

The *Trypanosoma cruzi* Neuraminidase Contains Sequences Similar to Bacterial Neuraminidases, YWTD Repeats of the Low Density Lipoprotein Receptor, and Type III Modules of Fibronectin

By Miercio E. A. Pereira, J. Santiago Mejia, Eduardo Ortega-Barria, David Matzilevich, and Reginaldo P. Prioli

From the Department of Medicine, Division of Geographic Medicine and Infectious Diseases, New England Medical Center Hospitals, Boston, Massachusetts 02111

Summary

Trypanosoma cruzi expresses a developmentally regulated neuraminidase (TCNA) implicated in parasite invasion of cells. We isolated full-length DNA clones encoding TCNA. Sequence analysis demonstrated an open reading frame coding for a polypeptide of 1,162 amino acids. In the N-terminus there is a cysteine-rich domain containing a stretch of 332 amino acids nearly 30% identical to the *Clostridium perfringens* neuraminidase, three repeat motifs highly conserved in bacterial and viral neuraminidases, and two segments with similarity to the YWTD repeats found in the low density lipoprotein (LDL) receptor and in other vertebrate and invertebrate proteins. This domain is connected by a structure characteristic of type III modules of fibronectin to a long terminal repeat (LTR) consisting of 44 full length copies of twelve amino acids rich (75%) in serine, threonine, and proline. LTR is unusual in that it contains at least 117 potential phosphorylation sites. At the extreme C-terminus is a hydrophobic segment of 35 amino acids, which could mediate anchorage of TCNA to membranes via a glycosylphosphatidylinositol linkage. This is the first time a protozoan protein has been found to contain a YWTD repeat and a fibronectin type III module. The domain structure of TCNA suggests that the enzyme may have functions additional to its catalytic activity such as in protein-protein interaction, which could play a role in *T. cruzi* binding to host cells.

Trypanosoma cruzi causes Chagas disease, a chronic debilitating, incurable disease that afflicts millions of people in Latin America. *T. cruzi* exists in three developmental forms: amastigotes and epimastigotes which are multiplying forms, and trypomastigotes, which do not multiply but instead transmit infection from insects to mammals, and vice-versa (1). The molecular mechanism underlying trypomastigote-host cell interaction is beginning to be understood, and it appears that specific recognition molecules mediate *T. cruzi* invasion of cells (2).

One protein suggested to play a role in *T. cruzi*-host cell interaction is the parasite neuraminidase (TCNA)¹ (3, 4). Polyclonal and monoclonal antibodies raised against TCNA enhance infection at concentrations that inhibit enzyme activity (5, 6). Enhancement of infection is probably a consequence of reduced TCNA activity because it can be abrogated

by exogenous neuraminidase (e.g., *Vibrio cholera*), whose activity is not inhibited by the TCNA antibodies (5, 6). *T. cruzi* infection is also augmented by high density lipoprotein (HDL) and low density lipoprotein (LDL), two inhibitors of TCNA activity (7, 8). Enhancement of infection by HDL and LDL is blocked by *V. cholera* neuraminidase, an enzyme that is not inhibited by HDL or LDL. It is not clear how neutralization of endogenous TCNA keeps *T. cruzi* infection in check, but it seems that desialylation removes the receptor for the trypomastigote adhesin (2).

TCNA is developmentally regulated, with maximal activity in trypomastigotes, minimum in epimastigotes, and not detectable in amastigotes (3, 4). The enzyme is located on the surface of the parasite (9, 10) where it can chemically modify, by desialylation, the surfaces of myocardial and vascular endothelial cells (11), and of erythrocytes (12). After transformation from amastigotes intracellularly, all trypomastigotes express neuraminidase (i.e., they are of the NA⁺ phenotype), as they do upon exiting ruptured host cells into the culture medium (13). Once in the extracellular milieu, NA⁺ trypomastigotes differentiate into NA⁻ trypomastigotes (i.e., parasites that do not express enzyme activity) to reach an

¹ Abbreviations used in this paper: BNR, bacterial neuraminidase; EGF, epidermal growth factor; GPI, glycosylphosphatidylinositol; HDL, high density lipoprotein; LDL, low density lipoprotein; LTR, long-terminal repeat; ORF, open reading frame; SAPA, shed acute phase antigen; TCNA, *Trypanosoma cruzi* neuraminidase.

equilibrium in which ~20% of the population is NA⁺ and ~80% is NA⁻ (3).

TCNA is highly polymorphic. The enzyme from metacyclic trypomastigotes has an apparent molecular weight of 66,000 (14), as does the NA from epimastigotes (Rosenberg, I., R.P. Prioli, and M.E.A. Pereira, unpublished observation), and, as assessed by immunoblotting using mono- and polyclonal antibodies, the TCNA from tissue culture trypomastigotes consists of a set of apparently glycosylated polypeptides whose range in size depends on the strain or clone analyzed (6). For example, in the Silvio X-10/4 clone, the enzyme is comprised of bands of apparent mol wt 120,000, 160,000, and a 200,000 doublet; and in the Y-H6 and MV-13 clones, TCNA is made up of bands ranging from 130,000 to 222,000. Under nondenaturing conditions, these polypeptides are held together as trimers by noncovalent bonds (Prioli, R.P., J.S. Mejia, and M.E.A. Pereira, manuscript in preparation).

The primary structure of TCNA is unknown. Here we report the isolation and sequencing of the DNA encoding one of the TCNA isoforms from tissue culture trypomastigotes. The deduced amino acid sequence reveals a catalytic domain in the N-terminus, which is structurally similar to bacterial neuraminidases, in particular the *Clostridium perfringens* enzyme, that contains two YWTD sequence motifs; a domain similar to fibronectin III modules containing GTP-binding consensus sequences; a long terminal tandem repeating structure rich (75%) in serine, threonine and proline residues; and, a hydrophobic stretch of 35 amino acids at the extreme C-terminus which could mediate anchorage of TCNA to the cell surface by a glycosylphosphatidylinositol (GPI) linkage.

Materials and Methods

Parasites. All studies were performed with the cloned Silvio X-10/4 and Montalvania-13 strains of *T. cruzi* (4). Epimastigotes were grown in liver-infusion medium containing 10% FCS, and trypomastigotes were propagated in Vero cells as previously described (4, 11).

Antibodies. TCNA mAbs TCN-1 and TCN-2 were isolated from mice infected with a clone of Montalvania-13 that expresses high enzyme activity (4). These antibodies inhibit neuraminidase activity and deplete TCNA from trypomastigote lysates after addition of protein A-Sepharose, and are specific for the TCNA from tissue culture trypomastigotes, as they do not react with the enzyme from epimastigotes.

Antibodies against 7F, a recombinant TCNA clone, were obtained as a modification of our previously published procedure (5). Bacterial lysates containing recombinant TCNA from 7F insert (see below) were run on a 10% SDS-PAGE, the lanes in the right and left corners cut out with a scalpel, blotted on nitrocellulose and probed with TCN-2 to visualize the relative position of the TCN-2 immunoreactive protein. The blots were then aligned with the polyacrylamide gel, and the portion of the gel corresponding to the immunoreactive band cut out, frozen at -70°C, lyophilized, and injected weekly into mice with Freund's adjuvant. Three injections of gel slices corresponding to those shown in Fig. 2 (lane 8 from left) sufficed to elicit antibodies that reacted with endogenous TCNA (Fig. 2, lane 4 from left). Similar results were obtained by eliciting antibodies against the recombinant TCNA from clone 10C.

Immunoblots. Endogenous TCNA was obtained by lysis of trypomastigotes and epimastigotes in 0.01 M PBS, pH 7.0, containing 1% Triton X-100 and protease inhibitors (5 μM pepstatin and leupeptin, 10 mM EDTA, 5 mM iodoacetamide, and 10 μg/ml soybean trypsin inhibitor), centrifuged at 1,000 g to remove nuclei and debris, and cleared at 100,000 g for 1 h. Amastigotes and epimastigotes were treated similarly as trypomastigotes. Such lysates (~10 μg/lane) were run on 10% SDS-PAGE, blotted to nitrocellulose, and allowed to react with TCN-2 or 7F antibodies, followed by an anti-mouse antibody conjugated to alkaline phosphatase, as previously described (6).

Recombinant TCNA was obtained from *Escherichia coli* XL-1 Blue that had been transformed with pBluescript containing TCNA inserts (7F, 10C, 16C, 19Y, and 20G), or without inserts. Bacteria were grown to an OD₅₅₀ = 1.0, washed once with PBS, lysed by freezing in liquid nitrogen in the presence of protease inhibitors (see above) and thawing at 37°C for 15 min, and centrifuged in a microfuge for 10 min. The supernatant was immediately analyzed by immunoblotting, sieving chromatography, or neuraminidase activity (see below). Preliminary experiments showed most recombinant TCNA to be present in supernatants of the lysed bacteria and very little in bacterial pellets. It should be emphasized that recombinant TCNA is highly susceptible to degradation, and only fresh samples were analyzed. Even storage of the enzyme in crude lysates at -20°C for several days gives rise to low mol wt bands in the immunoblots and to reduced enzyme activity.

Sieving Chromatography. Supernatants (0.2 ml) from lysates of bacteria transformed with 7F (prepared as above) were applied to an Ultrogel AcA-22 (Pharmacia-LKB; Uppsala, Sweden) column (40 × 1 cm) equilibrated with PBS at 4°C that had been calibrated with dextran (void volume), thyroglobulin (660 kd), ferritin (440 kd), catalase (230 kd) and aldolase (158 kd). Aliquots (0.02 ml) of each fraction (0.8 ml) were electrophoresed on 10% SDS-PAGE, blotted to nitrocellulose, and probed with TCN-2, followed by an alkaline phosphatase-conjugated goat anti-mouse IgG (6). The amount of immunoreactive material in the blots was quantitated with a laser scanning densitometer.

Neuraminidase Activity. Lysates of bacteria prepared as above were analyzed for neuraminidase activity by hydrolysis of the fluorogenic substrate 4 MU-NANA as previously described (5, 6). Lysates were also incubated with human erythrocytes and the degree of cell dehydration ascertained by peanut lectin hemagglutination (12).

Library Construction and Screening. A tissue culture trypomastigote (Silvio X-10/4) genomic DNA library was constructed in λZAP by Stratagene (La Jolla, CA), after shearing and sizing of the DNA to produce inserts of 4-9 kbp. The library was screened unamplified with mouse polyclonal TCNA antibodies and TCN-2 mAb (see above) as described by Huynh et al. (15). Clones 7F and 10C were isolated with the polyclonal antibody, and clones 16C, 19Y and 20G with TCN-2.

DNA Sequencing and Analysis. DNA inserts were automatically excised and subcloned into pBluescript as recommended by Stratagene. Clone 7F insert, despite being the largest of the isolated inserts (Fig. 1 B), was chosen for DNA sequencing because it consistently produced highest TCNA activity and a recombinant protein that reacted the strongest with TCN-1, TCN-2 (Fig. 2), and rabbit and mouse TCNA antibodies. Nucleotide sequence of 7F insert was determined by the chain-termination method of Sanger et al. (16) using double-stranded DNA as templates for T7 DNA polymerase (United States Biochemicals [Cleveland, OH] and Pharmacia-LKB [Uppsala, Sweden]), 7F templates were obtained by exonuclease III/S1 nuclease unidirectional digestion (17) starting from the 3' end of the insert. Ambiguities and gaps were corrected

by using oligonucleotide primers designed from previous sequences. Nucleotide sequence of the other clones was determined in their 3'- and 5'- ends using KS and SK primers (Stratagene), and in some specified regions (see Fig. 1 A), with oligonucleotide primers based on the 7F sequence.

Sequences were analyzed on an Apple MacIntosh computer using the DNA Strider program (18) or programs available from the Molecular Biology Computer Resource Center (MBCRR) at Dana-Farber Cancer Institute, Harvard University (Boston, MA); and from University of Wisconsin (Madison, WI) Genetics Computer Group's sequence analysis software package (GCG) (19) (GCG use was a courtesy of the GRASP Core Center at the New England Medical Center, Boston, MA). Databases searched were Genbank Release 65, Protein Identification Resource Release 25, and Swiss-Prot-16. Optimal alignments between two sequences were determined with the FASTA program (20). Statistical significance of comparisons was computed with the program ALIGN (21).

RNA and DNA Blot Hybridization. Total RNA was prepared from various developmental forms of *T. cruzi* and Vero cells using LiCl/guanidinium isothiocyanate (22). RNA transfer was performed as described (23). Probes were derived from the coding region of

TCNA (fragment 2 was an Exo III/S1 deletion mutant equivalent to bp -92/537 and fragment 1 a PstI/Pvu II digest of 7F, bp 1800/3751; see Fig. 1 B and Fig. 3), labeled by random priming (24), and hybridized under standard conditions. DNA (Southern) blots were performed exactly as described by Sambrook et al. (23). Markers for the RNA blot was the RNA ladder (BRL) and for the DNA blot, a Hind III digest of phage λ (New England Biolabs, Beverly, MA).

Results

Isolation and Nucleotide Sequence of DNA Encoding TCNA.

Mouse polyclonal antibody against TCNA (6) was used as a probe to screen a trypanostigote genomic DNA library constructed in the λ ZAP expression vector. Approximately 10^5 clones from the unamplified library were screened, and two immuno-reactive plaques containing ~9.0- and 8.5-kb inserts (7F and 10C, respectively) were isolated. Additional clones (20G, 19Y and 16C) were isolated using the TCNA-specific mAb TCN-2 (6) (see Fig. 1 B).

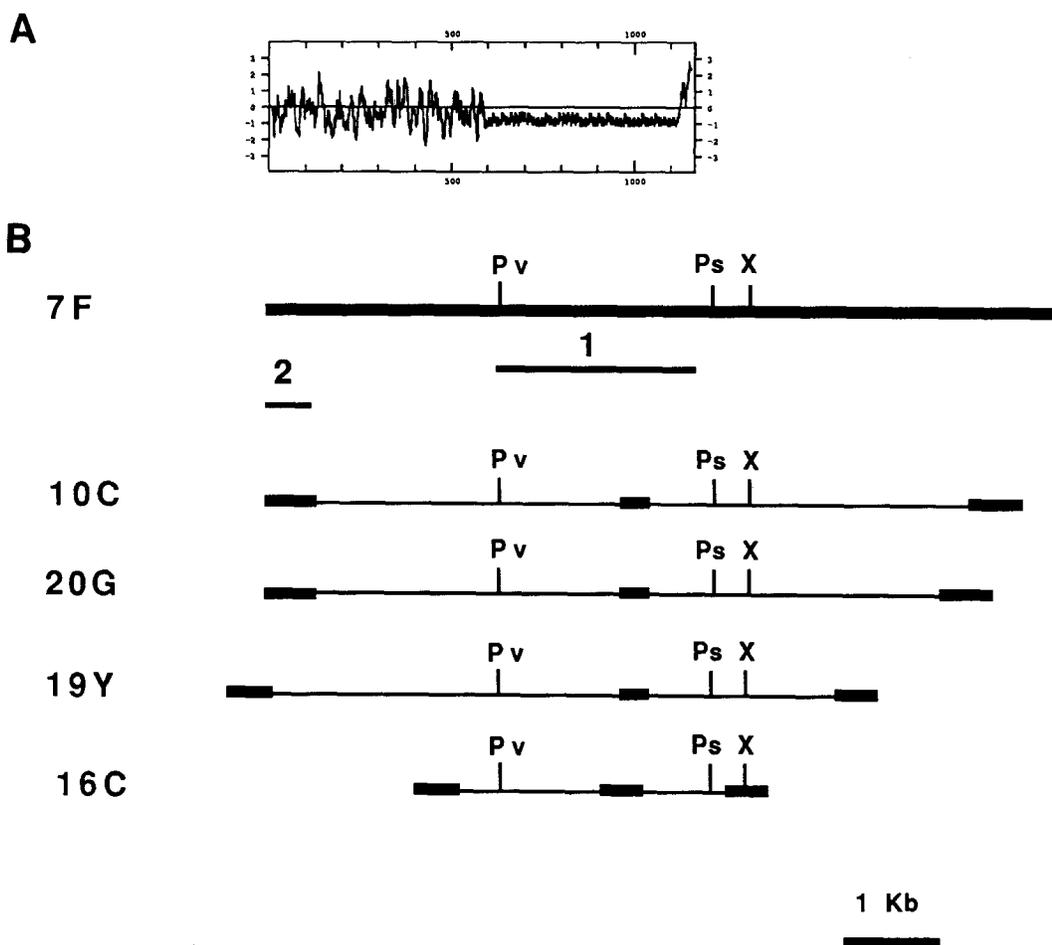


Figure 1. Hydrophobicity plot of the TCNA sequence and schematic representation of the TCNA DNA. (A) Hydrophobicity of the TCNA sequence. The plot is colinear with the TCNA DNA sequence, and was generated according to Kyte and Doolittle (26). Positive values indicate hydrophobic amino acids. The plot clearly shows the hydrophilic LTR and the hydrophobic segment in the C-terminus of TCNA, and the absence of N-terminal peptide sequence. (B). Plasmids and DNA fragments used for sequencing and Northern and Southern blots. Heavy lines indicate the region of each DNA clone that was sequenced. The fragments used for Northern and Southern blots are shown below the restriction map of 7F. Restriction sites marked are: Pv, Pvu II; Ps, Pst I; X, Xho I.

Of the 8,885 nt of 7F only 5,404 are displayed in Fig. 3 because the remaining 3,481 nt in the 3'-end of the insert contains an ORF of 1,061 nt that is transcribed in the three stages of *T. cruzi* (data not shown). The deduced amino acid sequence of this ORF is unrelated to TCNA. Details about the remaining 3,481 nt of 7F will be published elsewhere.

Hydrophobicity and RNA Blot Analysis. Hydrophobicity analysis (26) of TCNA does not reveal typical N-terminal signal peptide sequences (Fig. 1 A). Although most proteins do contain a peptide leader sequence (27), TCNA is not unique in lacking a signal peptide because several proteins such as human IL-1 (28) and bovine scavenger receptor (29) are synthesized without a leader peptide. Therefore such peptide is not an absolute requirement for proteins to be exported to the surface or the environment.

Hydrophobicity analysis also illustrates a hydrophilic LTR in the C-terminus, followed by a nonpolar region of 35 amino acids at the extreme C-terminus. This region is characteristic of proteins anchored to the cell surface by a GPI linkage (30, 31), suggesting that TCNA is GPI-anchored to the trypomastigote surface. This interpretation is consistent with findings showing that TCNA is released spontaneously into the culture medium (5) or by GPI-specific lipases (32).

RNA blotting (Fig. 4) shows expression of the cloned gene in trypomastigotes, epimastigotes, and amastigotes: three major bands of 5.3, 4.5, and 4.0 kb, and three minor bands of 3.1, 2.8, and 2.4 kb hybridized to a probe derived from the catalytic domain (Fragment 2, see Fig. 1 A). Hybridization was more intense with trypomastigotes than with epi- and amastigotes, in accordance with immunoblot analysis (see Fig. 2 A) and neuraminidase activity assays of the developmental forms of *T. cruzi* (3, 4). No hybridization was observed with Vero cells RNA (Fig. 4).

Primary Structure Analysis. The predicted protein sequence of TCNA (Fig. 3) can be considered in four domains (Fig. 5): (I) a N-terminal cysteine-rich region (residues 1-457) which should contain the catalytic site of the enzyme; (II) a fibronectin (FN) type III module (residues 458-588); (III) a C-terminal long tandem repeat (residues 589-1127); and (IV) a hydrophobic segment (residues 1128-1162) at the extreme C-terminus.

Domain I: The Cysteine-Rich Region. The N-terminus of TCNA contains six cysteine residues and exhibits 27% identity with the *C. perfringens* neuraminidase (33) in an alignment extending for about 90% of the bacterial enzyme (Fig. 6). Support for the probable homology between the cysteine-rich domain of TCNA and the *C. perfringens* neuraminidase is provided by statistical analysis using the program ALIGN (21), which gave a score of 10.33 standard deviations above the mean obtained by 100 random permutations of the two sequences. The TCNA domain includes three consensus sequences (X-X-S-X-D-X-G-X-TW-X-X) (Figs. 3 and 6) which are highly conserved in bacterial and viral neuraminidases (34). These bacterial neuraminidase repeats (BNR) are located in separate regions of TCNA, similar to their distribution in bacterial neuraminidases (34). Because of its structural similarity with the *C. perfringens* neuraminidase and with the

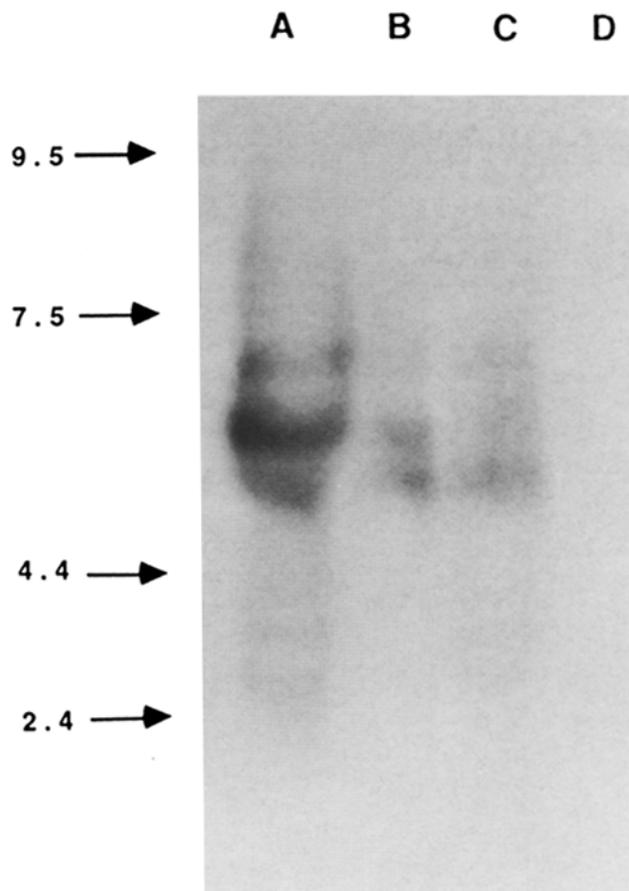


Figure 4. RNA blot analysis of TCNA transcripts. Total RNA (30 μ g) was separated on a 1.2% agarose-formaldehyde gel, blotted to a nylon membrane, and probed with fragment 2 (Fig. 1 B) derived from the catalytic domain, which was labeled with 32 P by random priming. RNA sources were: trypomastigotes (lane A), epimastigotes (lane B), amastigotes (lane C), and Vero cells (lane D).

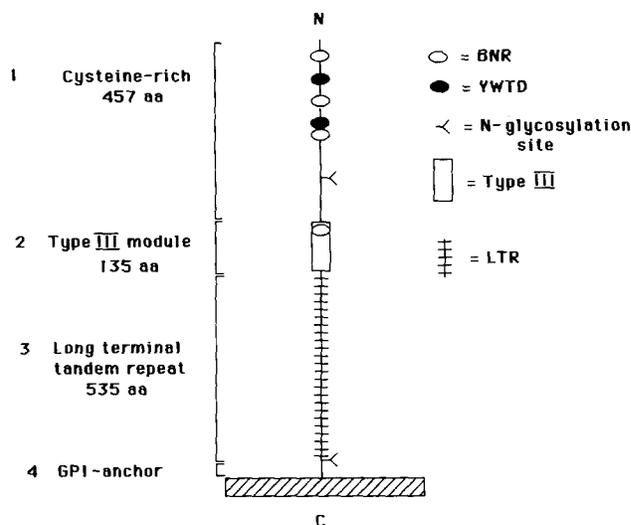


Figure 5. Schematic model of the predicted monomeric structure of TCNA. (○), BNR repeats; (●), YWTD repeats; (Y), N-glycosylation sites; (□), type III module; (|||), LTR domain.

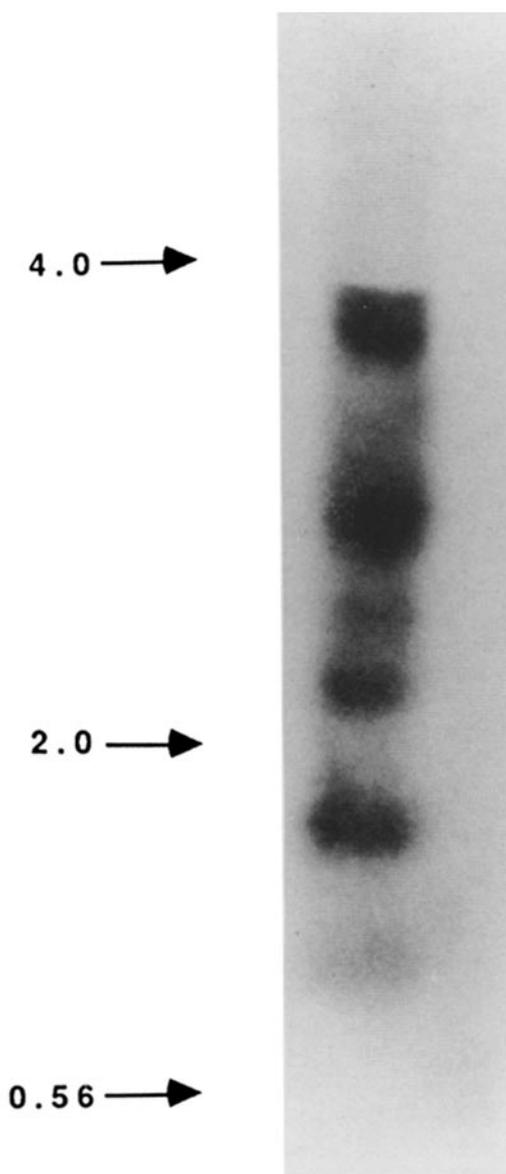


Figure 10. Southern blot of the LTR domain of TCNA. Trypomastigote genomic DNA (Silvio X-10/4) was digested with Pvu II and Pst I, electrophoresed on 1% agarose gel, transferred to a nylon membrane and probed with ^{32}P -labeled fragment 1 (Fig. 1 B) from the LTR domain.

Cysteine-Rich Domain. TCNA has structural features shared by other neuraminidases, the most striking of which is the presence of repeats conserved in bacterial and viral neuraminidases. All bacterial neuraminidases whose primary sequence is known have four repeats of consensus X-X-Ser-X-Asp-X-Gly-X-Thr-Irp-X-X, which are spaced from one another by 40–80 residues (34). TCNA also has four such BNRs (three in domain I and one in domain II) (Figs. 6 and 8) with consensus spacings similar to those in bacterial enzymes. In addition, domain I of TCNA is about 30% identical in sequence to the *C. perfringens* neuraminidase with an ALIGN score of 10.33 standard deviations above the mean, supporting the argument that TCNA domain I and the *C. perfringens*

enzyme are homologous to one another. TCNA, like the *C. perfringens* enzyme, does not have common amino acid sequence with bacterial neuraminidases, except for the BNRs.

TCNA has structural motifs which are unique when compared to other neuraminidases. In the cysteine-rich domain there are two repeats similar to the YWTD motifs found in vertebrate and invertebrate proteins. Little is known about their function, but mutagenesis of the YWTD repeats and flanking cysteine-rich motifs of the LDL receptor produces mutant receptors unable to bind LDL, to release β -VLDL at acid pH, and to recycle properly (59). It is of interest that HDL and LDL are specific inhibitors of TCNA activity (7, 8), and that binding of ^{125}I -HDL to the developmental forms of *T. cruzi* correlates with TCNA activity, suggesting that TCNA can serve as receptors for lipoproteins. If so, it could be that the YWTD motifs of TCNA are involved in binding to host lipoproteins.

FNIII Domain. The FNIII motif of TCNA connects the cysteine-rich domain to LTR (Fig. 5). Like the YWTD repeats, FNIII molecules have not been previously observed in microorganisms. They are found in multiple copies in several developmentally regulated proteins that participate in activities such as binding to cell surface integrins and to heparin. It is therefore possible that the FNIII domain contributes to TCNA binding to host cell-surface proteins during *T. cruzi*-host cell interactions. This hypothesis is consistent with the alteration of *T. cruzi* infection by TCNA antibodies (5, 6). Alternatively, the FNIII domain could mediate TCNA binding to endogenous *T. cruzi* proteins to form ligand-enzyme complexes. This hypothesis is in keeping with recent findings which reveal a protein of 110–115 kd that co-precipitates with TCNA during immunoprecipitation with monoclonal antibodies specific for the *T. cruzi* enzyme but not with isotype-matched control antibodies (Mejia, J.S., R.P. Prioli, E. Ortega-Barria, and M.E.A. Pereira, unpublished observations).

Nucleotide binding consensus motifs have been derived by sequence comparison of many functionally distinct GTP-binding protein families (49–51). The importance of these motifs was confirmed by the X-ray structure of the GTP-binding domain of the elongation factor Tu (60) and of ras protein (61). These studies showed that the motifs GXX-XXGK and DXXG, separated from one another by 40–80 amino acids, are involved in binding to phosphate of GTP, whereas the motif NKXD mediates interaction with the keto group of the guanine ring. The FNIII module of TCNA contains the phosphate binding motifs in correct spacing and order (residues 491–497 and 580–583; respectively), and, although it contains the guanine-ring binding motifs (residues 497–500 and 293–296), they are not spaced properly. Thus, although it is unlikely that the FNIII domain binds GTP, the conservation of GXXXXGK and DXXG raises the intriguing possibility that TCNA binds phosphate, either as sugar-phosphate or as nucleotides (other than GTP).

LTR Domain. The LTR domain is unique among known neuraminidases, but it is a common feature of parasitic protozoan proteins whose primary structures are known (62). Proteins with long tandem repeats usually have anomalous migration in SDS-PAGE, and the differences between the value

predicted by DNA sequence and the one obtained experimentally can vary by as much as 100% (62). The LTR domain helps therefore explain the discrepancy between the predicted mol wt of TCNA (119,959 kD) and the one observed by immunoblotting of recombinant TCNA (180,000 kD).

The function of LTRs in protozoan proteins has not been elucidated, but it has been suggested that tandem motifs, due to their intrinsic repeated structure and consequent enhanced immunogenicity, may form a "smokescreen" against the host immune response to protect domains involved in functions such as in binding to host cells or catalytic activity (62). This could well be the case for TCNA, because the repeat unit of LTR contains epitopes for monoclonal and polyclonal antibodies raised against the intact enzyme. For example, we have found that binding of the mAb TCN-2 to TCNA is competitively inhibited by the soluble synthetic peptide Asp-Ser-Ser-Ala-His-Gly-Thr-Pro-Ser-Thr-Pro-Val, that is repeated many times in LTR, and that TCN-2 and Chagasic sera bind to the peptide coupled to a protein carrier (ovalbumin) (Prioli, R.P., E. Ortega-Barria, J.S. Mejia, and M.E.A. Pereira, unpublished observations). TCN-2 epitopes present in the C-terminus is consistent with the finding that the antibody neutralizes desialylation of erythrocytes but not of soluble small molecules (6). We have not yet found antibodies, monoclonal or polyclonal that react with synthetic peptides from domain I (Fig. 5) or with the intact domain.

The amino acid composition of LTR contains a high content of α -helix- and β -sheet-breaking amino acids (serine, threonine, and proline) in invariant positions (Fig. 9) (63). Empirical predictive methods indicate that the probability of α -helix and β -sheet is less than 1%, and for β -turn more than 99% (64). TCNA may belong to a group of proteins such as RNA polymerase II and rhodopsin, that contain tandem repeats having 20–60% Pro with abundant Gly, Gln, and Ser, which can give rise to a novel secondary structure, polyproline β -turn helix (65). Many of these structures are consistent with a series of turns in an extended conformation.

LTR is similar to the C-terminal sequence of shed acute phase antigen (SAPA) of *T. cruzi*, which contains tandem units of 12 amino acids rich in Ser, Thr, and Pro, ending with a hydrophobic segment (66). SAPA has only 14 tandem repeats, compared to 44 for TCNA. The N-terminus sequence of SAPA is now known, and it is about 80% identical to that of TCNA (A.C.C. Frasch, personal communication). It is of interest that TCNA activity was detected in the serum of a patient accidentally infected with *T. cruzi*, but only in the acute phase of Chagas' disease (67). Since TCNA activity was present before noticeable parasitemia, neuraminidase activity assays were proposed to serve as an aid for the diagnosis of acute Chagas' disease, similar to what has been suggested for SAPA (66). SAPA and corresponding antibodies also appear to be extremely useful reagents to confirm congenital cases of Chagas' disease (68). Although it is not yet known whether SAPA has neuraminidase activity, it is possible that it belongs

to the TCNA family. Differences in their sequences could be attributed to strain variations, to the fact that the TCNA and SAPA sequence were derived from trypomastigote and epimastigote genomes, respectively, or to isoform variations.

TCNA Polymorphism. TCNA from tissue culture trypomastigotes is comprised of a set of high mol wt polypeptides that differ in size and number according to strains and clones of *T. cruzi*. Within a given cloned *T. cruzi* strain, the size of the polypeptides may vary by more than 80 kd (5, 6). The mechanism underlying such polymorphism is not clear, but postranslational modifications may not be one them because pulse-chase experiments failed to show a time-dependent interrelationship in the generation of the polypeptides (Mejia, J.S., E. Ortega-Barria, R.P. Prioli, and M.E.A. Pereira, unpublished observations). Alternatively the polypeptides could arise from distinct TCNA genes or by cotranslational modifications: RNA blots hybridized with DNA derived from the catalytic domain (Fragment 2, Fig. 1 B) show at least 7 bands (Fig. 4). Southern blotting experiments using DNA from the LTR domain as a probe, and genomic DNA digested with Pvu II and Pst I, that cut the repeated structure in its 5' and 3' ends, respectively (Fig. 1 A), reveal a multiple banding pattern. Thus, Southern blotting results are consistent with the notion that variable length polymorphism of the LTR DNA is a mechanism for the multiple banding of TCNA polypeptides.

TCNA polymorphism can also be rationalized on the basis of protein synthesis initiation at different methionine codons. Proteins can be synthesized from overlapping reading frames if the first AUG codon does not lie in an optimum context for ribosomal subunit recognition, such as with many viral proteins (69). Assuming that the initiation signal form high eukaryotes (GCCGCCAGCCAUGG) holds for *T. cruzi*, the first AUG codon of TCNA is in a good but not perfect context (7 out of 13 nucleotides). If the ribosomal subunit bypass the first AUG, it may stop and initiate protein synthesis at the next in-frame AUG (residues 364–366), which is an even better context (8 out of 13 nucleotides). And it is possible that some TCNA isoforms contain AUG initiation codon upstream of the one shown in Fig. 3. But whatever the mechanism for the polymorphism of TCNA may be, it does not seem to abolish enzyme activity since each polypeptide from Silvio X-10/4 (Fig. 2 A) exhibits neuraminidase activity after isolation by sieving chromatography under non-denaturing conditions (Prioli, R.P., J.S. Mejia, and M.E.A. Pereira, manuscript in preparation).

In conclusion, the structure of TCNA as deduced from DNA sequence suggests that the enzyme, in addition to catalyzing hydrolysis of sialic acid from glycoconjugates, may have other important functional properties, including protein-protein interactions and binding to nucleotides. The availability of TCNA DNA clones provides experimental tools that should be useful in further dissecting the structure and function of the enzyme.

We thank Jacqueline Sharon for critical reading of the manuscript and for useful suggestions, and Carlos Frasch and George Cross for communicating unpublished findings and for helpful discussions.

This work is dedicated to the late Professor Paulo Marcello, a superb clinician and scientist from Ceará, Brazil, for his encouragement and interest.

This work was supported by National Institutes of Health grant AI-18102.

Address correspondence to Mierao E. A. Pereira, New England Medical Center Hospital, Division of Geographic Medicine, Box 41, Boston, MA 02111.

The present address of D. Matzilevich is the Imperial College of Science, Technology and Medicine, Department of Pure and Applied Medicine, Prince Consort Road, London SW 2BB, England.

Received for publication 22 February 1991 and in revised form 26 March 1991.

References

1. Brener, Z. 1973. The biology of *Trypanosoma cruzi*. *Ann. Rev. Microbiol.* 27:349.
2. Pereira, M.E.A. 1990. Cell Biology of *Trypanosoma cruzi*. In *Modern Parasite Biology*. D. Wyler, editor. W.H. Freeman and Co., New York. pp. 64-78.
3. Pereira, M.E.A. 1983. A developmentally regulated neuraminidase activity in *Trypanosoma cruzi*. *Science (Wash. DC)*. 219:1444.
4. Pereira, M.E.A., and R. Hoff. 1986. Heterogeneous distribution of neuraminidase activity in strains and clones of *Trypanosoma cruzi* and its possible association with parasite myototropism. *Mol. Biochem. Parasitol.* 20:183.
5. Cavalleco, R., and M.E.A. Pereira. 1988. Antibody to *Trypanosoma cruzi* neuraminidase enhances infection *in vitro* and identifies a subpopulation of trypomastigotes. *J. Immunol.* 140:617.
6. Prioli, R.P., J.S. Mejia, and M.E.A. Pereira. 1990. Monoclonal antibodies against *Trypanosoma cruzi* neuraminidase reveal enzyme polymorphism, recognize a subset of trypomastigotes, and enhance infection *in vitro*. *J. Immunol.* 144:4384.
7. Prioli, R.P., J.M. Ordoval, I. Rosenberg, E.J. Schaeffer, and M.E.A. Pereira. 1987. Similarity of cruzin, an inhibitor of *Trypanosoma cruzi* neuraminidase, to high-density lipoprotein. *Science (Wash. DC)*. 238:1417.
8. Prioli, R.P., I. Rosenberg, and M.E.A. Pereira. 1990. High- and low-density lipoproteins enhance infection of *Trypanosoma cruzi* *in vitro*. *Mol. Biochem. Parasitol.* 38:191.
9. Souto-Padron, T., G. Harth, and W. De Souza. 1990. Immunohistochemical localization of the neuraminidase in *Trypanosoma cruzi*. *Infect. Immun.* 58:586.
10. Prioli, R.P., J.S. Mejia, T. Aji, M. Aikawa, and M.E.A. Pereira. 1991. *Trypanosoma cruzi*: localization of neuraminidase on the surface of trypomastigotes. *Trop. Med. Parasitol.* In press.
11. Libby, P., J. Alroy, and M.E.A. Pereira. 1986. A neuraminidase from *Trypanosoma cruzi* removes sialic acid from the surface of mammalian myocardial and endothelial cells. *J. Clin. Invest.* 77:127.
12. Pereira, M.E.A. 1983. A rapid and sensitive assay for neuraminidase using peanut lectin hemagglutination: application to *Vibrio cholera* and *Trypanosoma cruzi*. *J. Immunol. Meth.* 63:25.
13. Rosenberg, I., R.P. Prioli, J.S. Mejia, and M.E.A. Pereira. 1991. Differential expression of *Trypanosoma cruzi* neuraminidase in intra- and extracellular trypomastigotes. *Infect. Immun.* 59:464.
14. Harth, G., C.G. Haidaris, and M. So. 1987. Neuraminidase from *Trypanosoma cruzi*: analysis of enhanced expression of the enzyme in infectious forms. *Proc. Natl. Acad. Sci. USA.* 84:8320.
15. Huynh, T.V., R.A. Young, and R.W. Davis. 1985. Construction and screening cDNA libraries in lambda gt10 and lambda gt11. In *DNA Cloning Techniques: A Practical Approach*, D. Glover, ed. IRL Press. Oxford. pp. 49.
16. Sanger, F.S., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-termination inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463.
17. Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene (Amst)*. 28:351.
18. Marck, C. 1988. "DNA Strider": a "C" program for the fast analysis of DNA and protein sequences on the Apple McIntosh family of computers. *Nucleic Acids Res.* 16:1829.
19. Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387.
20. Pearson, W.R., and D.J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA.* 85:2444.
21. Dayhoff, M.Q., W.C. Barker, and L.T. Hunt. 1983. Establishing homologies in protein sequences. *Meth. Enzymol.* 91:524.
22. Chirgwin, J.M., A.E. Przybyla, L.J. MacDonald, and W.J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* 18:5294.
23. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*. New York, Cold Spring Harbor Laboratory, Cold Spring Harbor. pp. 1.3-18.86.
24. Feinberg, A.P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6.
25. Kozak, M. 1989. The scanning model for translation: an update. *J. Cell Biol.* 108:229.
26. Kyte, J., and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105.
27. Von Heijne, G. 1985. Signal sequences: the limits of variation. *J. Mol. Biol.* 184:99.
28. Auron, P.E., A.C. Webb, L.J. Rosenwasser, S.F. Mucci, A. Rich, S.M. Wolff, and C.A. Dinarello. 1984. Nucleotide sequence of human monocyte interleukin 1 precursor cDNA. *Proc. Natl. Acad. Sci. USA.* 81:7907.
29. Kodama, T., M. Freeman, L. Rohrer, J. Zabrecky, P. Mat-

- sudaira, and M. Krieger. 1990. Type I macrophage scavenger receptor contains α -helical and collagen-like coiled cells. *Nature (Lond.)*. 343:531.
30. Cross, G. 1990. Glycolipid anchoring of plasma membrane proteins. *Annu. Rev. Cell Biol.* 6:1.
 31. Ferguson, M., and A.F. Williams. 1988. Cell-surface anchoring of proteins via glycosylphosphatidylinositol structures. *Annu. Rev. Biochem.* 57:185.
 32. Rosenberg, I., R.P. Prioli, E. Ortega-Barria, and M.E.A. Pereira. 1991. Stage-specific phospholipase C mediated release of *Trypanosoma cruzi* neuraminidase. *Mol. Biochem. Parasitol.* In press.
 33. Roggentin, P., B. Rothe, F. Lottspeich, and R. Schauer. 1988. Cloning and sequencing of a *Clostridium perfringens* sialidase gene. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 238:31.
 34. Roggentin, P., B. Rothe, J.B. Kaper, J. Galen, L. Lawrisuk, E.R. Vimr, and R. Schauer. 1989. Conserved sequences in bacterial and viral sialidases. *Glycoconjugate J.* 6:349.
 35. Yamamoto, T., C.G. Davis, M.S. Brown, W.J. Schneider, M.L. Casey, J.L. Goldstein, and D.W. Russell. 1984. The human LDL receptor: a cysteine-rich protein with multiple Alu sequences in its mRNA. *Cell.* 39:27.
 36. Herz, J., V. Hamann, S. Rogne, O. Myklebost, H. Gausepohl, and K.K. Stanley. 1988. Surface location and high affinity for calcium of a 500 Kd liver membrane protein closely related to the LDL protein receptor suggest a physiological role as lipoprotein receptor. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:4119.
 37. Gray, A., T.J. Dull, and A. Ullrich. 1983. Nucleotide sequence of epidermal growth factor cDNA predicts a 128,000-molecular weight protein precursor. *Nature (Lond.)*. 303:722.
 38. Mann, K., R. Deutzmann, M. Aumailley, R. Timpl, L. Raimondi, Y. Yamada, T.-C. Pan, D. Conway, and M.-L. Chu. 1989. Amino acid sequence of mouse nidogen, a multidomain basement membrane protein with binding activity for laminin, collagen IV and cells. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:65.
 39. Norton, P.A., R.O. Hynes, and D.J.G. Rees. 1990. Sevenless: seven found? *Cell.* 61:15.
 40. Kemp, B.E., and R.B. Pearson. 1990. Protein kinase recognition sequence motifs. *TIBS (Trends Biochem. Sci.)*. 15:342.
 41. Skorstengaard, K., M.S. Jensen, P. Sahl, T.E. Petersen, and S. Magnusson. 1986. Complete primary structure of bovine plasma fibronectin. *Eur. J. Biochem.* 161:441.
 42. Jones, F.S., M.P. Gurgoon, S. Hoffman, K.L. Grossin, B.A. Cunningham, and G.E. Edelman. 1988. A cDNA clone for cytotoxin contains sequences similar to epidermal growth factor-like repeats and segments of fibronectin and fibrinogen. *Proc. Natl. Acad. Sci. USA.* 85:2186.
 43. Ranscht, B. 1988. Sequences of contactin, a 130 KD-glycoprotein concentrated in areas of interneuronal contact, defines a new member of the immunoglobulin supergene family in the nervous system. *J. Cell. Biol.* 107:1561.
 44. Streuli, M., N.X. Kruger, L.R. Hail, S.F. Schlossman, and H. Saito. 1988. A new member of the immunoglobulin superfamily that has a cytoplasmic region homologous to the leukocyte common antigen. *J. Exp. Med.* 168:1523.
 45. Cunningham, B.A., J.J. Hemperly, B.A. Murray, E.A. Prediger, R. Brackenbury, and G.E. Edelman. 1987. Neural cell adhesion molecule: structure, immunoglobulin-like domains, cell surface modulation, and alternative RNA splicing. *Science (Wash. DC)*. 236:799.
 46. Spring, J., K. Beck, and R. Chiquet-Ehrisman. 1989. Two contrary functions of tenascin: dissection of the active sites by recombinant tenascin fragments. *Cell.* 59:325.
 47. Pathy, L. 1990. Homology of a domain of the growth hormone/prolactin receptor family with type III modules of fibronectin. *Cell.* 61:13.
 48. Harrelson, A.L., and C.S. Goodman. 1988. Growth core guidance in insects: fascilin II is a member of the immunoglobulin superfamily. *Science (Wash. DC)*. 242:700.
 49. Dever, T.E., M.J. Glynias, and W.C. Merrick. 1987. GTP-binding domain: three consensus sequence elements with distinct spacing. *Proc. Natl. Acad. Sci. USA.* 84:1814.
 50. Bourne, H.R., D.A. Sanders, and F. McCormick. 1991. The GTPase superfamily: conserved structure and molecular mechanism. *Nature (Lond.)*. 349:117.
 51. Santos, E., and A.R. Nebreda. 1989. Structural and functional properties of ras proteins. *FASEB (Fed. Am. Soc. Exp. Biol.) J.* 3:2151.
 52. Corden, J.L., D.L. Cadena, J.M. Ahearn, and M.E. Dohmus. 1985. A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc. Natl. Acad. Sci. USA.* 82:7934.
 53. Corden, J.L. 1990. Tails of RNA polymerase II. *TIBS (Trends Biochem. Sci.)*. 15:383.
 54. Rosenberg, A., and C.-L. Schengrund. 1976. Sialidases. In *Biological Roles of Sialic Acid*. A. Rosenberg and C.-L. Schengrund, editors. Plenum Press, New York. pp 295-359.
 55. Schauer, R. 1982. Chemistry, metabolism, and biological function of sialic acid. *Adv. Carbohydr. Chem. Biochem.* 40:131.
 56. Cantz, M., J. Gehler, and J. Sprander. 1977. Mucopolidosis I: Increased sialic acid content due to a deficiency of an α -N-acetylneuraminidase in cultured fibroblasts. *Biochem. Biophys. Res. Commun.* 74:732.
 57. Colman, P.M., and C.W. Ward. 1985. Structure and diversity of influenza virus neuraminidase. *Curr. Top. Microbiol. Immunol.* 114:177.
 58. Russo, T.A., J.S. Thompson, V.G. Godoy, and M.H. Malamy. 1990. Cloning and expression of the *Bacteriodes fragilis* TAL 2480 neuraminidase gene, manH, in *Escherichia coli*. *J. Bacteriol.* 172:2594.
 59. Davis, C.G., J. Goldstein, T.C. Südhoff, R.G.W. Anderson, D.W. Russell, and M.S. Brown. 1987. Acid-dependent ligand dissociation and recycling of LDL receptor mediated by growth factor homology region. *Nature (Lond.)*. 326:760.
 60. la Cour, T.F.M., J. Nyborg, S. Thirup, and B.F.C. Clark. 1985. Structural details of the binding of guanosine diphosphate to elongation factor Tu from *E. coli* as studied by X-ray crystallography. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:2385.
 61. De Vos, A.M., L. Tong, M.V. Milburn, P.M. Matias, J. Jancaik, S. Noguchi, S. Nishimura, K. Miura, E. Ohtsuka, and S.H. Kim. 1989. Three-dimensional structure of an oncogene protein: catalytic domain of human c-H-ras p21. *Science (Wash. DC)*. 239:888.
 62. Anders, R.F., R.L. Coppel, G.V. Brown, and D.J. Kemp. 1988. Antigen with repeated amino acid sequences from the asexual blood stages of *Plasmodium falciparum*. *Prog. Allergy*. 4:148.
 63. Richardson, J.W., and D.C. Richardson. 1989. Principles and patterns of protein conformation. In *Prediction of Protein Structure and the Principles of Protein Conformation*. G.D. Fasman, editor. Plenum Press, New York. pp 1-98.
 64. Chou, D.Y., and G.D. Fasman. 1974. Conformational parameters for amino acids in helical, β -sheet, and random coil regions calculated from proteins. *Biochemistry*. 13:211.
 65. Matsushima, N., C.E. Creutz, and R.H. Kretsinger. 1990. Polyproline, β -turn helices. Novel secondary structures proposed for the tandem repeats within rhodopsin, synaptophysin,

- synexin gliadin, RNA polymerase II, hordein, and gluten. *Proteins*. 7:125.
66. Affranchino, J.L., C.F. Ibañez, A.O. Liguetti, A. Rassi, M.B. Reyes, R.A. Macina, L. Åslund, U. Pettersson, and A.C.C. Frasch. 1989. Identification of a *Trypanosoma cruzi* antigen that is shed during the acute phase of Chagas' disease. *Mol. Biochem. Parasitol.* 34:221.
67. de Titto, E., and F.G. Araujo. 1988. Serum neuraminidase activity and hematological alterations in acute human Chagas' disease. *Clin. Immunol. Immunopathol.* 46:157.
68. Reyes, M.B., M. Lorca, P. Muñoz, and A.C.C. Frasch. 1990. Fetal IgG specificities against *Trypanosoma cruzi* antigens in infected newborns. *Proc. Natl. Acad. Sci. USA.* 87:2846.
69. Kozak, M. 1986. Bifunctional messenger RNAs in eukaryotes. *Cell.* 47:481.