cruzi* DNA in λ gt11 that has been previously described (4). These clones were initially selected by hybridization with ³²P-labeled FL-160-1 PCR products (see 3' FL-160 probes below) using high stringency washes (0.2× SSPE [30 mM NaCl, 2 mM NaPO₄, 0.2 mM Na₂ EDTA, pH 7.4]/0.1% SDS at 65°C). Sequencing was performed in both directions along the open reading frame (ORF)¹ using PCR-mediated amplification of ~300–400 bp products (7). In brief, 1 ng of cloned DNA was amplified for 27 cycles of 94 for 1 min, 55 for 1 min, and 72 for 1 min. The reaction product was selected using a low melting point 2% agarose gel and purified from the agarose using glass absorption (Geneclean II™; Bio 101, Inc., La Jolla, CA). Then, one of the two primers in combination with the Dye Deoxy™ Terminator Cycle Sequencing kit (Applied Biosystems,

Inc., Foster City, CA) was thermocycled, the products purified on a spin column (CENTRI-SEP™, Princeton Separations, Princeton, NJ), and run on an automated DNA sequencer (model 373; Applied Biosystems, Inc.) at the Molecular Pharmacology Pharmacology Facility, University of Washington. Where this did not lead to an unambiguous sequence, single stranded DNA (ssDNA) was used for template. ssDNA was generated from pBS FL-160 subclones using m13 helper phage and purified as recommended by the manufacturer (Stratagene). Sequencing was carried out on ssDNA with either the fluorescence-labeled forward (21M13) primer or T3 primer (Dye-primer kit™, Applied Biosystems, Inc.) or unlabeled primer and the Dye-terminator™ kit, as necessary, to resolve ambiguities. Oligonucleotides used for sequencing were as follows (except as noted, number refers to first bases' relation to FL-160-2 for sense [S] or last bases' relation for antisense[AS]): sense oligos: S 1846 5'-GGGCTGTTTGTGTACGTG-3'; S 2812 5'-ACAAGCACAACGCCTGGG-3'; S 775 5'-CATGAGGGATCC-ACCCAGCC-3'; S 479 5'-CCAGGATCCCGGACGGTGTTG-3'; FL-160-3 S 17 5'-GAAGGATCCAAATTGTTGGTACAGGCG-3'; S 1126 5'-GTCGGATCCAGCACTCTGCTGTAC-3'; and antisense oligos: AS 450 5'-CCGTGAATTCAGTCGTTGACGTGCC-3'; AS 1324 5'-CGTCCTTAAGCTTCCATGCAGCC-3'; AS 911 5'-GTG-TCGAATTCCTCATACCAC-3'; FL-160-4 AS 297 5'-CACGTC-GTTCATTTGCATCC-3'; AS 1381 5'-AGCCACTGAGCCAA-TTATTC-3'; AS 2195 5'-GGTGATGGTGTGGCGTC-3'; AS 668 5'-CGACCATCGTCCTCCATC-3'; AS-120 5'-ATGTGTGTGTG-GGCGGTG-3'; AS 178 5'-GGGGATCCTGAAACGCCTGTACC-AAC-3'; AS 219 5'-GGCCACAACGACACCATTC-3'; AS 180 5'-GAGGATCCACGGAAAGAGTGAAACG-3'; AS 1757 5'-CTG-CCGTTGAGAATTCTCCTCCAC-3'; AS 1926 5'-CTGAATTC-CCTCGTAATCCTCATCTTC-3' and all primers used for PCR subcloning of FL-160-1 (5). Many primers have one to four base differences from the expected sequence to allow the addition of restriction endonuclease sites to the PCR products. Oligonucleotides were synthesized on a DNA synthesizer (model 381A; Applied Biosystems, Inc.) at the Molecular Pharmacology Facility.

Analysis of DNA sequencing data was carried out using the IntelliGenetics™ package of programs (IntelliGenetics Inc., Mountain View, CA). ORFs in GenBank and EMBL were translated in three frames and compared to the FL-160-2 amino acid sequence using a fast pairwise comparison of sequences (FastDB™ program; IntelliGenetics Inc.). Alignments were made within this program, and homology refers to identical matches within the overlapping ORFs.

PCR Amplification of FL-160 Transcripts. Amplification of FL-160 transcripts was accomplished by reverse transcription with antisense oligo FL-160-2 AS 450 for cDNA 1 and antisense oligo (FL-160-1 AS 488 (5) 5'-CTGAATTCAGTGGTTTTCCTCTG-CGG-3') for cDNA 2–5. 5 μg of total trypomastigote RNA was reverse transcribed as previously described (8) except that a 37°C incubation temperature was substituted for 42°C and 2 μM of antisense primer was used. Balanced PCR using 5% of the resultant cDNA, sense oligo related to the first 21 nucleotides of the *T. cruzi* spliced leader sequence (21-ME oligonucleotide sequence: 5'-CCC AAGCTTCGCTATTATTGATACAGTTTC-3') and antisense oligos: FL-160-2 AS 450 for cDNA 1, AS 1926 for cDNA 2, FL-160-2 AS 1381 for cDNA 3, FL-160-4 AS 297 for cDNA 4, and FL160-2 1324 for cDNA 5. PCR was carried out for 32 cycles of 96°C for 1 min, 54°C for 1 min, and 72°C for 2 min. Products were visualized by ethidium bromide staining of a 1% agarose gel, transferred to a nylon membrane (Hybond-N™; Amersham, Arlington Heights, IL), and probed with ³²P-labeled common oligonucleotide (FL-160-2 AS 911) and washed at high stringency.

¹ *Abbreviations used in this paper:* AS, antisense; CUB gene, calmodulin-ubiquitin-associated gene; GST, glutathione-S-transferase; ORF, open reading frame; pBS, Bluescript™; S, sense; Tc exoant, *T. cruzi* exoantigen.

FL-160 cDNA 1 PCR product was excised from a low melting point agarose gel, cut with HindIII and EcoRI, ligated into HindIII and EcoRI cut pBS (Stratagene), and a single clone sequenced using the splice leader, FL-160-2 AS 450, and FL-160-2 AS 219 primers, and the Dye-terminator™ sequencing process as described above.

FL-160 cDNA 3 and 4 PCR products were excised from a low melting point agarose gel, and directly sequenced with the 21-ME oligonucleotide and FL-160-2 AS 1381 for cDNA 3 or FL-160-4 AS 297 for cDNA 4 and the Dye-terminator™ sequencing process as described above. The FL-160 cDNA 3 and 4 PCR products were subcloned into the pCR plasmid vector (Invitrogen Corp., San Diego, CA). Plasmid was isolated from three colonies derived from cDNA 3 (FL-160 cDNA 3A-C) and two colonies derived from cDNA 4 (FL-160 4A and B). These plasmids used to prime a PCR reaction using oligonucleotides FL-160-2 AS 1381 for cDNA 3 or FL-160-4 AS 297 for cDNA 4 combined with FL-160-2 S 775. The resultant PCR products were excised from a low melting point gel and sequenced with oligonucleotides FL-160-2 AS 1381 for cDNA 3A-C or FL-160-4 AS 297 for cDNA 4A and B and the Dye-terminator™ sequencing process as described above.

Production of FL-160 NH₂-terminal Protein and Abs. Production of the recombinant protein representing the NH₂ terminus of FL-160-3 was accomplished using a PCR-generated product subcloned into pGEX-2 vector (9). The product was generated using primers FL-160-3 S17 and FL-160-2 AS 1757, cut with BamHI and EcoRI, and ligated onto BamHI and EcoRI cut GST pGEX-2. This yielded a glutathione-S-transferase (GST) fusion protein with the FL-160-3 fused to GST (see Fig. 1).

A clone of *Escherichia coli* XL-1Blue™ (Stratagene) containing the appropriate PCR product cloned into pGEX 2 was grown to log phase, then stimulated with 0.1 mM isopropyl β-D-thiogalactopyranoside for 2 h. The fusion protein was found mainly in inclusion bodies. The cell pellet was treated with 1 mg/ml lysozyme, washed with 0.1% deoxycholate, treated with 10 μg/ml DNase 1, washed with 1% NP-40, and the insoluble inclusion bodies solubilized in 8 M urea (10). This was dialyzed against PBS, and the resultant soluble fraction analyzed by SDS-PAGE. Over 90% of the stained protein migrated at 83 kD, as expected for the fusion protein (data not shown). A New Zealand White rabbit was immunized four times at monthly intervals with 250 μg of the fusion protein, and the immune serum taken 1 wk after the last immunization. Abs to GST were removed by absorption to GST coupled to Sepharose 4B (4), and Abs directed to the NH₂ terminus of FL-160-3 affinity purified using the NH₂-terminal fusion protein coupled to CNBr-activated Sepharose 4B (4).

Southern and Northern Blots and Gene-counting Analysis. Northern analysis was as previously described (4). Southern blots were performed using 10 μg genomic DNA of *T. cruzi* CL strain trypomastigotes electrophoresed on a 1% agarose gel and transferred to nylon membranes (Amersham) after 20-min exposure to 0.25 M HCl, 30 min in 0.5 M NaOH, and 30 min in 0.5 M Tris buffer, pH 8.0, with 1.5 M NaCl for neutralization. Gene-counting experiments were carried out comparing hybridization to 10–320 ng pFL-160-3-PBS (7,000 bp) and *T. cruzi* DNA 61–1,000 ng bound to nylon membranes (Amersham) using a slot blot apparatus (PR 600; Hoefer, San Francisco, CA) as described by the manufacturer. Hybridization at 42°C and high stringency wash at 65°C (0.2× SSPE/1% SDS) were identical to those described previously for Northern blot analysis (4). Radioactivity bound was quantitated using phosphoimaging analysis (Phosphoimager SP; Molecular Dynamics, Sunnyvale, CA) at the Markey Molecular Medicine Center (University of Washington). ³²P-labeled 3′ probes were prepared

from PCR products of FL-160-1 previously described (5): probe 3′ 1: 48× FL-160-1 611–761 (see Fig. 2 [5]) and probe 3′ 2: FL-160-1 750–926 (see Fig. 2 [5]). The 5′ probe prepared using a PCR product of FL-160-2 and oligonucleotides FL-160-3 S 17 and FL-160-2 AS 450. The location of these probes with relation to the FL-160 genes is shown (see Fig. 3). The labeled product was synthesized using a 100-μl reaction containing 1 ng gel-purified PCR product, 5 μCi α-[³²P]dCTP (New England Nuclear, Inc., Boston, MA), 1 μM concentration of each primer, 0.1 mM dNTP, 2 U TUB™ enzyme (Amersham), and buffer as supplied by the manufacturer (Amersham). Thermal cycle incorporation of label was by 27 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min. Labeled reaction product was recovered by precipitation in 2 M NH₄Ac/EtOH, and routinely >80% incorporation was obtained. Control probe for the gene count experiment was a 341-bp PCR product within the ORF from the calmodulin-ubiquitin–associated gene (CUB gene) (11), provided by Dr. John Swindle (University of Tennessee, Memphis, TN). The probe was made using the oligonucleotides UB10 5′-GCAGGCACGTGGAACAGTGAAGG-3′ and UB13 5′-GCATTTTGTGGTTCACTTTG-3′. This was hybridized to the pBS:2.65a′ plasmid (5,500 bp, 12.5–200 pg) containing the CUB gene on the 2.65-kb BglII fragment of the calmodulin-ubiquitin locus from *T. cruzi* CL strain, as well as *T. cruzi* CL strain DNA as above.

Immunofluorescence, Immunoperoxidase, Immunoelectron Microscopy, and Flow Cytometry Analysis. Live *T. cruzi* trypomastigotes were washed twice by centrifugation in DMEM, exposed to cold DMEM containing 5% goat serum and 10 μg/ml affinity-purified rabbit Abs at 4°C for 30 min, washed once by centrifugation, and fixed in 4% paraformaldehyde in PBS for 30 min at 4°C. Where competitions of Ab binding with recombinant proteins were performed, affinity-purified rabbit Abs (10 μg/ml) were preincubated with 500 μg/ml recombinant protein for 1 h at 4°C. For immunofluorescence and flow cytometry, the trypomastigotes were exposed to 1:300 diluted fluorescein-conjugated anti–rabbit IgG (Tago Inc., Burlingame, CA) in DMEM containing 5% goat serum for 30 min at 4°C. Parasites were washed twice by centrifugation, and evaluated using a fluorescence microscope (Carl Zeiss, Inc., Thornwood, NY) or flow cytometry. Flow cytometric analysis was carried out using a FACScan® (Becton Dickinson & Co. Immunocytometry Systems, San Jose, CA) interfaced with a computer (model 310; Hewlett-Packard, Palo Alto, CA). An excitation source of 488 nm was generated using a 15-mW air-cooled argon-ion laser. Fluorescence emission was collected through a 530/30 band pass filter. Autocomp software version 2.0 along with CaliBRITE beads (Becton Dickinson) were used to set up the instrument. The Consort-30 program (Becton Dickinson) was used for acquisition and analysis of data. For immunoelectron microscopy, Ab-opsonized fixed trypomastigotes were exposed to biotinylated goat anti–rabbit Ab and avidin-biotin-peroxidase complexes as described by the manufacturer (Vector Laboratories, Inc., Burlingame, CA), and processed as previously described (12).

Immunoperoxidase staining of mouse sciatic nerve was performed using a nerve from a normal 10-wk-old C3H/He female mouse. 7-μm cryostat sections were cut and stained as previously described (5) using 10 μg/ml of affinity-purified rabbit Abs to FL-160-3 NH₂ or FL-160-1 COOH terminus and biotinylated anti–rabbit IgG with avidin-biotin-peroxidase complexes (Vector Laboratories, Inc.). Negative controls included anti-GST Abs (5) and rabbit anti-SA85 1.1 affinity-purified Abs (13) (gift of Stuart Kahn, Department of Pediatrics, University of Washington).

```
   10        20        30        40        50        60        70
KLLVQAFHSF RAPSLVYVNG VVVATVEAHY TNSTDQKSCV SLAAKSMESD GGEWINGTAI VFDHYDVKID

   80        90       100       110       120       130       140
RLLSPTAIVD ERDSETNALV GGYGTSTTPL TEVTGDGYWR PRMADGVVPH GTDDEEKKEF KWQSRSTSGV

  150       160       170       180       190       200       210
PYDFWEDNST NRDRFKQFLG GGGAGIRMED DGRYVLPIQA LKNNEKAVSL VIIAKGFTYG WEFSNDTSHE

  220       230       240       250       260       270       280
GCTQPAVLEW KEKELIMMTS CDDGSRRVYR SSTMGNLWYE EYDTLSRVWG NSLKRTGHGV QGGFVSATIN

  290       300       310       320       330       340       350
GQKVILVSRP VYSENHKKET GRLHLWMTDM QRIYDVGLIS AENEDVAAST LLYATVEVPS LEGGEKKEEK

  360       370       380       390       400       410       420
KLYCLYEVAA AEDGKYNIAF VDLTEKLEDM KKVLAAWKEK DAQIAKEYRC GNEKNNWLSG CGEAINKGLV

  430       440       450       460       470       480       490
GLLSSKSTDS TWADEYLCVN ATVHGEVEST SDGGLTFKGP GAWAKWPVGD MGQTVPYHFA NSKFTLVATV

               500       510       520       530
           SIDKAPEKDS SPIPLMGVRM NDAQGTVLFG LSYTHDKEWR
```

Figure 1. Predicted amino acid sequence of the NH2 terminus of FL-160-3. The 5' terminus of the FL-160-3 gene was expressed as a fusion protein with GST (see Materials and Methods). The predicted amino acid sequence of the expressed NH2 terminus of FL-160-3 is shown.

Results

FL-160 Is a Member of a Large Family of Highly Related Genes.   A 5-kb genomic fragment (FL-160-2), bounded by BamHI and ApaI was found to contain the entire ORF of a gene that is highly related to the original gene fragment that defined the FL-160 gene (FL-160 in references 4 and 5 is now called FL-160-1). The sequence and the predicted ORF of FL-160-2 are demonstrated in Fig. 2. The ATG that is hypothesized to encode the start methionine is underlined. The evidence that this is the start codon of the gene is that: (a) the codon has a Kozak start sequence consensus (14); (b) the one other in-frame ATG upstream (also underlined) does not have a Kozak start sequence consensus; (c) multiple stop codons are found upstream in all three reading frames; and (d) mRNA sequence analysis demonstrates that the proposed start ATG is near the beginning of all five transcripts cloned and the first possible ATG in a related transcript (see FL-160 cDNA 1, below). As in all kinetoplast genes described to date, introns are not found in any of the genes for FL-160. The entire ORF of the FL-160-2 gene has a predicted molecular mass of 109,126 daltons. Thus, posttranslation modification, such as glycosylation, likely affects the mobility of the mature trypomastigote 160-kD protein. In support of this, there are eleven possible N-linked glycosylation sites. The sequence is serine and threonine rich, and these two amino acids comprise 18% of all amino acids. The sequence is acidic with a predicted pI of 4.2. Glutamic and aspartic acid residues are ~15% of all amino acids in the predicted ORF.

Three related clones with partial ORFs have been isolated and sequenced. The first of these is the original FL-160 genomic clone (FL-160-1) which contained 37% of the entire ORF. The other two contain 95% (FL-160-3) and 64% (FL-160-4) of the 3' end of the ORF. All four genes isolated to date are different with respect to both the nucleotide sequence (range 5–11% differences) and predicted amino acids (range 10–21% differences; see Figs. 2 and 3 *A*).

```
FL-160-2AA  411  A E D G K Y N I A F V D L T E X L E D N
FL-160-2   1230  TGCAGAGGACGGTAAGTACAACATTGCTTTTGTGGACCTGACGGAGAAGTTGGAGGACATG
FL-160-3   1105  ------------------------------------------------------------
FL-160-4    146  ----------------------------------------------T-------------

FL-160-2AA  431  K K V L A A W N E Q D A Y I A K K Y R C G
FL-160-2   1291  AAGAAGGTGTTGGCTGCATGGAATGAACAGGACGCGTACATTGCGAAGAAGTACCGCTGCG
FL-160-3   1166  -----------------------G--GA---------C-A---------G----------
FL-160-4    207  -G-----------------------------------C-A----------G-A---G----

FL-160-2AA  452  N E X   N N W L S D C G     E A L T K G L
FL-160-2   1352  GGAATGAAAAG   AATAATTGGCTCAGTGACTGTG   GTGAAGCCCTCACGAAAGGGCT
FL-160-3   1227  -----------   -------------G-----      -----------AT-------
FL-160-4    268  -----------GATGCA---GAA-GAC---------ATG---GT-A------T--------

FL-160-2AA  470  V G F L S N K S T G K T W A D E Y L C V
FL-160-2   1407  GGTTGGCTTCTTATCCAACAAGTCAACTGGGAAAACATGGGCAGACGAGTACCTCTGCGTG
FL-160-3   1282  -------C-------G-G-----------AC-GC--G-----------------------
FL-160-4    329  -----------------------------AA--GC--G----------------------

FL-160-2AA  490  N A T V H G E V E S T P D G G L T F K G P
FL-160-2   1468  AATGCAACCGTTCACGGGGAAGTGGAAAGTACTCCTGATGTGGATTGACGTTCAAAGGAC
FL-160-3   1343  -----------------------------------T------------------------
FL-160-4    390  ---A--------------------------------------C-----------------

FL-160-2AA  511  G A W A K W P V G D N R Q T V P Y H F A
FL-160-2   1529  CCGGAGCGTGGGCGAAGTGGCCTGTTGGCGACATGAGGCAGACTGTGCCGTACCACTTTGC
FL-160-3   1404  -------------------------------------G----------A-----------
FL-160-4    451  -------AG-----GT---------------------G-----A----------------

FL-160-2AA  531  N S X F T L V A P V S I D X A P E K G S
FL-160-2   1590  GAACTCTAAGTTTACTCTTGTGGCGCCGGTGTCCATCGATAAGGCTCCGGAGAAAGGCAGC
FL-160-3   1465  --------------------------A----------------------------A----
FL-160-4    512  --------G---------------A--------------------------C--------

FL-160-2AA  551  S P I P L N G V R N N D A Q G T V L F G L
FL-160-2   1651  AGCCCCATCCCTTTGATGGGTGTGAGGATGAACGATGCACAAGGCACGGTGCTTTTTGGTC
FL-160-3   1526  ------------------------------------------------------------
FL-160-4    573  --------T---------------------------------------------------

FL-160-2AA  572  S Y T H D K E W R I I L N G S L L T L P
FL-160-2   1712  TGTCGTACACCCACGACAAAGAGTGGAGGATAATCCTCAACGGCAGTCTTTTGACGCTTCC
FL-160-3   1587  ---------------------------------T----------------G--
FL-160-4    634  -----------------A---A-----------T----------G------T-G--

FL-160-2AA  592  A H G E D A G W E A N K K Y H V V L Q N
FL-160-2   1773  GGCGCATGGTGAGGATGCCGGATGGGAAGCAAACAAGAAATATCATGTGGTACTGCAAATG
FL-160-3   1648  -------------A-----------C--------G-------------------------
FL-160-4    695  --------   --A------------------C---------------------------

FL-160-2AA  612  D Y S N G L F V Y V N G X X I Y D S E F D
FL-160-2   1834  GATTATAGTAATGGGCTGTTTGTGTACGTGAATGGAAAGAAGATATACGACAGCGAATTTG
FL-160-3   1709  -------A-G-------------------------G----------C----------
FL-160-4    753  -------A------------------G-----G----G-----G----C-----   -
FL-160-1      1                                              ----C-

FL-160-2AA  633  Y E V E D Y E D Y E S F Y G R L Q K L L
FL-160-2   1895  ATTATGAAGTTGAGGATTACGAGGATTATGAGAGCTTTTATGGCAGGTTACAAAAACTGTT
FL-160-3   1770  -----------A------------------------C-C---A----------------
FL-160-4    811  ---------------------------C---A----------A-----A----------
FL-160-1      8  ---------A----------------GC---A--------------------C-------

FL-160-2AA  653  S S H S I S H F Y I G G D G K S S D S
FL-160-2   1956  GAGCTCCCACAGCATCTCGCACTTTTACATTGGTGGGGACGGCAAGAGCAGCTCGGACAGC
FL-160-3   1831  ------------------------------------------A----------------
FL-160-4    872  -GAG----------------------------------------------A----G---
FL-160-1     69  ------------------------------C------------------------A-

FL-160-2AA  673  I H V T V S N V L L Y N R L L K D D E L N
FL-160-2   2017  ATTCATGTGACGGTGAGCAACGTCCTGCTGTACAACCGCTTATTGAAGGATGATGAGCTCA
FL-160-3   1892  -----A-----------------------------------C-A-------------
FL-160-4    933  G-----------------------------------------A--A-----------
FL-160-1    130  -----------------------------------------C-A-------------

FL-160-2AA  694  P P N K T N A A A A A E A E V P A P E G
FL-160-2   2098  ACCCACCAATGAAGACGAACGCGGCCGCTGCTGCAGAGGCGGAGGTGCCGGCCCCGGAAGG
FL-160-3   1953  ---------------------------C-------------T---------------
FL-160-4    994  ------T-A-----G-----T----------C------------T--------A----
FL-160-1    191  ------T-----------G-----------C-G-----------T-----------

FL-160-2AA  714  A P E K S H L S Q P S E K D T T L S P Q
FL-160-2   2139  TGCGCCTGAAAAAAGTCATTTGAGTCAACCCTCTGAAAAAGACACCACACTATCACCACAA
FL-160-3   2014  ------C----C---------------------------------------C-------
FL-160-4   1055  ------C----C---C------------------------G------C---------
FL-160-1    252  ------C----C--A------G-A---G-T----------------------

FL-160-2AA  734  K Q G L S P E K S K N E K Y S A G S G Q T
FL-160-2   2200  AAACAAGGATTATCGCCGGAAAAATCAAAAAATGAAAAGTATTCAGCAGGTTCCGGGCAGA
FL-160-3   2075  ----------------------------------------C-----------------
FL-160-4   1116  ----G-AC-----A-----------------------------------------G
FL-160-1    313  --T-----------A-----------------------------------------

FL-160-2AA  755  T S T D P F T G L Y T S A A E G E V E E V
FL-160-2   2261  CAACTTCTACTGATTTTTACTGGCTTGTACACCTCAGCCGCCGAGGGGGAAGTGGAAGAAGT
FL-160-3   2136  -----------------CG---------------------------------------
FL-160-4   1177  --C-----G-----CCCG------C--C------------------A--------T--
FL-160-1    374  --------G-----C---------------------

FL-160-2AA  775  A A S S A G S A S S L T P T P G E G S S
FL-160-2   2322  GGCAGCCAGCAGCGCTGGCTCGGCGTCTTCATTGACTCCAACACCGGGAGAAGGCAGTTCT
FL-160-3   2197  -------------------------------------------G------------C--
FL-160-4   1238  ----T--G----T-T--------C-----G------A------G---------C--
FL-160-1    435  ----T-------T-T----------C------G------A------G-------C--
```

```
FL-160-2AA  795  Q X A L E T D V P A T E D Q F R S E Q V H
FL-160-2   2383  CAAAAAGCACTTGAAACCGATGTGCCTGCCACTGAGGATCAGTTTGAAAGTGAACAGGTGC
FL-160-3   2258  ------------------------------------------------------------
FL-160-4   1299  --------------------------T------------C---------------A--
FL-160-1    496  ----------C------------------------------C--------C---------A--

FL-160-2AA  816  S S L S V V R P N T E Q T D E A V V A T
FL-160-2   2444  ACTCTTCTCTCAGCGTGGTTCGGCCGATGACGGAGCAGACGGACGAGGCGGTTGTCGCCAC
FL-160-3   2319  ------------------A----AA---------------------------A------
FL-160-4   1360  ----G-----------------A------------------G-T--A----T----------
FL-160-1    557  ----G-----------------A------------------G-T--A--------------

FL-160-2AA  836  P Q R K T T E D R P Q H S T L S D A S E
FL-160-2   2505  TCCGCAGAGGAAAACCACTGAAGATCGGCCGCAGCACAGCACTCTATCTGATGCTTCAGAG
FL-160-3   2380  ------------------------------------------------------------
FL-160-4   1421  ------------------------------------------------------------
FL-160-1    618  ------------------------------------------------------------

FL-160-2AA  856  N N E E S S S H S A T L T S D E Q N F G P
FL-160-2   2566  AACATGGAAGAGTCGTCGTCCCATTCCGCCACGCTCACCAGCGACGAACAAAACTTTGGCC
FL-160-3   2441  -------------------------C----------------------------------
FL-160-4   1482  G--------------A-----------------------------------C-G---A--
FL-160-1    679  G--------------A-----------------------------------C-G---A--

FL-160-2AA  877  E G R K D A H P H T A V S A S S G P D S
FL-160-2   2627  CTGAAGGGAGGAAGGACGCACATCCCCACACTGCCGTGAGTGCGAGCAGCGGCCCCGACTC
FL-160-3   2502  -----------------------------------------T----------A------
FL-160-4   1543  -C--G-A----------A--T----------G----G-------------------
FL-160-1    740  -C--G-A----------A--A----------------G-----------T--------

FL-160-2AA  897  S H S T E V A P V D G A T A A H E P S T
FL-160-2   2688  ATCCCACAGCACTGAGGTGGCGCCCGTGGACGGTGCCACAGCTGCCCATGAGCCCAGCACT
FL-160-3   2563  ----  --------------CG--------------------  --------------
FL-160-4   1604  -----------------AGCG-------------C-G-----------------
FL-160-1    801  --------------------T----------------

FL-160-2AA  917  D P E T A Q G H D E V L D G D D A A P G N
FL-160-2   2749  GACCCGGAAACTGCTCAAGGCCATGATGAAGTGTTGGATGGCGATGATGCCGCACCGGGTA
FL-160-3   2621  ------------------------------T-------------------------
FL-160-4   1665  --------------------G-----------------------------------
FL-160-1    842  ------C----------------------------------G----------AA--A-T

FL-160-2AA  938  T S T T P G E T K I P S E S N A T S H L
FL-160-2   2810  ATACAAGCACAACGCCTGGGGAGACCAAGATCCCATCGGAGTCCAACGCAACATCGCACTT
FL-160-3   2682  ------------------------------------------------------------
FL-160-4   1726  -------------------C-A--------------------------------C-
FL-160-1    902  --TTGC------------------A-----------------C------T--C

FL-160-2AA  958  D H D L L L E H G Q     F D L S A N A L I
FL-160-2   2871  GGACCATGACCTTTTGCTTGAGCACGGGCA   GTTCGACTTGTCGGCCATGGCTCTAATT
FL-160-3   2743  -----------G---------------     --------------------------
FL-160-4   1787  ---A------G---G-----------------     ----TT----T-C-----------
FL-160-1    963  ----------A------------------TTT-AG---T---G--T-----------

FL-160-2AA  977  G D S T V H G C V S R V F G L L L L G L W
FL-160-2   2929  GGTGACAGCACCGTGCACGGGTGTGTGTCTCGGGTGTTTGGGCTGCTGCTTCTGGGGCTGT
FL-160-3   2801  ----------------------------------------TT--------------C-
FL-160-4   1845  -----G------------------------------GTT-----------------
FL-160-1   1024  -----------------T-------------------GTT----------------

FL-160-2AA  998  G T A A L F .  1003
FL-160-2   2990  GGGGCACCGCGGCTCTCTTCTGAGGTGTGAGGGAGGAGTGGAATGGTGTTTCTTCGTGGAA
FL-160-3   2862  ---------------G------------------C-GC-T----------
FL-160-4   1906  --C------------G-----A----------------------C-GC-TA--------
FL-160-1   1085  ----------------G------G------A-------T---T-C-GC-T--T-------

FL-160-2   3051  CAACAACTGTGCAGCTCCC AGCATCACTCACTCAACTGGGTCAGGGATACTTTCTTT 3107
FL-160-3   2923  ---------------------C---------------TT---------- 2980
FL-160-4   1967  ------------G----------------------T-----A--A-C---A-C----C- 2024
FL-160-1   1146  ---------CAGCTC--------CT-------------TTT---A--------CT--- 1203
```

Figure 2. DNA sequence of FL-160 genes and cDNA an amino acid sequence of FL-160-2. Nucleic acid sense strand sequences (5' to 3') for genomic DNA clones of FL-160-2, FL-160-1, FL-160-3, and FL-160-4, and the FL-160 cDNA clone 1 are shown aligned. (*Dashes*) Sequence is identical to FL-160-2. (*Spaces*) Gaps are placed to maintain the best homology. The nucleic acid sequence for partial FL-160 cDNA clone 1 is also shown aligned to FL-160-2. The first 11 bases of FL-160 cDNA clone 1 correspond to the 3' 11 bases of splice leader sequence for *T. cruzi*. The ATG hypothesized to initiate translation is underlined at base 1 of FL-160-2. Another in-frame ATG, at base -111, is underlined. The -111 ATG is hypothesized not to initiate transcription because of the lack of Kozak consensus sequence for translation initiation, as well as the lack of ATG in the same spot in FL-160 cDNA 1. The predicted amino acid sequence of FL-160-2 is shown over the sequence. The 12-amino acid peptide previously mapped to be the epitope in FL-160-1, responsible for antigenic mimicry (5) with mammalian nervous tissue, is underlined with a double line. Note that the sequences are all identical in this region and that the long ORF is the same for all four sequences (data not shown). The 11 potential N-glycosylation sites are underlined with a heavy line. The nucleotide sequence data has been submitted to EMBL/GenBank/DDBJ under the accession numbers X70947, X70948, X70949, and X70950.
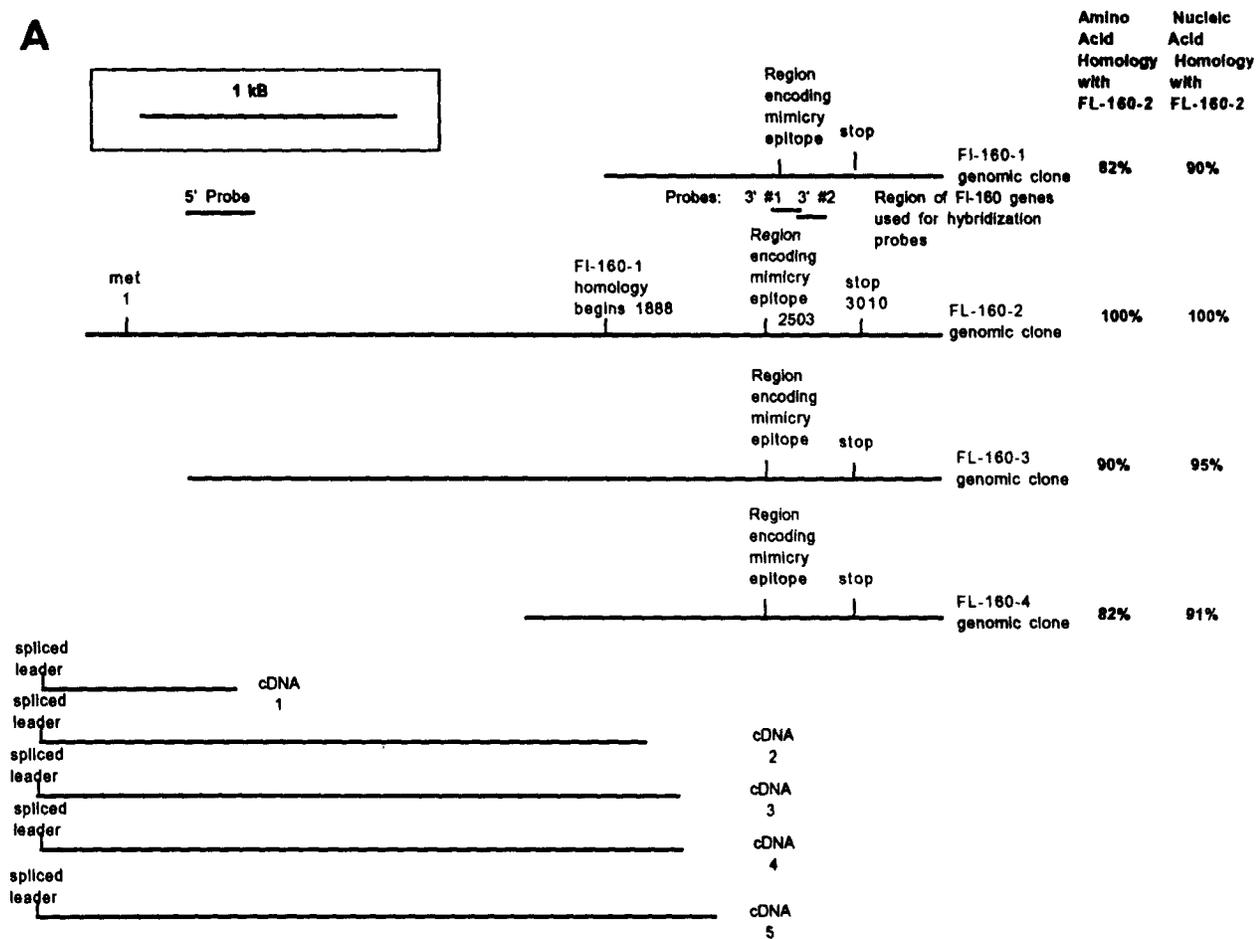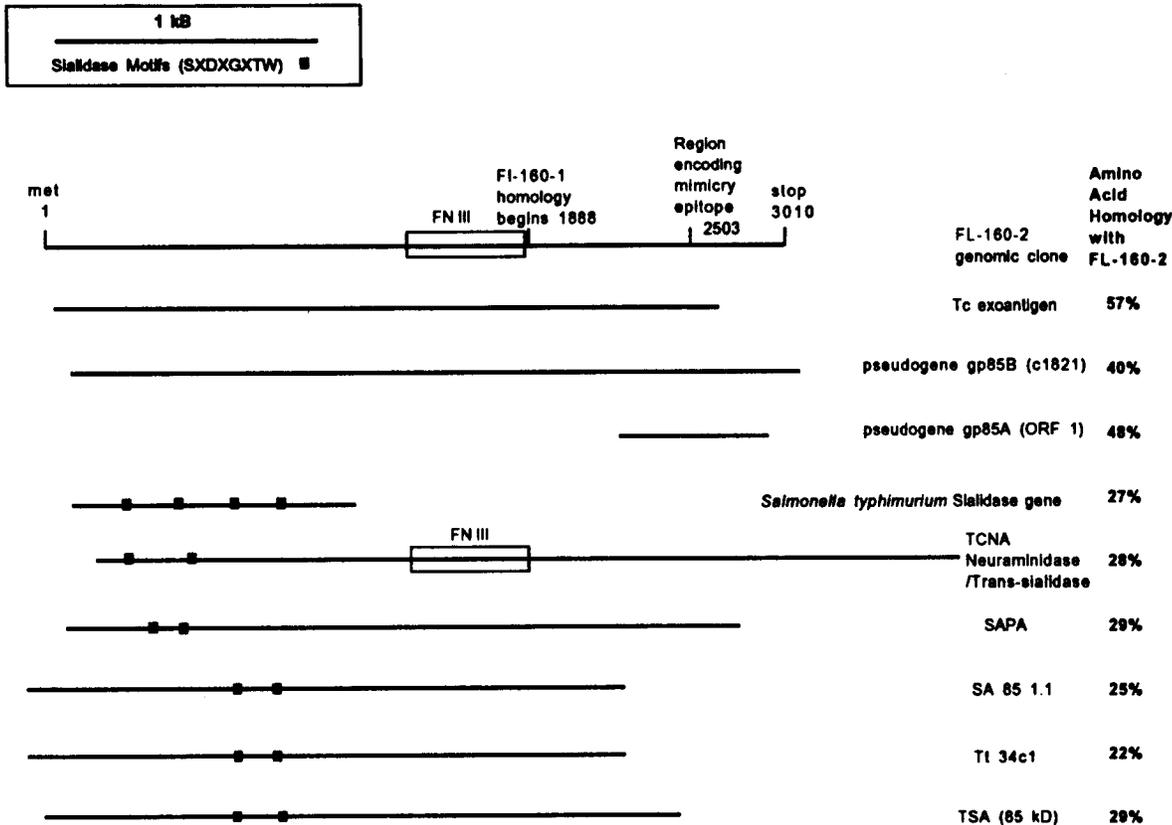
A

1 kB

| | Amino Acid Homology with FL-160-2 | Nucleic Acid Homology with FL-160-2 |

5' Probe

Region encoding mimicry epitope stop

Fl-160-1 genomic clone — 82% 90%

Probes: 3' #1   3' #2   Region of Fl-160 genes used for hybridization probes

met 1

Fl-160-1 homology begins 1888

Region encoding mimicry epitope 2503   stop 3010

FL-160-2 genomic clone — 100% 100%

Region encoding mimicry epitope stop

FL-160-3 genomic clone — 90% 95%

Region encoding mimicry epitope stop

FL-160-4 genomic clone — 82% 91%

spliced leader — cDNA 1

spliced leader — cDNA 2

spliced leader — cDNA 3

spliced leader — cDNA 4

spliced leader — cDNA 5

Figure 3. Alignments of FL-160 genes and cDNAs and related sequences. (A) Alignment of the FL-160 genes and cDNAs. The region of the FL-160 genes used for hybridization probes for Southern, Northern, and slot blot hybridization are shown above FL-160-2. The stop codon for the longest ORF and the region encoding the 12-amino acid molecular mimicry epitope is marked on each genomic clone. To the right of each genomic clone is the percent that the clone is homologous to the amino acid and nucleotide sequence of FL-160-2. (Bottom) The alignment of cDNAs from Figs. 2 and 5 shown schematically. (B) Alignment of the predicted amino acid sequence of FL-160-2 with the predicted amino acid of genes and pseudogenes listed in GenBank 72. Tc exoant and pseudogenes gp85A and B are the most highly related, and may represent different members of the same gene and pseudogene family. The other genes are more related to FL-160-2 in their 5' end, and much of this relatedness is found within the homologous region to bacterial sialidases. The sialidase motif (SXDXGXTW), when complete, is shown as a black box within the sequence. Though these motifs are incomplete in FL-160-2, two to three amino acids of the sialidase motif can be found in regions aligned with the motifs in TCNA and the SA85 genes. The location of the fibronectin III domain of TCNA and FL-160 is shown as an open box.

Comparison of the FL-160-2 predicted amino acid sequence with the DNA databases, translated into all three reading frames, using a pairwise comparison, demonstrated many sequences that are significantly related. These are shown schematically in Fig. 3 B. The sequence most related was T. cruzi exoantigen gene (Tc exoant) (Jazin, E. E., L. Aslund, D. O. Sanchez, J. Henriksson, U. Pettersson, and A. C. Frasch, unpublished data, deposited in EMBL, accession number M88337) which gave a significance of >139 SD over the mean. This sequence appears to be highly related at the amino acid (57% homology on alignment) and the nucleic acid level (64% homology on alignment). Two recently described pseudogenes are closely related to the FL-160 genes, and are likely to be members of the FL-160 gene family (15). These are

48 and 40% homologous in the amino acid sequence in the longest ORFs (significance 12 and 29 SD over the mean).

The four FL-160 genes, the T. cruzi exoantigen gene, and pseudogenes gp85A and B (FL-160 family) are much more highly related to each other than the 85-kD surface protein (SA85) family (SA85 1.1 [16], TSA 1 [17], and Tt34c1 [18]) or the trans-sialidase (trans-sialidase [19]/neuraminidase [20]/ shed acute-phase antigen [SAPA] [21]) family (Fig. 3 B). Yet these two families are significantly related to the FL-160 genes (Fig. 3 B, significance >6 SD above the mean). The FL-160-2 predicted ORF contains a fibronectin type III domain (Fig. 3 B), similar to that found in T. cruzi neuraminidase as first reported by Pereira et al. (20). This domain is defined by conserved aromatic and hydrophobic amino acids separated by

B

1 kB

Sialidase Motifs (SXDXGXTW) ■

		Amino Acid Homology with FL-160-2
met 1 ... FN III ... FI-160-1 homology begins 1868 ... Region encoding mimicry epitope 2503 ... stop 3010	FL-160-2 genomic clone	FL-160-2
	Tc exoantigen	57%
	pseudogene gp85B (c1821)	40%
	pseudogene gp85A (ORF 1)	48%
	Salmonella typhimurium Sialidase gene	27%
FN III	TCNA Neuraminidase /Trans-sialidase	28%
	SAPA	29%
	SA 85 1.1	25%
	Tt 34c1	22%
	TSA (85 kD)	29%

variable numbers of amino acids (20). The predicted amino acid sequence of the FL-160-2 genes is highly related to a bacterial sialidase gene (22) (27% homology, significance >6 SD above the mean). Note that the neuraminidase reported by Pereira et al. (20) also had 27% homology to the bacterial sialidases. Though the FL-160 genes do not contain complete sialidase motifs (SXDXGXTW), two to three of the sialidase motif amino acids can be found at the sites where the bacterial sialidases align to the FL-160-2 gene.

Southern analysis of *T. cruzi* CL genomic DNA reveals that a large number of genes are related to the FL-160 family. Southern analysis demonstrates that the FL-160 related genes vary extensively in restriction sites (Fig. 4 *a*). We focused on three probes, one from the 5' end of the gene, and two from the 3' end of the gene (Fig. 2 diagrammatically shows locations of probes). One of the 3' probes includes the coding sequence for the peptide responsible for antigenic mimicry (3' probe 1). The other 3' probe is just 3' to this probe (3' probe 2). These probes are <50% homologous to SA85 genes and trans-sialidase genes, such that hybridizing and washing at high stringency should identify only members of the FL-160 family. The pattern on Southern hybridization for most restriction enzymes appears to be similarly complex whether a probe from the 5' end of the FL-160 genes is used or whether the probe is from the 3' end. The EcoRI digest shown in Fig. 4 *a* is an example of this. Most other enzymes tested gave a wide range of hybridization bands (e.g., AvaI, EagI,

EcoRV, HincII, HindIII, NcoI, SacII, ScaI, and XhoI, data not shown). Some restriction enzymes give less complex patterns. For example, restriction of genomic DNA with PstI and hybridization with a probe from the 5' end of the gene
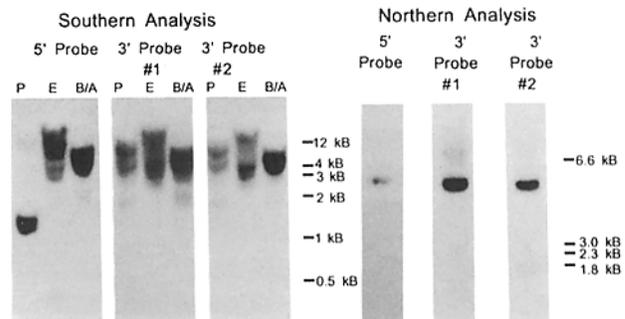
Southern Analysis   **Northern Analysis**

Southern Analysis: 5' Probe / 3' Probe #1 / 3' Probe #2 — P E B/A P E B/A P E B/A

Northern Analysis: 5' Probe / 3' Probe #1 / 3' Probe #2

—12 kB
—4 kB
—3 kB
—2 kB
—1 kB
—0.5 kB

—6.6 kB
—3.0 kB
—2.3 kB
—1.8 kB

Figure 4. Southern and Northern hybridization data using probes from different regions of the FL-160-2 gene. Southern hybridization was performed using digests of *T. cruzi* CL strain genomic DNA and PstI, EcoRI, and BamHI combined with ApaI enzymes. The blots were hybridized with labeled PCR fragments, one from the 5' end of the gene and two from the 3' end of the gene (*3' probes 1 and 2*). The 3' probe 1 encodes a region of the FL-160 genes shown previously to be responsible for antigenic mimicry. More than 20 bands of different molecular weight could be seen for each digest and probe combination except one. The PstI digest probed with the 5' probe gave a predominately 1-kb band. Northern blots were 10 μg of trypomastigote total RNA hybridized with the same three probes. A single band of ~4–5 kb was seen on all three autoradiograms.
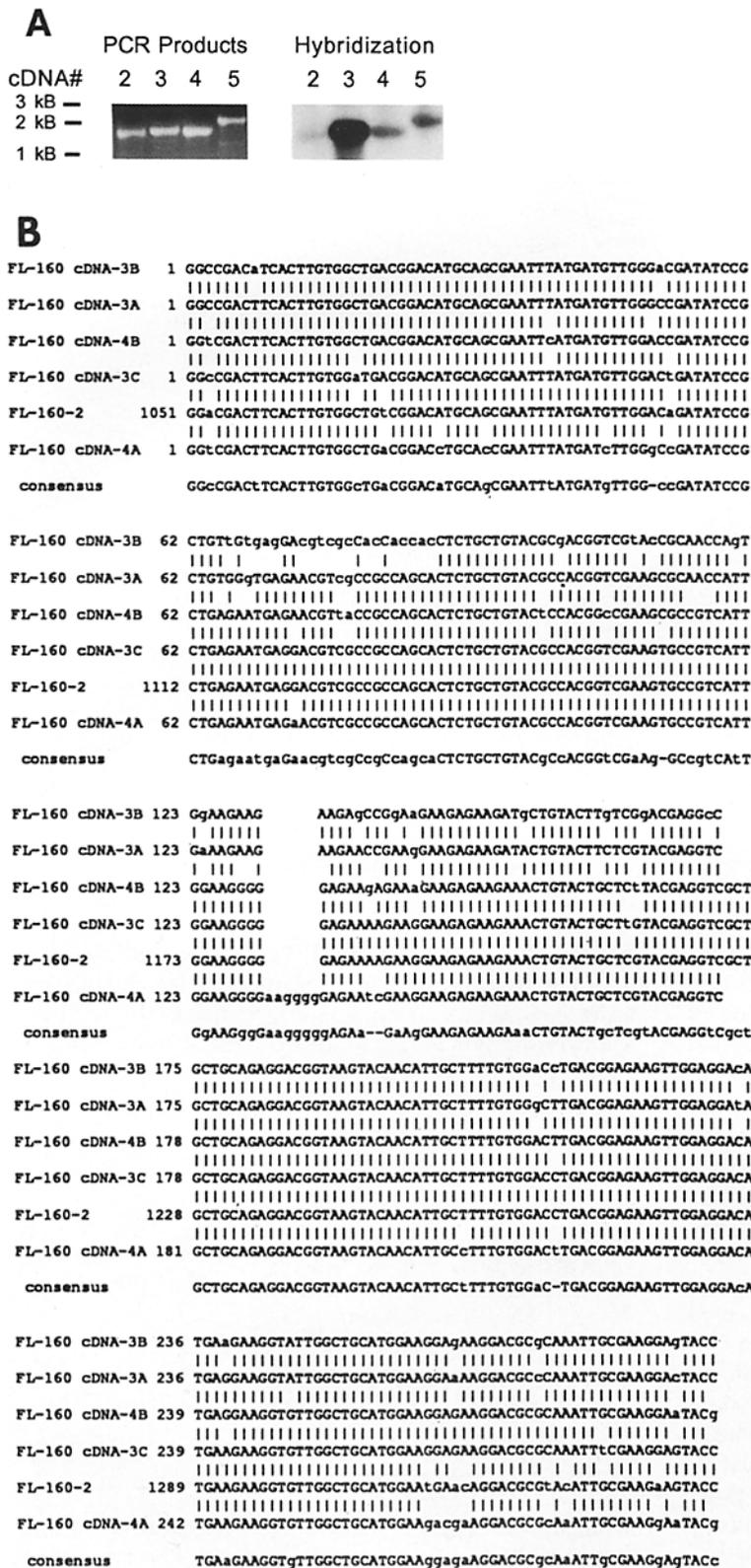
A

PCR Products   Hybridization

cDNA#   2 3 4 5   2 3 4 5
3 kB —
2 kB —
1 kB —

B

```
FL-160 cDNA-3B     1 GGCCGACaTCACTTGTGGCTGACGGACATGCAGCGAATTTATGATGTTGGGaCGATATCCG
                     IIIIIII IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII IIIIIIIIII
FL-160 cDNA-3A     1 GGCCGACTTCACTTGTGGCTGACGGACATGCAGCGAATTTATGATGTTGGGCCGATATCCG
                     II IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII IIIIIIIIII IIIIIIIIII
FL-160 cDNA-4B     1 GGtCGACTTCACTTGTGGCTGACGGACATGCAGCGAATTcATGATGTTGGACCGATATCCG
                     II IIIIIIIIIIIIIII IIIIIIIIIIIIIIIIIIIII IIIIIIIIII IIIIIIIII
FL-160 cDNA-3C     1 GGcCGACTTCACTTGTGGaTGACGGACATGCAGCGAATTTATGATGTTGGACtGATATCCG
                     II IIIIIIIIIIIIIIII IIIIIIIIIIIIIIIIIIIIIIIIIIIIII IIII I IIIIIIIIII
FL-160-2        1051 GGaCGACTTCACTTGTGGCTGtCGGACATGCAGCGAATTTATGATGTTGGACaGATATCCG
                     II IIIIIIIIIIIIIIIII IIIIII IIIII IIIIIIIIIIIII IIII I IIIIIIIII
FL-160 cDNA-4A     1 GGtCGACTTCACTTGTGGCTGaCGGACcTGCAcCGAATTTATGATcTTGGgCcGATATCCG

consensus            GGcCGACtTCACTTGTGGcTGaCGGACaTGCAgCGAATTtATGATgTTGG-ccGATATCCG
```

```
FL-160 cDNA-3B    62 CTGTtGtgagGAcgtcgcCacCaccacCTCTGCTGTACGCgACGGTCGtAcCGCAACCAgT
                     IIII III  I    II      I I  IIIIIIIIIIIII IIIIIIII I IIIIIIII I
FL-160 cDNA-3A    62 CTGTGGgTGAGAACGTcgCCGCCAGCACTCTGCTGTACGCCACGGTCGAAGCGCAACCATT
                     III I IIIIIIIIII IIIIIIIIIIIIIIIIIIIIII IIIIIII IIIIIIIII  IIII
FL-160 cDNA-4B    62 CTGAGAATGAGAACGTtaCCGCCAGCACTCTGCTGTACtCCACGGcCGAAGCGCCGTCATT
                     IIIIIIIIII IIII  IIIIIIIIIIIIIIIIIIIIIIIII IIIIII  III IIIIIIIII
FL-160 cDNA-3C    62 CTGAGAATGAGGACGTCGCCGCCAGCACTCTGCTGTACGCCACGGTCGAAGTGCCGTCATT
                     IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
FL-160-2        1112 CTGAGAATGAGGACGTCGCCGCCAGCACTCTGCTGTACGCCACGGTCGAAGTGCCGTCATT
                     IIIIIIIIIII IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
FL-160 cDNA-4A    62 CTGAGAATGAGaCGTCGCCGCCAGCACTCTGCTGTACGCCACGGTCGAAGTGCCGTCATT

consensus            CTGagaatgaGaacgtcgCcgCcagcaCTCTGCTGTACGCcACGGtCGaAg-GCcgtCAtT
```

```
FL-160 cDNA-3B   123 GgAAGAAG     AAGAgCCGgAaGAAGAGAAGATgCTGTACTTgTCGgACGAGGcC
                     I IIIIII     IIII III I IIIIIIIIIII IIIIIIIII III IIIIIII I
FL-160 cDNA-3A   123 GaAAGAAG     AAGAACCGAAgGAAGAGAAGATACTGTACTTCTCGTACGAGGTC
                     I IIII I     IIII III IIIIIIIIIIIII IIIIIIII III IIIIIIIIII
FL-160 cDNA-4B   123 GGAAGGGG     GAGAAgAGAAaGAAGAGAAGAAACTGTACTGCTCtTACGAGGTCGCT
                     IIIIIIII     IIIII IIII IIIIIIIIIIIIIIIIIIIIIIIII III IIIIIIIIII
FL-160 cDNA-3C   123 GGAAGGGG     GAGAAAAGAAGGAAGAGAAGAAACTGTACTGCTtGTACGAGGTCGCT
                     IIIIIIII     IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIH III IIIIIIIIII
FL-160-2        1173 GGAAGGGG     GAGAAAAGAAGGAAGAGAAGAAACTGTACTGCTCGTACGAGGTCGCT
                     IIIIIIII     IIIII   IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
FL-160 cDNA-4A   123 GGAAGGGGaaggggGAGAAtcGAAGGAAGAGAAGAAACTGTACTGCTCGTACGAGGTC

consensus            GgAAGGgGaagggggAGAa--GaAgGAAGAGAAGAaaCTGTACTgcTcgtACGAGGtCgct
```

```
FL-160 cDNA-3B   175 GCTGCAGAGGACGGTAAGTACAACATTGCTTTTGTGGaccTGACGGAGAAGTTGGAGGAcA
                     IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII I IIIIIIIIIIIIIIIIIIIII
FL-160 cDNA-3A   175 GCTGCAGAGGACGGTAAGTACAACATTGCTTTTGTGGgCTTGACGGAGAAGTTGGAGGAtA I
                     IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
FL-160 cDNA-4B   178 GCTGCAGAGGACGGTAAGTACAACATTGCTTTTGTGGACTTGACGGAGAAGTTGGAGGACA
                     IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
FL-160 cDNA-3C   178 GCTGCAGAGGACGGTAAGTACAACATTGCTTTTGTGGACCTGACGGAGAAGTTGGAGGACA
                     IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
FL-160-2        1228 GCTGCAGAGGACGGTAAGTACAACATTGCTTTTGTGGACCTGACGGAGAAGTTGGAGGACA
                     IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII IIIIIIIII IIIIIIIIIIIIIIII
FL-160 cDNA-4A   181 GCTGCAGAGGACGGTAAGTACAACATTGCtTTGTGGACtTGACGGAGAAGTTGGAGGACA

consensus            GCTGCAGAGGACGGTAAGTACAACATTGCttTTGTGGaC-TGACGGAGAAGTTGGAGGAcA
```

```
FL-160 cDNA-3B   236 TGAaGAAGGTATTGGCTGCATGGAAGGAgAAGGACGCgCAAATTGCGAAGGAgTACC
                     III IIIIIIIIIIIIIIIIIIIIIIIIII III IIIIIIIII IIIIIIIIIIIIIIII
FL-160 cDNA-3A   236 TGAGGAAGGTATTGGCTGCATGGAAGGAaAAGGACGCcCAAATTGCGAAGGAcTACC
                     IIIIIIIIII IIIIIIIIIIIIIIIII IIIIIIII IIIIIIIIIIIIIIIIII III
FL-160 cDNA-4B   239 TGAGGAAGGTGTTGGCTGCATGGAAGGAGAAGGACGCGCAAATTGCGAAGGAaTACg
                     III IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII IIIIIII III
FL-160 cDNA-3C   239 TGAAGAAGGTGTTGGCTGCATGGAAGGAGAAGGACGCGCAAATTtCGAAGGAGTACC
                     IIIIIIIIIIIIIIIIIIIIIIIIIIIII II IIIIIIII II IIIIIIIIIII III
FL-160-2        1289 TGAAGAAGGTGTTGGCTGCATGGAAtGAacAGGACGCGtAcATTGCGAAGaAGTACC
                     IIIIIIIIIIIIIIIIIIIIIIIIII  IIIIIIIII I IIIIIIIIII I III
FL-160 cDNA-4A   242 TGAAGAAGGTGTTGGCTGCATGGAAgacgaaGGACGCGcAaATTGCGAAGgAaTACg

consensus            TGAaGAAGGTgTTGGCTGCATGGAAggagaAGGACGCgcAaATTgCGAAGgAgTACc
```

Figure 5. Reverse transcription and sequence analysis demonstrates transcription of at least five members of FL-160. (*A*) (*Left*) Ethidium bromide-stained cDNA PCR products run on a 1% agarose TBE gel. RNA from trypomastigotes was reverse transcribed using a primer homologous to the 3' end of the FL-160 genes. The cDNA was then amplified using PCR. The 5' sense primer in each case was identical to the first 21 bases of the spliced leader sequence. Since the spliced leader sequence is trans-spliced onto mRNA in a post-transcriptional process, use of this primer should allow only amplification of mature mRNAs. The 3' primers were: FL-160-2 AS 1324 for cDNA 2, FL-160-2 AS 1381 for cDNA 3, FL-160-4 AS 297 for cDNA 4, and FL160-2 AS 1926 for cDNA 5. Primer FL-160-4 AS 297 (*PCR cDNA product 3*) is homologous only to the FL-160-4 gene, whereas primer FL-160-2 AS 1381 (*PCR cDNA product 4*) is not homologous to FL-160-4, but is homologous to FL-160-2 and FL-160-3. This demonstrates that each oligonucleotide is able to prime a PCR product from mRNA, and two of these oligonucleotides are specific for different mRNAs. (*Right*) The same products transferred to a nylon membrane and hybridized with a labeled oligonucleotide (FL-160-2 AS 911) homologous to the FL-160 genes within the region amplified. That binding occurs to each of the four PCR products confirms that the PCR products are all members of the FL-160 family. That fragments 3 and 4 are the same size and hybridize to the FL-160 internal probe is consistent with the hypothesis that at least two different FL-160 genes are transcribed into mRNA. (*B*) The partial sequence of three subclones of FL-160 cDNA 3 (A–C) and 2 subclones of FL-160 4 (A and B) are aligned with the sequence of FL-160-2 genomic clone. (*Uppercase letters*) Homologous bases with sequence above; and (*lowercase letters*) nonhomologous. A consensus sequence is shown below each alignment. (*Uppercase letters*) Identical bases at that position for all six members. None of the six sequences are identical and none of the cDNAs are completely homologous to the FL-160-3 or FL-160-4 (data not shown). More than 80% homology between the various members of the FL-160 family are observed and an ORF with more than 80% predicted amino acid identity with FL-160 is preserved in each cDNA. The nucleotide sequence data has been submitted to EMBL/GenBank/DDBJ under the accession numbers X71867, X71868, X71869, X71870, and X71871.

gives a strong band at ~1 kb. A more complex pattern of several PstI bands at 3–8 kb and a minor band at 1.8 kb is seen with 3' end probes. A wide band is found at 4–6 kb with BamHI (data not shown) and BamHI + ApaI digestion (Fig. 4 a), whether the probe is from the 3' or the 5' end of the molecule. This is probably due to conservation of sequence among the various members of the family. This sequence conservation probably occurs in the noncoding regions because our full-length clone is bounded by BamHI and ApaI sites and none of our clones have a BamHI site in the coding sequence. The FL-160 genes do not seem to be arranged in tandemly repeating arrays of genes, as partial digestion with BamHI does not give hybridizing bands at 4, 8, 12, 16 kb etc. (data not shown). Definitive proof that the genes are not arranged in tandem arrays awaits cloning large pieces of DNA encoding FL-160 genes and chromosomal localization.

More than One Member of the FL-160 Gene Family Is Transcribed. Fig. 4 b shows a Northern blot of *T. cruzi* trypomastigote total RNA hybridized to several fragments of the FL-160 gene. Both the 5' (derived from FL-160-3 gene) and the 3' fragments (derived from the FL-160-1 gene) recognize an ~4–5 kb transcript. None of the probes hybridized to epimastigote RNA (data not shown).

T. cruzi and other kinetoplasts are unique in that they have a 39 base spliced leader sequence added by a trans-splicing mechanism in a posttranscriptional process to the 5' end of the mature mRNA (23). This spliced leader can be used to obtain mature mRNA. The spliced leader sequence was used as a 5' sense primer in a PCR reaction using cDNA for template. First, cDNA was made using an antisense primer that is complementary to all of the sequenced FL-160 genes near the 3' end of the coding sequence (FL-160-1 AS 488 [5]). This cDNA was then PCR amplified using the sense spliced leader oligonucleotide and an antisense oligonucleotide predicted from the FL-160 genes.

A short cDNA transcript (FL-160 cDNA 1, 642 bp) was cloned and sequenced (Figs. 2 and 3 A). As expected, the first 11 bp of the 5' end were identical to the spliced leader sequence found on *T. cruzi* mRNA (23). Furthermore, the sequence was 79% homologous to the FL-160-2 genomic clone. An ATG exists at the identical position as the predicted start of FL-160-2 (Fig. 2).

Two antisense primers, each of which binds only a subset of FL-160 genes, were used to demonstrate that at least two subsets of genes are being transcribed. Fig. 5 A shows cDNA-driven PCR products that are made using antisense oligonucleotides homologous to the ORF of FL-160 genes combined with the sense oligonucleotide homologous to the spliced leader sequence. These cDNAs are depicted schematically in Fig. 3 A. Oligonucleotides were synthesized homologous to the region of FL-160-2 bases 1362–1381 and FL-160-4 bases 278–297 (Fig. 2). These regions had >60% difference between FL-160-2 and FL-160-4. Antisense oligonucleotide FL-160-2 AS 1381, which only binds cDNA homologous to FL-160-2, but not FL-160-4, was used to amplify cDNA 3. Antisense oligonucleotide FL-160-4 AS 297, which only binds cDNA homologous to FL-160-4, but not FL-160-2 was used

to amplify cDNA 4. Note that the products are the same length, about 1,700 bp (Fig. 5), as predicted from the antisense primers binding the same region.

To demonstrate that PCR products cDNA 3 and 4 were from the expected region of FL-160 genes, PCR cDNA products were made that were expected to be shorter (cDNA 2) or longer (cDNA 5) than PCR products cDNA 3 and 4. Antisense primers homologous to all of the FL-160 genes were used to generate cDNA PCR products 2 and 5: FL-160-2 AS 1324 for cDNA 2, and FL-160-2 AS 1926 for cDNA 5. Products of the expected molecular weight are seen in Fig. 5 A, and a diagram of the expected result is shown in Fig. 3 A. The location of the spliced leader in the cDNA 1 corresponds to that of cDNAs 2–5, taking into account the expected differences in length giving the homologous regions of the oligonucleotides used (shown diagrammatically in Fig. 3 A). This supports the idea that in each case the sense spliced leader oligonucleotide and the antisense primers gave the cDNA corresponding to the expected products.

Fig. 3 shows that all PCR products hybridize with a common oligonucleotide expected to bind to all members of the FL-160 gene family (FL-160-2 AS 911). There was no apparent binding to irrelevant DNA such as pBS (data not shown).

FL-160 cDNA PCR products 3 and 4 were directly sequenced and found to represent multiple alleles, that is, in more than 10% of positions, multiple nucleotides were found in sequence analysis. When sequencing from the 5' end, the first 11 bp were homologous to the spliced leader sequence, demonstrating that the PCR products were derived from reverse transcription of mRNA and not contaminating genomic DNA. The FL-160 cDNA PCR products 3 and 4 were subcloned and five of these were sequenced as demonstrated in Fig. 5 B. Each of the five subclones is unique in sequence but has an ORF with about 80% homology to FL-160-2. Note that in comparison with the other sequences: an insertion of 6 bp occurs in FL-160 cDNA 4A; a deletion of 3 bp occurs in FL-160 cDNAs 3A, 3B, and 4A; and a 24-bp region of only 22% homology to the other cDNAs occurs in FL-160 cDNA 3B.

Taken together, the results demonstrate that mRNA homologous with unique regions from at least five FL-160 genes are synthesized. Thus, multiple genes appear to be transcribed.

Quantitative Hybridization Suggests that 750 Copies of FL-160 Genes Are Present in the T. cruzi Genome. Fig. 6 shows an autoradiogram of hybridization of the 5' and the two 3' probes to slot blots of *T. cruzi* genomic DNA and to FL-160-3 plasmid. These probes were chosen because they had <50% homology to *T. cruzi* sequences other than those in the FL-160 family. This blot was quantitatively scanned using phosphoimaging analysis (Fig. 6). The dose–response of both the FL-160-3 plasmid and the genomic *T. cruzi* DNA gave a linear increase in cpm with increased DNA except at the highest dose of DNA used. Since the number of plasmids and genome equivalents of DNA applied to the filter was known, an estimate of the number of copies of the gene in the *T. cruzi* genome was possible. The pBS-FL-160-3 plasmid is ~7,000 bp, which means each nanogram of DNA represents

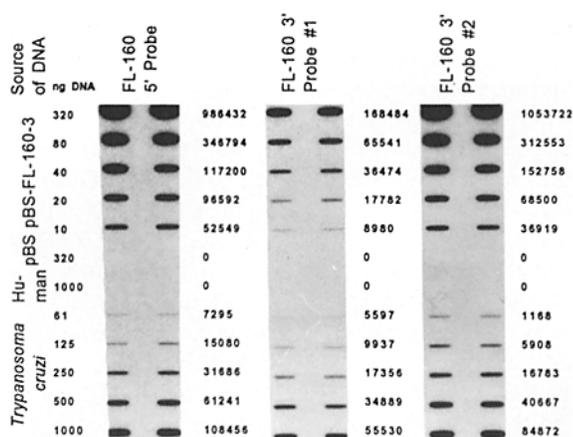| Source of DNA	ng DNA	FL-160 5' Probe		FL-160 3' Probe #1		FL-160 3' Probe #2	
Human pBS pBS-FL-160-3	320		986432		168484		1053722
	80		346794		65541		312553
	40		117200		36474		152758
	20		96592		17782		68500
	10		52549		8980		36919
	320		0		0		0
	1000		0		0		0
Trypanosoma cruzi	61		7295		5597		1168
	125		15080		9937		5908
	250		31686		17356		16783
	500		61241		34889		40667
	1000		108456		55530		84872

Figure 6. Quantitative hybridization to *T. cruzi* genomic DNA and FL-160-3 plasmid DNA suggests ∼750 copies of FL-160 are present in the genome. An autoradiograph of slot blot hybridization of the FL-160 5' probe and the two 3' probes to duplicate quantities of *T. cruzi* genomic DNA (*1,000–61 ng/slot*), human DNA (*1,000 ng/slot*), pBS (*320 ng/slot*), and pBS-FL-160-3 plasmid DNA (*10–320 ng/slot*). The radioactivity per sample was quantitated using phosphoimager analysis. The values for human DNA were subtracted from the *T. cruzi* values and the values for pBS were subtracted from the pBS-pBS-FL-160-3 values. The mean of the two values is shown under each pair of slots. Using simple linear regression analysis of these values and the estimate that there is 160 fg of DNA/*T. cruzi* cell, gives values of 752 genes/cell for the 5' probe, 720 genes/cell for the 3' probe 2, and 2,374 genes/cell for the 3' probe 1. In experiments not shown, using an identical analysis for the CUB gene of *T. cruzi*, 2.4 genes/cell were observed. Work by Ajioka and Swindle (11) had shown that two copies of the CUB gene/cell are present, demonstrating the ability of the quantitative hybridization technique to estimate the number of genes within an 80% accuracy for the CUB gene.

1.25 × 10⁸ plasmids. Simple linear regression analysis was performed using the values for the plasmid hybridization and the plasmid equivalents applied. The R² of the linear regression analysis for each of the probes was >0.99, demonstrating a linear relationship. The values obtained by hybridization to *T. cruzi* DNA within the values in the linear range were then entered into the linear equation, and the number of genes present estimated. Each *T. cruzi* cell has ∼160 femtogram (fg) of DNA (24) which means that each microgram of DNA represents 6.3 × 10⁶ genome equivalents. The number of gene equivalents determined by hybridization was divided by the genome equivalents of *T. cruzi* DNA. A mean of 752 estimated copies per cell was found for the 5' probe (range 732–768) and a mean of 730 was found for the 3' probe 2 (range 718–742). The 3' probe 1 gave a mean of 2,374 copies per cell (range 2076–2554), suggesting that this probe identifies more members of the family or superfamilies relating to FL-160. Similar values were obtained when the experiment was repeated (data not shown). As a control, an identical analysis was performed using a probe and a reference plasmid containing the CUB gene, previously determined to have two copies per *T. cruzi* genome (11). Again, a linear relationship was found with the quantity of radioactivity bound with CUB probe hybridized to *T. cruzi* genomic DNA (0.5–1.5 μg) and to CUB reference plasmid (25–100 pg) (data not shown). A

mean of 2.4 copies per cell was found for the CUB probe (range 2.04–2.58), demonstrating that the quantitative hybridization technique gave a result within 80% of the expected number of genes for a two copy per cell gene. Taken together, these experiments suggest the FL-160 genes are part of a 750-copy multigene family.

Abs to the COOH and the NH₂ Termini of FL-160 Both Recognize a 160-kD Flagellar-pocket Surface Protein, but Each Binds a Different Cross-reactive Epitope of Mammalian Nervous Tissues. Abs were prepared to the COOH terminus of the gene as previously published (4–5) and also to the NH₂ terminus of the gene. The NH₂-terminal recombinant protein was made by PCR amplification of the 5' end of FL-160-3 and splicing the product into pGEX-2 (9). The protein was expressed as a recombinant fusion protein with *Schistosoma japonicum* GST. Rabbits were immunized and the Abs were affinity purified. Fig. 7 demonstrates affinity-purified Abs directed to either the COOH region (the FL-160-1 protein product) or the NH₂ terminus recognized a 160-kD protein on Western blots of trypomastigotes. Both Abs recognized a protein distributed along the flagellum of the live trypomastigotes by immunofluorescence (Fig. 8, *b* and *c*) or immunoelectron microscopy (Fig. 8, *e* and *f*). If trypanosomes were fixed and treated with 0.1% Tween detergent before staining, more diffuse cytoplasmic staining was noted in trypomastigotes and amastigotes (data not shown). This suggests that anti-FL-160 Abs react with trypomastigote flagellar membranes, but that the antigen is more evenly distributed in permeabilized trypomastigotes and amastigotes. Flow cytometry analysis demonstrated that anti-FL-160 Abs bind live trypomastigotes (Fig. 9). A monomorphic curve of fluorescence was observed, suggesting that all trypomastigotes bind both anti-NH₂ and anti-COOH-terminal FL-160 Abs on their surface. The Ab binding demonstrated by flow cytometry was abrogated by excess recombinant fusion protein but not GST recombinant protein, showing antigen specificity.
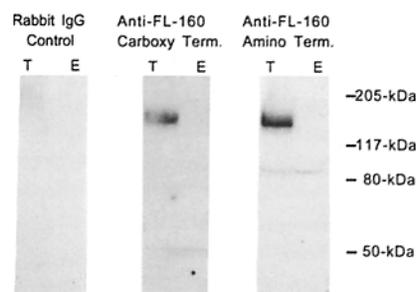
Rabbit IgG Control		Anti-FL-160 Carboxy Term.		Anti-FL-160 Amino Term.	
T	E	T	E	T	E

–205-kDa
–117-kDa
– 80-kDa
– 50-kDa

Figure 7. Western blots demonstrate antibodies to the 5' end and 3' end of FL-160 bind to a 160-kD protein of mammalian forms of *T. cruzi*. An autoradiograph of a Western blot. Each lane has 5 × 10⁶ parasite equivalents of lysates, run on a 10% SDS-PAGE gel, and transferred to nitrocellulose. (*T*) Lanes have trypomastigote lysates; (*E*) lanes have epimastigote lysates. Lanes were reacted with 20 μg/ml of affinity-purified anti-FL-160 Ab (directed either to the NH₂ or COOH terminus) or nonimmune rabbit IgG, and then ¹²⁵I-protein A. Autoradiography was for 18 h for the anti-FL-160 COOH-terminal lanes or 120 h for the anti-FL-160 NH₂ terminal and rabbit IgG control lanes.
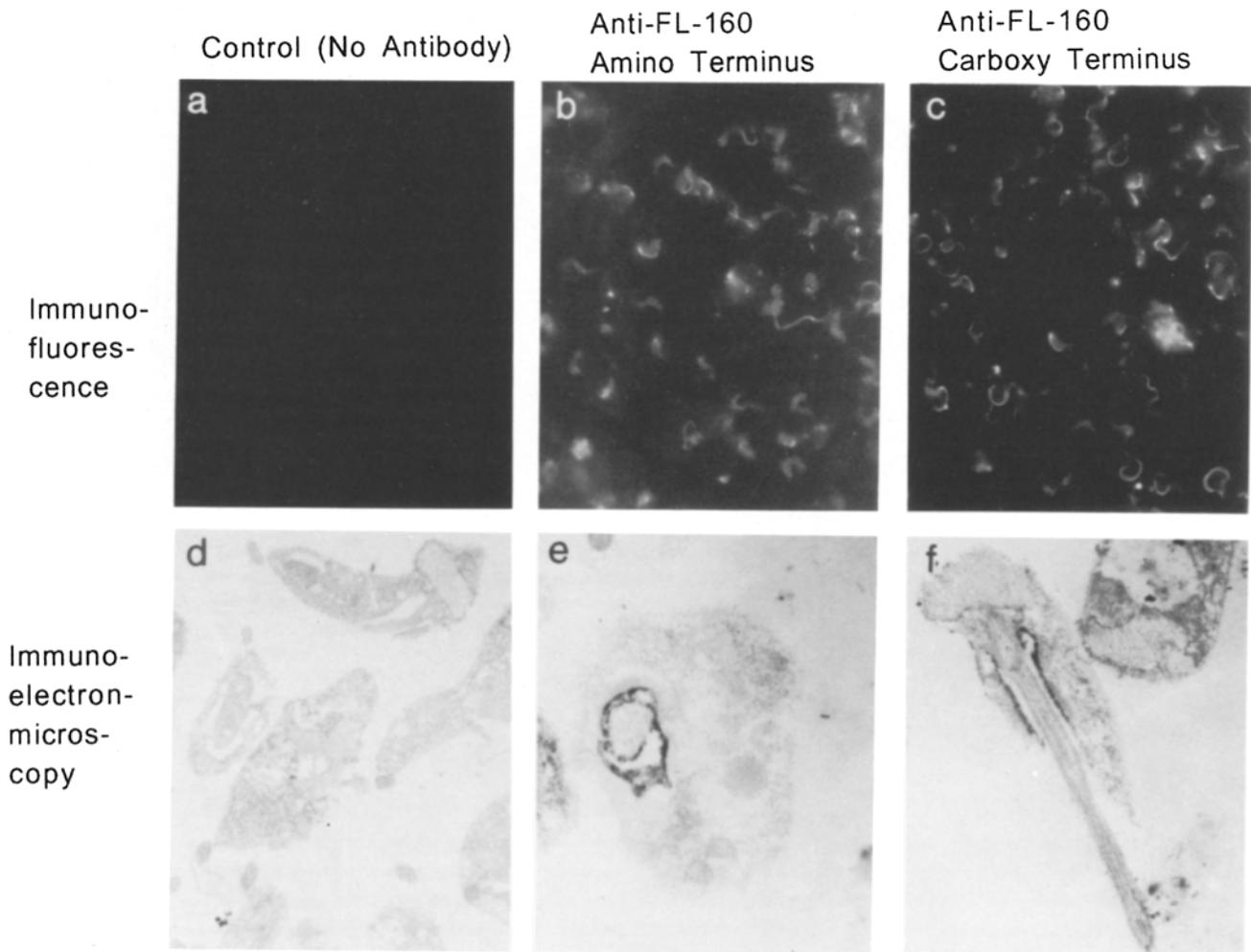
Control (No Antibody)   Anti-FL-160 Amino Terminus   Anti-FL-160 Carboxy Terminus

Immuno-fluores-cence

Immuno-electron-micros-copy

Figure 8. Immunofluorescence and immunoelectron microscopy demonstrate anti-FL-160 antibody binding limited to the flagellar pocket of trypomastigotes. (a–c) Immunofluorescent-labeled trypomastigotes; (d–f) immunoperoxidase-labeled trypomastigotes viewed by transmission electron microscopy. (a and d) Controls without Ab; (b and e) reacted with anti-NH_2-terminal FL-160 Ab; (c and f) reacted with anti-COOH-terminal FL-160 Ab. In experiments not shown, neither nonimmune rabbit IgG nor anti-GST rabbit IgG reacted with the parasites. (a–c) ×670; (d) ×11,200; (e) ×27,600; and (f) ×16,700.
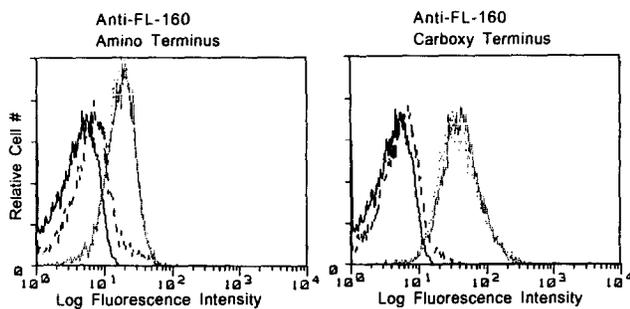
Figure 9. Flow cytometry demonstrates staining of live parasites by Abs to the NH_2 and COOH termini of FL-160. Two tracings of relative cell number (linear scale) and fluorescence intensity (log scale) are shown after indirect immunofluorescence was performed on live trypomastigotes. (*Left*) Tracing shows trypomastigotes reacted with: only fluoresceine-conjugated anti-rabbit Ab (FCARA) (_____); anti-NH_2 terminal FL-160 Ab and then FCARA (. . . . .); anti-NH_2 terminal FL-160 Ab plus 250 μg/ml FL-160-3 NH_2-terminal GST fusion protein and then FCARA (----); anti-

Frozen sections of murine sciatic nerve were exposed to affinity-purified anti-FL-160 Abs and binding demonstrated by immunoperoxidase. Abs to the NH_2 terminus of FL-160 react with epineurium and occasional linear densities within the nerve (Fig. 10). In contrast, anti-FL-160 COOH-terminal Abs do not bind to epineurium, but rather bind a much higher number of linear elements within peripheral nerve, which probably correspond to peripheral nerve axons (Fig. 10). Thus, the epitopes the anti-FL-160 NH_2 and COOH-terminal Abs recognize are distinct. This experiment has been repeated four

NH_2-terminal FL-160 Ab plus GST and then FCARA (····). (*Right*) tracing shows trypomastigotes reacted with: only FCARA (_____); anti-COOH-terminal FL-160 Ab and then FCARA (. . . . .); anti-NH_2-terminal FL-160 Ab plus 250 μg/ml FL-160-1 COOH-terminal-GST transferase fusion protein and then FCARA (----); anti-NH_2-terminal FL-160 Ab plus 250 μg/ml GST and then FCARA (····).

691   Van Voorhis et al.

| Control | Anti-FL-160 Carboxy Term. | Anti-FL-160 Amino Term. |
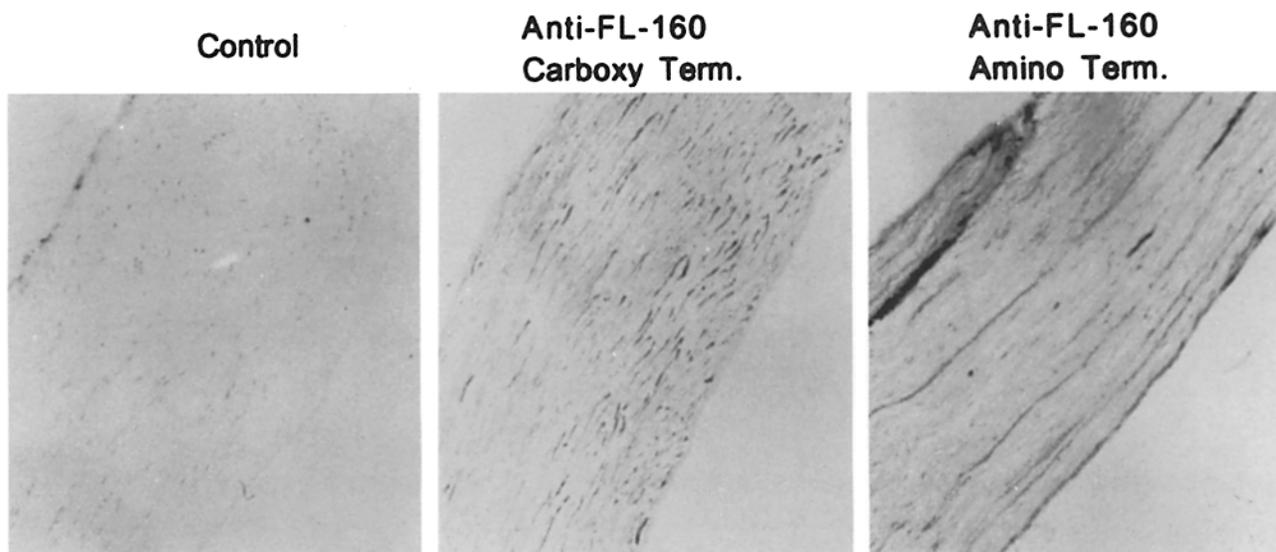
Figure 10. Abs directed to the NH₂ terminus of FL-160 bind to distinct structures of sciatic nerve compared with Abs directed to the COOH terminus of FL-160. Immunoperoxidase-stained sciatic nerve frozen sections counterstained with methyl green are shown. (*Center*) Exposed to affinity-purified anti-COOH-terminal FL-160 Abs, and developed with biotinylated anti–rabbit Ab, avidin-peroxidase, and diaminobenzidine. Note the multiple streaks throughout the longitudinal section of the nerve in the same pattern as axons. (*Right*) Exposed first to anti-NH₂-terminal FL-160 Abs and developed in the same way. Note the strong reaction with the epineurium and the scattered linear staining within the nerve. The control (*left*) is mouse nerve that was not exposed to rabbit Abs, but was otherwise treated the same as the other sections. Note that little staining is seen. A pattern identical to this was seen if the Abs were reacted with an excess of their immunizing recombinant protein, or if anti-SA85 1.1 is used. ×232.

times, and an identical result obtained each time, i.e., the staining of sciatic nerve is different with the anti-FL-160 COOH and NH₂-terminal Abs. Abs to the SA85 1.1 surface protein (13, 16), provided by Stuart Kahn, did not bind murine sciatic nerve (data not shown).

In experiments not shown, recombinant fusion protein of the FL-160 NH₂ terminus-GST eliminates the staining of murine sciatic nerve with anti-NH₂-terminal FL-160 Abs. Recombinant FL-160 COOH terminus-GST or GST protein do not inhibit the nerve staining of anti-NH₂-terminal FL-160 Abs. Thus the binding of anti-NH₂-terminal FL-160 Abs to nerve can be competed only by recombinant proteins containing the NH₂ terminus of FL-160.

Discussion

These findings suggest that FL-160, the flagellum-associated 160-kD protein of *T. cruzi* trypomastigotes, is expressed from a family of about 750 genes and or pseudogenes. PCR-mediated cDNA synthesis with two oligonucleotides specific for different genes shows that at least five copies of the genes are transcribed. Abs, directed to two different regions of the FL-160 protein, bind to a 160-kD protein on the outer surface of the cell membrane in the flagellar pocket of trypomastigotes. Finally, both the NH₂ and COOH terminus of FL-160 protein have distinct molecular mimicry epitopes that are homologous to mammalian nervous tissues.

The predicted amino acid sequence of FL-160-2 has recognizable domains. The NH₂ terminus has a putative signal sequence (25) followed by a region with 27% homology to

bacterial sialidase. The next region is homologous to the fibronectin type III domain. The COOH terminus lacks the amino acid repeats noted in many other surface antigens of *T. cruzi* but has the 12-amino acid molecular mimicry epitope. Finally, the COOH terminus has a putative signal for addition of a glycosyl-phosphotidylinositol anchor (26, 27).

The genes of the FL-160 family are highly related to one another (>80% homology of amino acid sequence among sequenced genes and mRNAs reported here). The FL-160 genes are also highly related to several *T. cruzi* genes deposited in the DNA databases. The similarity of the Tc exoant (unpublished sequence in EMBL, see above) sequence to FL-160-2 (Fig. 3; 57% homology of the amino acid sequence) suggests it is a member of the same family as FL-160 and alignment to FL-160-2 suggests Tc exoant is missing 47 amino acids from the COOH terminus. Pseudogenes gp85B (c1821) (15) and gp85A (ORF 1) (15) are also highly related to FL-160-2 (Fig. 3, 40 and 48% homology of the amino acid sequences, respectively). The FL-160 genes reported here are cloned from a different strain than Tc exoant, gp85A, or gp85B and some of the differences in sequence could be attributed to differences between different strains of parasites. The gp85A and gp85B sequences are designated pseudogenes because no initiation codon could be identified and no homologous mRNA could be identified (15). Abs directed to the translated product of the ORF 1 of gp85A bound a 160-kD protein (15), consistent with these pseudogenes being part of the same family as FL-160 genes. Thus, it is possible that some of the 750 copies of the FL-160 family, identified by quantitative hybridization studies, are pseudogenes.

FL-160 is also related to bacterial sialidases and to *T. cruzi* 85-kD surface protein (SA85) and SAPA/neuraminidase/trans-sialidase genes. The SA85 and trans-sialidase families have also been shown to be related to bacterial sialidase genes (15–22). The homology between the FL-160, SA85, and trans-sialidase families are highest in the region of homology to bacterial sialidases. For this reason, we propose that the group of FL-160, SA85, and trans-sialidase genes are a superfamily of genes related to bacterial sialidases as proposed by Campatella et al. (28) for the latter two families.

Abs to the NH_2 or the COOH terminus of the FL-160 protein localize the protein on the surface of the cellular membrane of the flagellar pocket. Western blots with both Abs demonstrate binding to a 160-kD protein of trypomastigotes. Flow cytometry demonstrates that both Abs bind all of the parasites in the population and that this binding is competed away by recombinant FL-160 protein. Ab binding to live parasites demonstrates that the Abs bind to the surface membrane in the flagellar pocket by fluorescence microscopy and ultrastructural localization. These Abs are directed to the recombinant proteins derived from two different FL-160 genes that are ~80% related. These experiments, taken together, demonstrate that the Abs bind to proteins of identical molecular weight and localization. The role of the FL-160 protein in the physiology of the trypomastigote and the flagellar apparatus is unknown at present. One possibility is that the region with sialidase homology is employed for carbohydrate binding and/or cleavage.

Like *Trypanosoma brucei*, *T. cruzi* has large numbers of genes for its mammalian stage surface proteins. *T. brucei* generally expresses only one member of this gene family at a time. Unlike *T. brucei*, *T. cruzi* mammalian forms express many members of its surface proteins at once. The 85-kD surface protein (SA85) family has been shown to be a large gene family of which at least nine members are transcribed and three are translated simultaneously (13). At least five of the FL-160 genes appear to be transcribed. By extrapolation to the results with the SA85 family, we hypothesize that many of the FL-160 gene family are transcribed and translated. If this is true, it would imply that the 160-kD protein that is localized to the outer membrane of the flagellar pocket is actually made up of a number of related proteins. These proteins would

have 10–20% amino acid differences, yet it is interesting that they are retain the molecular mimicry epitope for mammalian nervous tissues found within the COOH terminus of the FL-160 proteins. Whether this sequence is conserved by chance, for a functional purpose, or for immune evasion via molecular mimicry is unknown.

There are many possible reasons for the apparent diversity of expression of SA85 and FL-160 surface proteins. One possibility is that the diversity in surface proteins somehow leads to immune evasion. The various peptides generated by the variations in sequence, bound in MHC molecules, may be responsible for the polyclonal lymphocyte proliferative response observed in acute Chagas' disease (8, 29). The polyclonal lymphocyte proliferative response may be responsible for the general suppression in lymphocyte responses after this proliferative response, and the general suppression in lymphocyte responses may be responsible for the persistence of the infection in the mammalian host. The variations in sequence may also direct the immune response away from functional domains, such as domains responsible for trans-sialidase function. Another possibility is that the SA85 and FL-160 surface proteins are involved with host cell binding or interaction. Diverse binding proteins may be required to bind the wide variety of cells and mammalian hosts that *T. cruzi* infects.

Both the NH_2 and the COOH termini of FL-160 appear to have different molecular mimicry epitopes for mammalian tissues. This is evidenced by the distinct pattern of binding of sciatic nerve by Abs directed to the NH_2 and COOH termini of FL-160. Furthermore, binding by anti-FL-160 NH_2-terminal Abs could be blocked by recombinant protein bearing the NH_2 terminus of FL-160, and not the recombinant COOH terminus of FL-160. This suggests that there are two distinct molecular mimicry epitopes for mammalian nervous tissue on FL-160. In contrast, affinity-purified Abs to *T. cruzi* surface antigen recombinant SA85 1.1 protein did not react with mammalian nervous tissue. This suggests that the molecular mimicry of mammalian nervous tissues by FL-160 may be unique, and not a general property of *T. cruzi* surface antigens. It is possible that the cross-reactive immune response to FL-160 has a role in the pathophysiology of the damage to nervous tissue observed in chronic Chagas' disease.

References

1. World Health Organization. 1991. Control of Chagas' Disease. WHO technical series #811. World Health Organization, Geneva. 93 pp.
2. Tafuri, W.L. 1979. Pathogenesis of *Trypanosoma cruzi* infections. *In* Biology of Kineoplastia Vol. 2. W.H.R. Lumsden and D.A. Evans, editors. Academic Press Limited, London. 547.
3. Eisen, H., K. Petry, and W. Van Voorhis. 1990. The origin of autoimmune pathology associated with *Trypanosoma cruzi* infection. *In* Immune Recognition and Evasion: Molecular Aspects of Host–Parasite Interaction. Academic Press Inc., New York. 91–103.
4. Van Voorhis, W.C., and H. Eisen. 1989. Fl-160: a surface antigen of *Trypanosoma cruzi* that mimics mammalian nervous tissue. *J. Exp. Med.* 169:641.
5. Van Voorhis, W.C., L. Schlekewy, and H. Le Trong. 1991. Molecular mimicry by *Trypanosoma cruzi*: the Fl-160 epitope that mimics mammalian nerve can be mapped to a 12 amino acid peptide. *Proc. Natl. Acad. Sci. USA.* 88:5993.
6. Plata, F., F.G. Pons, and H. Eisen. 1984. Antigenic polymorphism of *Trypanosoma cruzi*: clonal analysis of trypomastigote surface antigens. *Eur. J. Immunol.* 14:392.
7. Tracey, T.E., and L.S. Mulcahy. 1991. A simple method for direct automated sequencing of PCR fragments. *Biotechniques.* 11:68.
8. Van Voorhis, W.C. 1992. Coculture of human peripheral blood mononuclear cells with *Trypanosoma cruzi* leads to proliferation of lymphocytes and cytokine production. *J. Immunol.* 148:239.
9. Smith, D.B., and K.S. Johnson. 1988. Single step purification of polypeptides expressed in *Escherichiae coli* as fusions with glutathionine-S-transferase. *Gene (Amst.).* 67:31.
10. Harlow, E., and D. Lane. 1988. Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 90–91.
11. Ajioka, J., and J. Swindle. 1992. The calmodulin-ubiquitin associated genes of *Trypanosoma cruzi*: their identification and transcription. *Mol. Biochem. Parasitol.* 57:127.
12. Farr, A.G., and P.K. Nakane. 1981. Immunohistochemistry with enzyme-labeled antibodies: a brief review. *J. Immunol. Methods.* 47:129.
13. Kahn, S., W.C. Van Voorhis, and H. Eisen. 1990. The major 85-kDa surface antigen of the mammalian form of *Trypanosoma cruzi* is encoded by a large heterogeneous family of simultaneously expressed genes. *J. Exp. Med.* 172:589.
14. Kozak, M. 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* 12:857.
15. Takle, G.B., J. O'Connor, A.J. Young, and G.A.M. Cross. 1992. Sequence homology and absence of mRNA defines a possible pseudogene member of the gp85/sialidase multigene family.

Mol. Biochem. Parasitol. 56:117.
16. Kahn, S., T.G. Colbert, J.C. Wallace, N.A. Hoagland, and H. Eisen. 1991. The major 85-kDa surface antigen of the mammalian-stage forms of *Trypanosoma cruzi* is a family of sialidases. *Proc. Natl. Acad. Sci. USA.* 88:4481.
17. Fouts, D.L., B.L. Ruef, P.T. Ridley, R.A. Wrightsman, D.S. Peterson, and J.E. Manning. 1991. Nucleotide sequence and transcription of a trypomastigote surface antigen gene of *Trypanosoma cruzi. Mol. Biochem. Parasitol.* 46:189.
18. Takle, G.B., and G.A.M. Cross. 1991. An 85-kilodalton surface antigen gene family of *Trypanosoma cruzi* encodes polypeptides homologous to bacterial neuraminidases. *Mol. Biochem. Parasitol.* 48:185.
19. Pollevick, G.D., J.L. Affrachino, A.A.C. Frasch, and D.O. Sanchez. 1991. The complete sequence of a shed acute-phase antigen of *Trypanosoma cruzi. Mol. Biochem. Parasitol.* 47:247.
20. Pereira, M.E.A., J.S. Mejia, E. Ortega-Barria, D. Matzilevich, and R.P. Prioli. 1991. The *Trypansoma cruzi* neuraminidase contains sequences similar to bacterial neuraminidases, YWTD repeats of the low density lipoprotein receptor, and type III modules of fibronectin. *J. Exp. Med.* 174:179.
21. Uemura, H., S. Schenkman, V. Nussenzweig, and D. Eichinger. 1992. Only some members of a gene family in *Trypanosoma cruzi* encode proteins that express both *trans*-sialidase and neuraminidase activities. *EMBO (Eur. Mol. Biol. Organ.) J.* 11:3837.
22. Hoyer, L.L., A.C. Hamilton, S.M. Steenbergen, and E.R. Vimr. 1992. Cloning and sequencing and distribution of *Salmonella typhimurium* LT2 sialidase gene, *nanH*, provides evidence for interspecies gene transfer. *Mol. Microbiol.* 6:873.
23. De Lange, T., T.M. Berkcvens, H.J.G. Veerman, A.C.C. Frasch, J.D. Berry, and P. Borst. 1984. Comparison of the genes coding for the common 5' terminal sequence in three trypanosomal species. *Nucleic Acids Res.* 12:4431.
24. Kooy, R.F., F. Asnall, M. Van der Ploeg, and J.P. Overdulve. 1989. On the DNA content of *Trypanosoma cruzi. Mol. Biochem. Parasitol.* 36:73.
25. von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 14:4683.
26. Ferguson, M.A.J., and A.F. Williams. 1988. Cell surface anchoring of proteins via anchoring of proteins via glycosylphosphotidylinositol structures. *Annu. Rev. Biochem.* 57:285.
27. Cross, G.A.M. 1990. Glycolipid anchoring of plasma membrane proteins. *Annu. Rev. Cell Biol.* 6:1.
28. Campetella, O., D. Sanchez, J.J. Cazzulo, and A.C.C. Frasch. 1992. A superfamily of *Trypanosoma cruzi* surface antigens. *Parasitol. Today.* 8:378.
29. Minoprio, P.M., H. Eisen, L. Forni, M.R. D'Imperio Lima, M. Joskowicz, and A. Coutinho. 1986. Polyclonal lymphocyte responses to murine *Trypanosoma cruzi* infection: quantitation of both T- and B- cell responses. *Scand. J. Immunol.* 24:661.