THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 2003 by The American Society for Biochemistry and Molecular Biology, Inc.
Vol. 278, No. 18, Issue of May 2, pp. 15484–15494, 2003
Printed in U.S.A.

Biochemical Characterization of Proline Racemases from the Human Protozoan Parasite *Trypanosoma cruzi* and Definition of Putative Protein Signatures*

Received for publication, October 23, 2002, and in revised form, January 30, 2003
Published, JBC Papers in Press, February 11, 2003, DOI 10.1074/jbc.M210830200

Nathalie Chamond‡§, Christophe Grégoire‡§¶, Nicolas Coatnoan‡, Catherine Rougeot‡, Lucio Holanda Freitas-Junior‖, José Franco da Silveira, Wim M. Degrave‡‡, and Paola Minoprio‡§§**

*From the Departments of ‡Immunology and ‖Parasitology, Institut Pasteur, Paris, 75724, France, **Department of Microbiology, Immunology and Parasitology, Unifesp/Escola Paulista de Medicina, Sao Paulo 04023-062, Brazil, and ‡‡Department of Biochemistry and Molecular Biology, Instituto Oswaldo Cruz, Rio de Janeiro 21045-900, Brazil*

Proline racemase catalyzes the interconversion of L- and D-proline enantiomers and has to date been described in only two species. Originally found in the bacterium *Clostridium sticklandii*, it contains cysteine residues in the active site and does not require co-factors or other known coenzymes. We recently described the first eukaryotic amino acid (proline) racemase, after isolation and cloning of a gene from the pathogenic human parasite *Trypanosoma cruzi*. Although this enzyme is intracellularly located in replicative non-infective forms of *T. cruzi*, membrane-bound and secreted forms of the enzyme are present upon differentiation of the parasite into non-dividing infective forms. The secreted form of proline racemase is a potent host B-cell mitogen supporting parasite evasion of specific immune responses. Here we describe that the *TcPRAC* genes in *T. cruzi* encode functional intracellular or secreted versions of the enzyme exhibiting distinct kinetic properties that may be relevant for their relative catalytic efficiency. Although the K_m of the enzyme isoforms were of a similar order of magnitude (29–75 mM), V_{max} varied between 2×10^{-4} and 5.3×10^{-5} mol of L-proline/s/0.125 μM of homodimeric recombinant protein. Studies with the enzyme-specific inhibitor and abrogation of enzymatic activity by site-directed mutagenesis of the active site Cys^{330} residue reinforced the potential of proline racemase as a critical target for drug development against Chagas' disease. Finally, we propose a protein signature for proline racemases and suggest that the enzyme is present in several other pathogenic and nonpathogenic bacterial genomes of medical and agricultural interest, yet absent in mammalian host, suggesting that inhibition of proline racemases may have therapeutic potential.

D-Amino acids have long been described in the cell wall of Gram-positive and especially Gram-negative bacteria, where they constitute essential elements of the peptidoglycan and as substitutes of cell wall techoic acids (1). Moreover, various types of D-amino acids were discovered in a number of small peptides made by a variety of microorganisms through nonribosomal protein synthesis (2) that function mainly as antibiotic agents. However, these examples were considered exceptions to the rule of homochirality, and a dogma persisted that only L-amino acid enantiomers were present in eukaryotes, apart from a very low level of D-amino acids from spontaneous racemization because of aging (3). Recently, an increasing number of studies have reported the presence of various D-amino acids either as protein-bound (4) or under free forms (5) in a wide variety of organisms, including mammals. The origin of free D-amino acids is less clear than that of protein bound D-amino acids. For instance, in mammals, free D-amino acids may originate from exogenous sources (as described in Ref. 6), but the recent discovery of amino acid racemases in eukaryotes has also uncovered an endogenous production of D-amino acids, questioning their specific functions. Thus, the level of D-aspartate is developmentally regulated in rat embryos (7); the binding of D-serine to *N*-methyl-D-aspartate mouse brain receptors promotes neuromodulation (8, 9) and D-aspartate appears to be involved in hormonal regulation in endocrine tissues (10). All amino acid racemases require pyridoxal phosphate as a cofactor except proline and hydroxyproline racemases, which are co-factor-independent enzymes. Two reports have been published addressing the biochemical and enzymatic characteristics of the proline racemase from the Gram-positive bacterium *Clostridium sticklandii* (11, 12). A reaction mechanism was proposed whereby the active site Cys^{256} forms a half-reaction site with the corresponding cysteine of the other monomer in the active, homodimeric enzyme.

Although a variety of racemases and epimerases has been demonstrated in bacteria and fungi, we recently described the first eukaryotic amino acid (proline) racemase isolated from the infective metacyclic forms of the parasitic protozoan *Trypanosoma cruzi*, the causative agent of Chagas' disease in humans (13). This parasite-secreted proline racemase (*Tc*PRAC) was shown to be a potent mitogen for host B cells and plays an important role in *T. cruzi* immune evasion and persistence through polyclonal lymphocyte activation (13). This protein, previously annotated as TcPA45, with monomer size of 45 kDa, is only expressed and released by infective metacyclic forms of the parasite. The genomic organization and transcription of *TcPRAC* proline racemase gene indicated the presence of two homologous genes per haploid genome (*TcPRACA* and *TcPRACB*). Furthermore, localization studies using specific antibodies directed to 45-kDa *TcPRAC* protein revealed that an intra-

* This work was supported in part by Institut Pasteur and CNRS URA 1960. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AY 1409447.

§ Contributed equally to this work.
¶ Fellow of the Pasteur Institute.
§§ To whom correspondence should be addressed. Tel./Fax: 33-1-45-68-86-15; E-mail: pmm@pasteur.fr.

This paper is available on line at http://www.jbc.org

cellular and/or membrane-associated isoform, with monomer size of 39 kDa, is expressed in non-infective epimastigote forms of the parasite. Computer-assisted analysis of the *TcPRACA* gene sequence suggested that it could give rise to both isoforms (45 and 39 kDa) of parasite proline racemases through a mechanism of alternative *trans*-splicing, one of which would contain a signal peptide (13). In addition, preliminary analysis of putative *TcPRACB* gene sequences had revealed several differences that include point mutations as compared with *TcPRACA* but that also suggest that *TcPRACB* gene could only encode an intracellular isoform of the enzyme as the gene lacks the export signal sequence. Any of these molecular mechanisms *per se* would ensure the differential expression of intracellular and extracellular isoforms of proline racemases produced in different *T. cruzi* developmental stages.

Primarily it was essential to elucidate whether *TcPRACB* gene could encode a functional proline racemase. To answer this question we have expressed *TcPRACA* and *TcPRACB* paralogue genes in *Escherichia coli* and performed detailed studies on biochemical and enzymatic characteristics of the recombinant proteins. We show here that *TcPRACB* indeed encodes a functional proline racemase that exhibits slightly different kinetic parameters and biochemical characteristics when compared with *Tc*PRACA enzyme. Enzymatic activities of the respective recombinant proteins showed that the 39-kDa intracellular isoform of proline racemase produced by *TcPRACB* construct is more stable and has a higher rate of D-/L-proline interconversion than the 45-kDa isoform produced by *TcPRACA*. Additionally, the dissociation constant of the enzyme-inhibitor complex (K_i) obtained with pyrrole-2-carboxylic acid, the specific inhibitor of proline racemases, is lower for the recombinant *Tc*PRACB enzyme. Moreover, we show that Cys^{330} is a key amino acid of the proline racemase active site, because the activity of the enzyme is totally abolished by site-directed mutagenesis of this residue. Finally, multiple alignment of proline racemase amino acid sequences allowed the definition of protein signatures that can be used to identify putative proline racemases in other microorganisms. The significance of the presence of proline racemase in eukaryotes, particularly in *T. cruzi*, is discussed, as well as the putative consequences of this enzymatic activity in the biology and infectivity of the parasite.

EXPERIMENTAL PROCEDURES

Cloning and Automated Sequencing—λ phage and plasmid DNA were prepared using standard techniques, and direct sequencing was accomplished with the Big dye Terminator kit (PerkinElmer Life Sciences) according to the manufacturer's instructions. Extension products were run for 7 h in an ABI 377 automated sequencer. Briefly, to obtain the full length of the *TcPRAC* gene, ^{32}P-labeled 239-bp PCR product was used as a probe to screen a *T. cruzi* clone CL Brener λ Fix II genomic library (see details in Ref. 13). We isolated four independent positive phages. Restriction analysis and Southern blot hybridization showed two types of genomic fragments, each represented by two phages. Complete sequence and flanking regions of representative phages for each pattern was done. Complete characterization of *TcPRACA* gene, representing the first phage type, was described previously in Ref. 13. Full sequence of the putative *TcPRACB* gene, representing the second phage type, was then performed, and primers internal to the sequence were used for sequencing as described before (13).

Chromoblots—Epimastigote forms of *T. cruzi* (clone CL Brener) are maintained by weekly passage in LIT medium. Agarose (0.7%) blocks containing 1×10^7 cultured parasites were lysed with 0.5 M EDTA/10 mM Tris/1% sarcosyl, pH 8.0, digested by proteinase K and washed in 10 mM Tris/1 mM EDTA, pH 8.0. Pulsed field gel electrophoresis was carried out at 18 °C using the Gene Navigator apparatus (Amersham Biosciences) in 0.5× TBE. Electrophoresis were performed as described (14). Gels were then stained with ethidium bromide, photographed, exposed to UV light (265 nm) for 5 min, and further blotted under alkaline conditions to a nylon filter (HybondN$^+$; Amersham Biosciences). DNA probe, obtained by PCR amplification of *TcPRACA* gene with Hi-45 (5′ CTC TCC CAT GGG GCA GGA AAA GCT TCT G 3′) and Bg-45 (5′ CTG AGC TCG ACC AGA T(CA)T ACT GC 3′) oligonucleotides (as described in Ref. 13), was labeled with αdATP^{32} using a Megaprime DNA labeling system (Amersham Biosciences). The chromoblot was hybridized overnight in 2× Denhardt's/5× saline/sodium phosphate/EDTA/1.5% SDS at 55 °C and washed in 2× saline/sodium phosphate/EDTA/0.1% SDS followed by 1× saline/sodium phosphate/EDTA at 60 °C. Autoradiography was obtained by overnight exposure of the chromoblot using a PhosphorImager cassette (Molecular Dynamics).

Plasmid Construction and Protein Purification—*TcPRACA* gene fragment starting at codon 30 was obtained by PCR, using Hi- and Bg-45 primers, and cloned in frame with a C-terminal His_6 tag into the pET28b(+) expression vector (Novagen-Tebu, Le Perray en Yvelines, France). The fragment encoding for the *Tc*PRACB consisted of a *Hin*dIII digestion of *TcPRACB* gene fragment obtained by similar PCR and cloned in-frame with a C-terminal His_6 tag into the pET28b(+) expression vector. Respective recombinant proteins *Tc*PRACA and *Tc*PRACB were produced in *E. coli* BL21 (DE3) (Invitrogen) and purified using immobilized metal affinity chromatography on nickel columns (Novagen-Tebu, Le Parray en Yvelines, France) following the manufacturer's instructions.

Size Exclusion Chromatography—r*Tc*PRACA and r*Tc*PRACB proteins were purified as described above and dialyzed against phosphate-buffered saline, pH 7.4, or 0.2 M NaOAc, pH 6.0, elution buffers in dialysis cassettes (Slide-A-lyzer 7K; Pierce), overnight at 4 °C. The final protein concentration was adjusted to 2 mg/ml, and 0.5 ml of the solution were loaded onto Amersham Biosciences Superdex 75 column (HR10 × 30), calibrated previously with a medium range protein calibration kit (Amersham Biosciences). Size exclusion chromatography was carried out using a fast protein liquid chromatography system (AKTA Purifier; Amersham Biosciences). Elution was performed at a constant flow rate of 0.5 ml/min, protein fractions of 0.5 ml were collected, and the absorbance was monitored at 280 nm. Each fraction was assayed in racemization assays as described below. Fractions B1 and B5 were reloaded in the Superdex 75 column and submitted to a further size exclusion chromatography to verify the purity of the fractions.

Racemization Assays—The percent of racemization with different concentrations of L-proline, D-proline, L-hydroxy (OH)-proline, D-hydroxy (OH)-proline was calculated as described (13) by incubating a 500-μl mixture of 0.25 μM dimeric protein and 40 mM substrate in 0.2 M sodium acetate, pH 6.0, for 30 min or 1 h at 37 °C. The reaction was stopped by incubating for 10 min at 80 °C and freezing. Water (1 ml) was then added, and the optical rotation was measured in a polarimeter 241MC (PerkinElmer Life Sciences) at a wavelength of 365 nm, in a cell with a path length of 10 cm, at a precision of 0.001 °. The percent of racemization of 40 mM L-proline as a function of pH was determined using 0.2 M sodium acetate, potassium phosphate, and Tris-HCl buffers; reactions were incubated 30 min at 37 °C as described above. All reagents were purchased from Sigma.

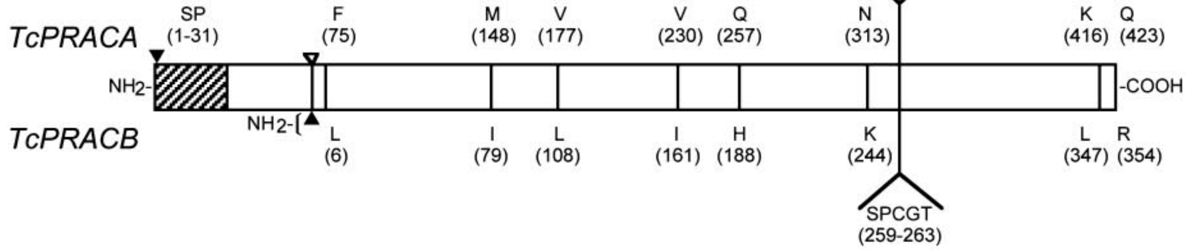
Kinetic Assays—Concentrations of L- and D-proline were determined polarimetrically from the optical rotation of the solution at 365 nm in a cell of 10-cm path length, thermostated at 37 °C. Preliminary assays were done with 40 mM L-proline in 0.2 M sodium acetate, pH 6, in a final volume of 1.5 ml. Optical rotation was measured every 5 s during 10 min and every 5 min to 1 h. After determination of the linear part of the curve, velocity in 5–160 mM substrate was measured every 30 s during 10 min to determine K_m and V_{max}. Calculations were done using the Kaleïdagraph program. Inhibition assays were done by incubating 0.125 μM dimeric protein, 6.7 μM-6 mM pyrrole-2-carboxylic acid (PAC),[1] 20 to 160 mM L-proline, as described above. Graphic representation and linear curve regression allowed the determination of K_i as [PAC]/[(slope with PAC/slope without PAC) − 1]. All reagents were purchased from Sigma.

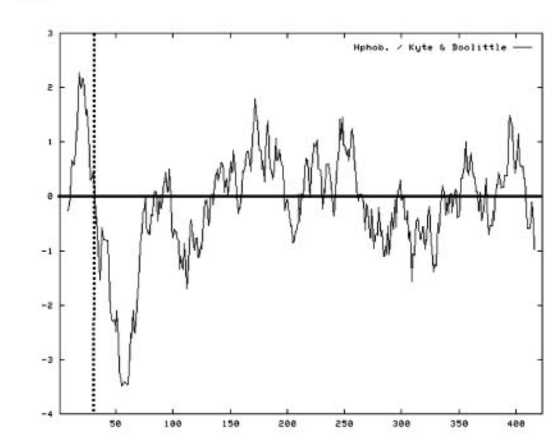
Site-directed Mutagenesis of ^{C330S}TcPRACA—Site-directed mutagenesis was performed by PCR, adapting the method of Higuchi *et al.* (15). Briefly, mutation of Cys^{330} of the proline racemase active site was produced by two successive polymerase chain reactions based on site-directed mutagenesis using two overlapping mutagenic primers, act-1 (5′ GCG GAT CGC TCT CCA **AGC** GGG ACA GGC ACC 3′) and act-2 (5′ GGT GCC TGT CCC **GCT** TGG AGA GCG ATC CGC 3′), designed to introduce a single codon mutation in the active site by replacement of the cysteine (TGT) at position 330 by a serine (AGC). A first step

[1] The abbreviations used are: PAC, pyrrole-2-carboxylic acid; CsPR, *C. sticklandii* proline racemase.

A

```
Tc-A    1 -------------------------------CCTTTTTCTTTTT-------------------AAAAACAAAAAAAATTCCGGGGGGAATATGGAACAGG-   50
Tc-B    1 GTGTGTTCAACAGTTTTGTTTCCTTTTTTCTCTTTTTCTCTTTCCATCATACATACATACATACATACATATATATATCT-GCGTAGATATGC-ACATGC   98

Tc-A   51 GTATATGCGTAAA-AGTGTCTGTCCCAAACAAAAATTTTTTTTTTCCGCCTTCCCATTTTTTTTTTTTTTTTTGTGTGTTT-CCCTTGATCTCTCGAACAG  148
Tc-B   99 GTATATGCGTGAAGAGTGTCTGTCCCAA-------------------------CATTTTTTTTTTTTTTTTTGTGTGTTTTCCCTTGATTCC-CGAACG  171

Tc-A  149 GGCAGGAAAAGCTTCTGTTTGACCAAAAATATAAAATTATTAAGGGCGAGAAAAAAGAAAAGAAAAAAAAATCAACGAGCAAACAGGAGAGAACACCAACA  248
Tc-B  172 GGCAGGAAAAGCTTCTGTTTGACCAAAAATATAAAATTATTAAGGGCGAGAAAA--GAAA---AAAAAAATCAACCG------AGGAGACAACACCAACA  260

Tc-A  249 AAAAAGGGAAATTATGCGATTTAAGAAATCATTCACATGCATCGACATGCATACGGAAGGTGAAGCAGCACGGATTGTGACGAGTGGTTTGCCACACATT  348
Tc-B  261 AAAAAGGGAAATTATGCGATTTAAGAAATCATTGACATGCATCGACATGCATACGGAAGGTGAAGCAGCACGGATTGTGACGAGTGGTTTGCCACACATT  360

Tc-A  349 CCAGGTTCGAATATGGCGGAGAAGAAAGCATACCTGCAGGAAAACATGGATTATTTGAGGCGTGGCATAATGCTGGAACCACGTGGTCATGATGATATGT  448
Tc-B  361 CCAGGTTCGAATATGGCGGAGAAGAAAGCATACCTGCAGGAAAACATGGATTATTTGAGGCGTGGCATAATGCTGGAGCCACGTGGTCATGATGATATGT  460

Tc-A  449 TTGGAGCCTTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCATGGTATTCATGGATACCGGTGGCTATTTAAATATGTGTGGACATAACTCAAT  548
Tc-B  461 TTGGAGCCTTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCATCGTATTCATGGATACCGGTGGCTATTTAAATATGTGTGGACATAACTCAAT  560

Tc-A  549 TGCAGCGGTTACGGCGGCAGTTGAAACGGGAATTGTGAGCGTGCCGGCGAAGGCAACAAATGTTCCGGTTGTCCTGGACACACCTGCGGGGTTGGTGCGC  648
Tc-B  561 TGCAGCGGTTACGGCGGCAGTGGAAACGGGAATTTTGAGCGTGCCGGCGAAGGCAACAAATGTTCCGGTTGTCCTGGACACACCTGCGGGGTTGGTGCGC  660

Tc-A  649 GGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAGGTGTCAAATGCGAGTATTATCAATGTACCCTCATTTTTGTATCAGCAGGATGTGGTGGTTGTGT  748
Tc-B  661 GGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAGGTGTCAAATGCGAGTATTATCAATGTGCCCTCATTTTTGTATCAGCAGGATGTGGTGATTGTTT  760

Tc-A  749 TGCCAAAGCCCTATGGTGAAGTACGGGTTGATATTGCATTTGGAGGCAATTTTTTCGCCATTGTTCCCGCGGAGCAGTTGGGAATTGATATCTCCGTTCA  848
Tc-B  761 TGCCAAAGCCCTATGGTGAGGTACGGGTTGATATTGCATTTGGAGGCAATTTTTTCGCCATTGTTCCCGCGGAGCACTTGGGAATTGATATCTCCGTTCA  860

Tc-A  849 AAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTTCTGCGTACTGAAATCAATCGCAGTGTGAAGGTTCAGCACCCTCAGCTGCCCCATATTAACACTGTG  948
Tc-B  861 AAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTTCTGCGTACTGAAATCAATCGCAGTGTGAAGGTTCAGCACCCTCAGCTGCCCCATATTAACACTGTG  960

Tc-A  949 GACTGTGTTGAGATATACGGTCCGCCAACGAACCCGGAGGCAAACTACAAGAACGTTGTGATATTTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGA 1048
Tc-B  961 GACTGTGTTGAGATATACGGTCCGCCAACGAACCCGGAGGCAAAATACAAGAACGTTGTGATATTTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGA 1060

Tc-A 1049 CAGGCACCAGCGCCAAGATGGCAACACTTTATGCCAAAGGCCAGCTTCGCATCGGAGAGACTTTTGTGTACGAGAGCATACTCGGCTCACTCTTCCAGGG 1148
Tc-B 1061 CAGGCACCAGCGCCAAGATGGCAACACTTTATGCCAAAGGCCAGCTTCGCATCGGAGAGACTTTTGTGTACGAGAGCATACTCGGCTCACTCTTCCAGGG 1160

Tc-A 1149 CAGGGTACTTGGGGAGGAGCGAATACCGGGGGTGAAGGTGCCGGTGACCAAAGATGCCGAGGAAGGGATGCTCGTTGTAACGGCAGAAATTACTGGAAAG 1248
Tc-B 1161 CAGGGTACTTGGGGAGGAGCGAATACCGGGGGTGAAGGTGCCGGTGACCAAAGATGCCGAGGAAGGGATGCTCGTTGTAACGACAGAAATTACTGGAAAG 1260

Tc-A 1249 GCTTTTATCATGGGTTTCAACACCATGCTGTTTGACCCAACGGATCCGTTTAAGAACGGATTCACATTAAAGCAGTAGATCTGGTAGAGCACAGAAACTA 1348
Tc-B 1261 GCTTTTATCATGGGTTTCAACACCATGCTGTTTGACCCAACGGATCCGTTCTTAAACGGATTCACACTAAAGCGGTAGATCTGGTAGAGCACAGAAACTA 1360

Tc-A 1349 TTGGGGAACACGTGCGAACAGGTGCTGCTACGTGAAGGGTATTGAATGAATCGTTTTTTTTTATTTTTATTTTTTATTTTTATTAGTGCATTATTATTAA 1448
Tc-B 1361 TTGGGGAACACGTGCGAACAGGTGCTGCTACGTAAAGGGTATTGAATGAATCGTTTTTTTTTTTTTTT---------TTTTATTAGTGCATTATTTTT-- 1449

Tc-A 1449 ATTTTTTTTTTGTTTTGGGGTTTCAACGGTACCGCGTTGGGAGCAGGGAAGCGATAGCGGCCGGACAATTTTTTGCTTTTATTTTCATTTTCATCTTCCT 1548
Tc-B 1450 -TTTTTTTTTTGTTTTGGGGTTTCAACGGTACCACGTTGGGAGCAGGGAAACGATAGCGGCCGGACAATTTTTTACTTTTATTTTCATTTTCACCTTCCT 1548

Tc-A 1549 ACCCAACCCCCTTGGTTCCACCGGTCGCGGCGGGGTCTTGTGGGTGGAGG 1598
Tc-B 1549 ACCCAACCCCCTTGGTTCCACCGGTCGCGGCGGGG--------------- 1583
```





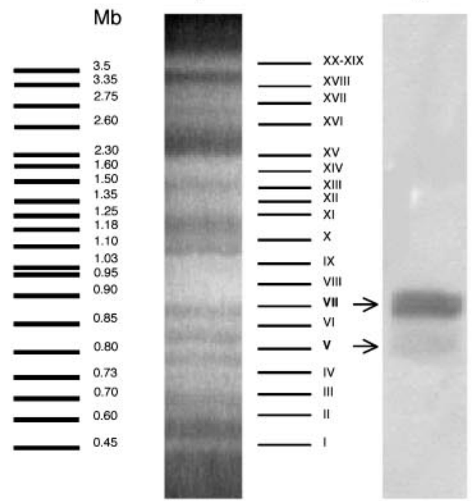


FIG. 1. **Comparative analysis of sequences of *T. cruzi* *Tc*PRACA and *Tc*PRACB proline racemase isoforms.** *A*, alignment of *TcPRACA* (*Tc-A*) and *TcPRACB* (*Tc-B*) nucleotide sequences. Non-coding sequences are shown in *italics*; trans-splicing signals are *underlined*; putative spliced leader acceptor sites are *double underlined*; the region encoding the computer-predicted signal peptide is indicated by a *double-headed*

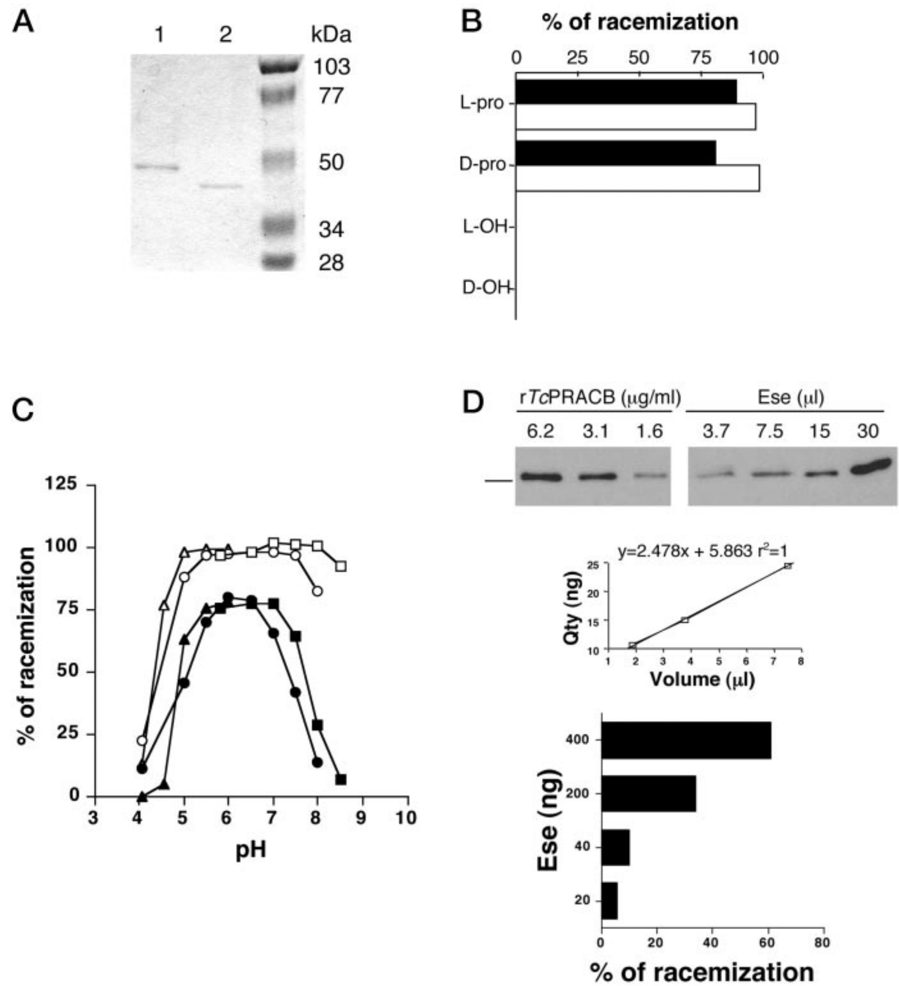


FIG. 2. **Biochemical characterization of *T. cruzi* proline racemase isoforms and substrate specificities.** *A*, SDS-PAGE analysis of purified r*Tc*PRACA (*lane 1*) and r*Tc*PRACB (*lane 2*) recombinant proteins. An 8% polyacrylamide gel was stained with Coomassie Blue. *Right margin*, molecular masses. *B*, percent of racemization of L-proline, D-proline, L-hydroxy (OH)-proline, and D-hydroxy (OH)-proline substrates by r*Tc*PRACB (*open bar*) as compared with r*Tc*PRACA (*closed bar*). Racemase activity was determined with 0.25 µM of each isoform of proline racemase and 40 mM substrate in sodium acetate buffer, pH 6.0. *C*, percent of racemization as a function of pH. Racemase assays were performed in buffer containing 0.2 M Tris-HCl (*squares*), sodium acetate (*triangles*), and potassium phosphate (*circles*), 40 mM L-proline and 0.25 µM of purified r*Tc*PRACA (*closed symbols*), and r*Tc*PRACB (*open symbols*). After 30 min at 37 °C, the reaction was stopped by heat inactivation and freezing. *D*, 39-kDa intracellular isoform was isolated from soluble (*Ese*) extracts of non-infective epimastigote forms of the parasite. Western blots of serial dilutions of the soluble suspension were compared with known amounts of *rTcPRACB* protein and used for protein quantitation using Quantity One® software. Racemase assays were performed in sodium acetate buffer, pH 6, using 40 mM L-proline and the equivalent depicted amounts of 39 kDa (ng) contained in Ese extract.

standard PCR amplification was performed using the *TcPRACA* DNA as template and a mixture of act-1 primer and the reverse C terminus primer, Bg-45 (5′ CTG AGC TCG ACC AGA T(C/A)T ACT GC 3′) (codon 423) or a mixture of act-2 primer and the forward N terminus primer, Hi-45 (5′ CTC TCC CAT GGG GCA GGA AAA GCT TCT G 3′) (codon −53) (see Fig. 5). Resulting amplified fragments of 316 and 918 bp, respectively, were purified by a Qiagen PCR extraction kit as prescribed and further ligated by T4 ligase to generate a template consisting of the full-length of a potentially mutated *TcPRACA** coding sequence used for the second step PCR. Amplification of this template was performed using forward Hi-45 and reverse Bg-45 primers, and the resulting *TcPRACA** fragment encoding for the mature proline racemase was purified and cloned in pCR®2.1-TOPO® vector (Invitrogen). TOP10 competent *E. coli* were transformed with the pCR®2.1-TOPO®-*TcPRACA** construct and plasmid DNA isolated from individual clones prepared for DNA sequencing. Positive mutants were then subcloned in-frame with a C-terminal His_6 tag into the *Nco*I/*Sac*I sites of the pET 28b(+) expression vector (Novagen-Tebu, Le Parray en Yvelines, France). Subclones of pET28b(+)-*TcPRACA** produced in *E. coli* (DH5α) were sequenced again to confirm the presence of the mutation (not shown). Soluble recombinant ^{C330S}*Tc*PRACA protein was produced in *E. coli* BL21(DE3) (Invitrogen) and purified using a nickel column (Novagen-Tebu), according the manufacturer's instructions.

RESULTS

Expression of a Functional Intracellular Isoform of Proline Racemase—We have previously characterized (13) a *TcPRAC* gene from *T. cruzi* and demonstrated *in vivo* and *in vitro* that it encodes a proline racemase enzyme. Analysis of the genomic organization and transcription of the *TcPRAC* gene indicated the presence of two paralogue gene copies per haploid genome, named *TcPRACA* (GenBank™ accession number AF195522) and *TcPRACB* (GenBank™ accession number AY140947). We showed that *TcPRACA* encodes a functional co-factor-independent proline racemase, closely resembling the *C. sticklandii* proline racemase (CsPR) (GenBank™ accession number E101199) (11). We now sequenced the full-length of *TcPRACB*, and, as can be observed in Fig. 1*A*, *TcPRACA* and *TcPRACB* genes both possess the characteristic trypanosome polypyrimidine-rich motifs in the intergenic region that are crucial *trans*-splicing signals when located upstream of an (AG)-dinucleotide used as acceptor site. As in other *T. cruzi* genes, UUA triplets are found at the end of the 3′ untranslated region preceding the

arrow; initiation of translation for *Tc*PRACA and *Tc*PRACB are shown by *single-headed arrows*; nucleotides are shaded in *light* and *dark gray* and represent silent mutations or point mutations, respectively; the proline racemase active site is in a *box*; UUA triplets are *underlined in bold* and precede polyadenylation sites that are *double-underlined*. *B*, schematic representation of amino acid sequence alignments of *T. cruzi Tc*PRACA (*Tc-A*) and *Tc*PRACB (*Tc-B*) proline racemases. The common scale is in amino acid residue positions along the linear alignment. ▼ and ▲ represent the initiation codons for *Tc*PRACA and *Tc*PRACB proteins, respectively; ▽ represents an alternative *Tc*PRACA putative initiation codon. Amino acid differences are indicated *above* and *below* the *vertical lines*, and their positions in the sequence are shown in *parentheses*. *SP*, signal peptide; the N-terminal domain of *Tc*PRACA extends from positions 1 to 69. *SPCGT*, conserved active sites of *Tc*PRACA and *Tc*PRACB proline racemases; N terminus and C terminus are indicated for both proteins. *C*, hydrophobicity profile of *Tc*PRACA. *Dotted line* depicts the cleavage site as predicted by Von Heijne's method (amino acids 31–32). *D*, ethidium bromide-stained gel of chromosomal bands of *T. cruzi* CL Brener clone after separation by pulsed-field gel electrophoresis (*lane 1*) and Southern blot hybridization with *TcPRAC* probe (*lane 2*). The sizes (*Mb*) of chromosomal bands are indicated, as are the region chromosome numbers (in *roman numerals*).

TABLE I
Stability of recombinant TcPRACA and TcPRACB proline racemases under different storage conditions

After purification on nickel-nitrilotriacetic acid-agarose column, recombinant proteins were kept for 10 days in nickel column buffer (20 mM Tris/500 mM NaCl/500 mM imidazol, pH 8.0) at room temperature (RT) or at +4 °C or diluted either in 50% glycerol and maintained at −20 °C (Gly/−20 °C) or in optimum pH buffer (NaOAc, pH 6.0) at 4 °C. Recombinant enzymes were precipitated in $(NH_4)_2SO_4$ and kept in solution at 4 °C or pellet dried at −20 °C. Racemase assays were performed for 30 min at 37 °C. Percent of preservation was determined polarimetrically using 0.25 μM of either purified r*Tc*PRACA or r*Tc*PRACB enzymes and 40 mM of L-proline, as compared with results obtained with freshly purified proteins (CTRL). This results are representative of at least two independent experiments.

Protein	% of preservation of proline racemase activity						
	Column				NaOAc, pH6	$(NH_4)_2SO_4$	
	CTRL	RT	+4 °C	Glyl-20 °C	4 °C	4 °C	−20 °C
r*Tc*PRACA	100.0	16.0	66.5	62.9	31.0	53.9	100.0
r*Tc*PRACB	100.0	100.0	34.0	93.6	77.6	98.4	100.0

polyadenylation site. Comparison between the two sequences revealed 14 point mutations (resulting in 96% identity) giving rise to seven amino acid differences. When expressed, the *Tc-PRACB* is predicted to produce a shorter protein (39 kDa) whose translation would start at the ATG codon at position 274 located downstream of the (AG)-spliced leader acceptor site (at position 175). In comparison, *TcPRACA* has an open reading frame that encodes a peptide with an apparent molecular mass of 45 kDa. The schematic protein sequence alignment of the two proteins *TcPRACA* and *TcPRACB* depicted in Fig. 1*B* reveals that *Tc*PRACB proline racemase lacks the amino acid sequence corresponding to the signal peptide observed in the *Tc*PRACA protein (*hatched box* in the figure; see predicted cleavage site in Fig. 1*C*). Therefore the *TcPRACB* would produce a 39-kDa, intracellular, and non-secreted isoform of the protein. As with CsPR (11) and *Tc*PRACA (13) (Fig. 1*B*), the active site of proline racemase is conserved in *TcPRACB* sequence. Furthermore, while differing by only seven amino acids, both the *TcPRACA* and *TcPRACB* sequences display around 50% homology to the CsPR (13). In accordance with other protein-coding genes in *T. cruzi*, *TcPRAC* genes are located on two different chromosomal bands of which one contains three or more chromosomes of similar size; see Fig. 1*D*. Thus, hybridization of blots containing *T. cruzi* CL Brener chromosomal bands separated by pulsed field gel electrophoresis revealed that sequences recognized by an homologous probe to both *TcPRACA* and *TcPRACB* are mapped in neighboring migrating bands of ~0.9 and 0.8 Mb, corresponding to regions VII and V, respectively, according to the numbering system of Cano *et al.* (14).

To verify whether the *TcPRACB* gene could encode a functional proline racemase, we expressed both *T. cruzi* paralogs in *E. coli* to produce C-terminal His_6-tagged recombinant proteins. After purification by affinity chromatography on nickel-nitrilotriacetic acid-agarose column, recombinant proteins were separated by SDS-gel electrophoresis revealing single bands with the expected sizes of 45.8 and 40.1 kDa for the r*Tc*PRACA and r*Tc*PRACB proteins, respectively (Fig. 2*A*). To determine whether r*Tc*PRACB displays proline racemase enzymatic activity, biochemical assays were employed to measure the shift in optical rotation of L- and D-proline substrates, as described (13). As can be seen in Fig. 2*B*, r*Tc*PRACB racemizes both L- and D-proline but not L-hydroxy-proline, like r*Tc*PRACA. In a similar manner, r*Tc*PRACB is a co-factor-independent proline racemase as described for CsPR (11) and r*Tc*PRACA (13) proline racemases. The rate of conversion of L- into D-proline was measured at various pH values using both recombinant enzymes. As illustrated in Fig. 2*C*, r*Tc*PRACA activity clearly shows a pH dependence with an optimal activity from pH 5.5 to 7.0. In contrast, the optimum activity of r*Tc*PRACB can be observed in a large pH spectrum varying from pH 4.5 to 8.5. These results revealed that translation of the open reading frame of both *TcPRAC* genes copies result in functional proline racemase isoforms. As described previously (13), Western blot analysis of non-infective epimastigote parasite extracts using antibodies raised against the 45-kDa secreted proline racemase had previously revealed a 39-kDa protein mostly in the soluble cellular fraction, only weakly in the cellular insoluble fraction, and absent from culture medium. To demonstrate that the intracellular 39-kDa isoform of the protein was equally functional *in vivo*, soluble cellular extracts were obtained from 5×10^8 epimastigote, non-infective parasites, and the levels of 39-kDa soluble protein were quantified by Western blot comparatively to known amounts of *rTcPRACB* enzyme. As can be observed in Fig. 2*D*, the intracellular isoform of the protein is indeed functional *in vivo*, because proline racemase enzymatic activity was displayed, and levels of racemization were dependent on protein concentration.

Functional Analysis and Kinetic Properties of Recombinant T. cruzi Proline Racemases—Because the *TcPRAC* gene copies encode for secreted and non-secreted isoforms of proline racemase with distinct pH requirements for activity, we investigated whether other biochemical properties differ between r*Tc*PRACA and r*Tc*PRACB proteins. Such differences might reflect the cellular localization of the protein during parasite differentiation and survival in the host. Both r*Tc*PRACA and r*Tc*PRACB enzyme activities are maximal at 37 °C and can be abolished by heating for 5 min at 80 °C. However, the stability of the two recombinant enzymes differs considerably when analyzed under different storage conditions. Thus, as shown in Table I, purified r*Tc*PRACB is highly stable, because its activity is maintained for at least 10 days at room temperature in 0.5 M imidazole buffer, pH 8.0, as compared with r*Tc*PRACA, which loses 84% of its activity under such conditions. In contrast, most of the enzymatic activity of r*Tc*PRACA is maintained at 4 °C (65%), compared with that of r*Tc*PRACB (34%). Both enzymes can be preserved in 50% glycerol at −20 °C or diluted in sodium acetate buffer at pH 6.0, but under these storage conditions r*Tc*PRACA activity is impaired. However, best preservation of both recombinant proline racemases was undoubtedly obtained when proteins were kept at −20 °C as ammonium sulfate precipitates.

Both recombinant enzymes exhibited Michaelis-Menten kinetics (Fig. 3*A*), and r*Tc*PRACB had a higher activity than r*Tc*PRACA. Indeed, as can be observed in Fig. 3*B*, analysis of L → D conversion of serial dilutions of L-proline catalyzed by a constant amount of each enzyme showed that r*Tc*PRACB enzyme (K_m of 75 mM and V_{max} of 2×10^{-4} mol·s^{-1}) has a higher velocity as compared with r*Tc*PRACA (K_m of 29 mM and V_{max} of 5.3×10^{-5} mol·s^{-1}). To determine the K_i values for PAC, the specific and competitive inhibitor of CsPR (16), assays were performed with both recombinant proteins. These assays revealed that PAC is comparably effective as inhibitor of r*Tc*PRACA (Fig. 3*C*) and r*Tc*PRACB (not shown), and K_i values obtained were 5.7 and 3.6 μM, respectively. The difference in K_i

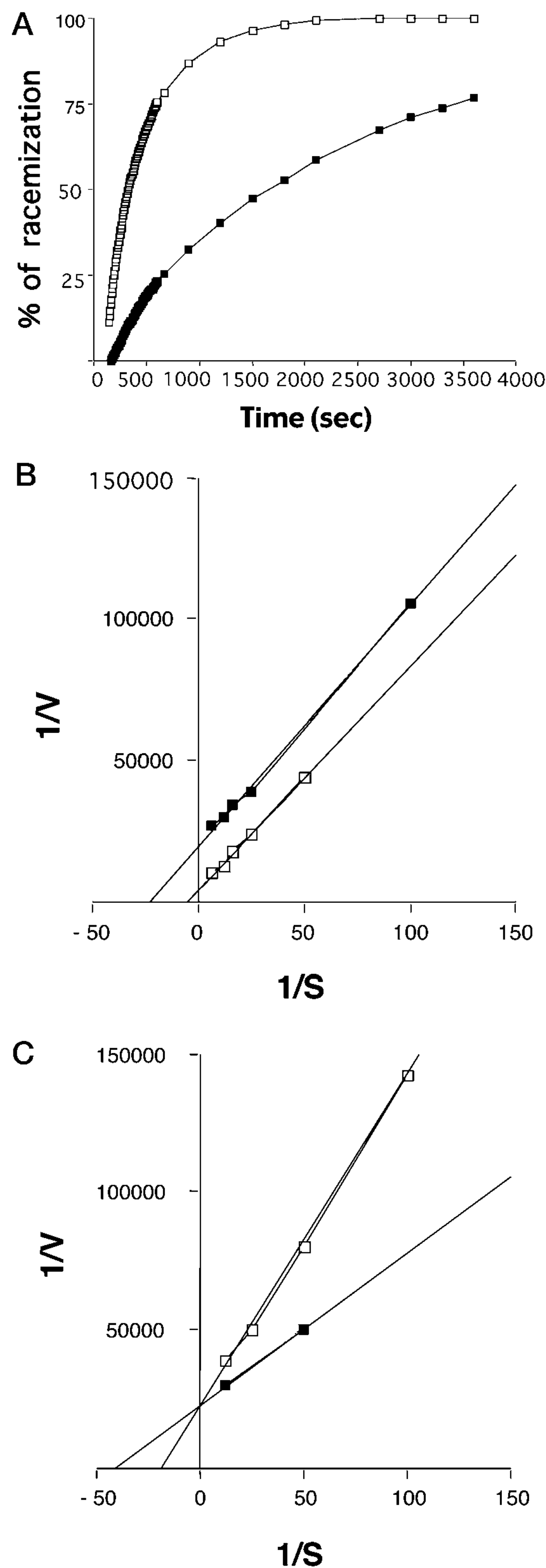


FIG. 3. **Kinetic parameters of L-proline racemization catalyzed by r*Tc*PRACA and r*Tc*PRACB proline racemase isoforms.** The progress of racemization reaction was monitored polarimetrically, as described previously (13). *A*, the determination of the linear part of the curve was performed at 37 °C in medium containing 0.2 M sodium acetate, pH 6.0, 0.25 μM purified enzyme, and 40 mM L-proline. r*Tc*PRACA reactions are represented by *black squares*, and r*Tc*PRACB reactions are represented by *white squares*. *B*, initial rate of racemase activity was assayed at 37 °C in medium containing 0.2 M sodium acetate, pH 6.0, 0.125 μM r*Tc*PRACA (*solid squares*), or r*Tc*PRACB (*open squares*) purified enzymes and different concentrations of L-proline. Lineweaver-Burk double reciprocal plots were used to determine values for K_m and V_{max} where 1/V is plotted in function of 1/[S], and the slope of the curve represents K_m/V_{max}. Values obtained were confirmed by using the Kaleïdagraph® program and Michaelis-Menten equation. The values are representative of six experiments with different enzyme preparations. *C*, double reciprocal plot kinetics of 0.125 μM r*Tc*PRACA proline racemase isoform in the presence (*open*) or absence (*closed*) of 6.7 μM PAC competitive inhibitor in function of L-proline concentration. For comparison, K_m reported for the proline racemase of *C. sticklandii* was 2.3 mM; kinetic assays using the native protein obtained from a soluble epimastigote fraction revealed a K_m of 10.7 mM and a K_i of 1.15 μM.

values reflects almost perfectly the difference in K_m values reported for both enzymes, which are similar to that of the native protein (not shown). These K_i values indicate that the affinity of PAC inhibitor is higher for r*Tc*PRACA and r*Tc*-PRACB than for CsPR (K_i of 18 μM).

Requirement of a Dimeric Structure for Proline Racemase Activity—When r*Tc*PRACA was submitted to size exclusion chromatography on a Superdex 75 column at pH 6.0, two peaks of protein were eluted around 80 kDa (B2 fraction) and 43 kDa (B4 fraction), respectively, presumably corresponding to dimeric and monomeric forms of the enzyme (Fig. 4). Western blot analysis of whole *T. cruzi* epimastigote extracts using non-denaturing PAGE had previously indicated a molecular mass of 80 kDa for the native protein (not shown), whereas a 45-kDa band was obtained by SDS-PAGE (13). To eliminate cross-contamination, B1 and B5 fractions, eluted at the start and at the end of the predicted dimer (B2) or monomer (B4) peaks, respectively, were reloaded onto the column, and the profiles obtained (see Fig. 4, *insets*) confirmed the purity of the fractions. Enzyme activity resides in the 80-kDa peak but not in the 43-kDa peak (Table II). These results corroborated that two subunits of the protein are necessary for racemase activity. At neutral pH (7.4 or above), the r*Tc*PRACA gives rise to high molecular weight aggregates that are not observed with r*Tc*PRACB, consistently with its broader optima pH spectrum (not shown).

Abrogation of Proline Racemase Activity by Mutation of Cys^{330} *of the Catalytic Site*—*C. sticklandii* proline racemase is described as a homodimeric enzyme with subunits of 38 kDa and a single proline binding site for every two subunits, where two cysteines at position 256 might play a crucial role in catalysis by the transfer of protons from and to the bound substrate (12). We have shown previously (16) that mitogenic properties of the *T. cruzi* proline racemase are dependent on the integrity of the enzyme active site, as inhibition of B-cell proliferation is obtained by substrate competition and specific use of analogues (PAC) resembling the structure assumed by the substrate proline in its transition state. To verify the potential role of the cysteine residues at the active site of the *T. cruzi* proline racemase, we replaced Cys^{330} by a serine residue through site-specific mutation of *TcPRACA*. The choice of serine as the substituting amino acid was made to avoid further major disturbances on three-dimensional structure of the protein (see strategy in Fig. 5 and "Experimental Procedures"). After confirmation of the single codon mutation through sequencing of the construct (not shown), the ^{C330S}r*Tc*PRACA mutant proline racemase was expressed in *E. coli* and purified in the same manner as wild-type r*Tc*PRACA. We then used ^{C330S}r*Tc*PRACA in racemization assays to verify the effects of the mutation on the enzymatic activity of the protein. As can be observed in Table III, a total loss of proline racemase activity is observed as compared with the wild-type enzyme, establishing that proton transfer during proline racemization is specifically dependent on the presence of the cysteine residue in the active site.

Proline Racemase Protein Signatures and Putative Proline Racemases in Sequence Databases—The conservation of critical

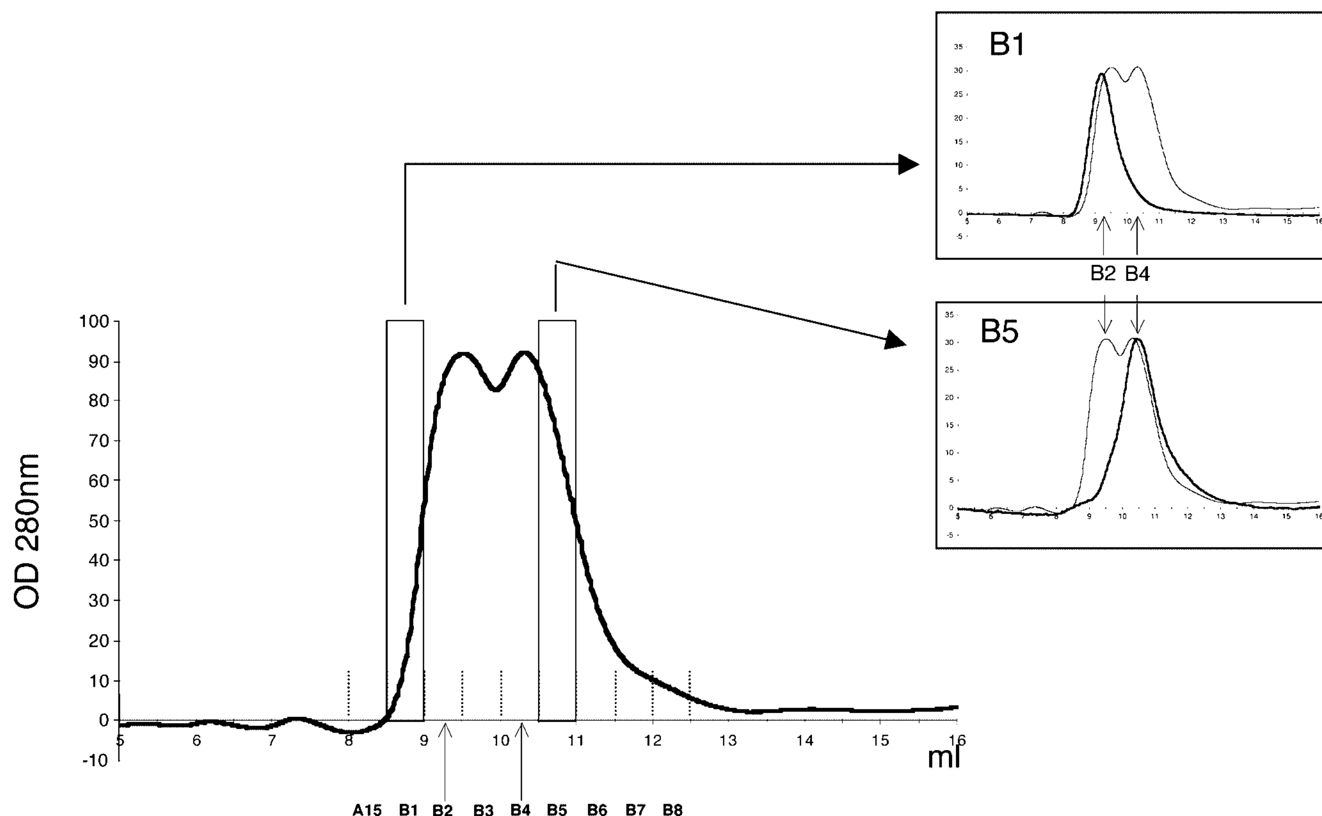


FIG. 4. **Size exclusion chromatography of r*Tc*PRACA protein using a Superdex 75 column.** Fractions were eluted by high pressure liquid chromatography at pH 6.0; *B2* and *B4* peaks correspond to r*Tc*PRACA dimer and monomer species, respectively. *B1* and *B5* eluted fractions were reloaded onto the column (*bold*; see *insets*) using the same conditions and compared with previous elution profile (*not bold*).

TABLE II
Racemase activity of recombinant TcPRACA fractions after size exclusion chromatography

After elution from Superdex 75 column, 20 μl of each peak (A15 to B7; see Fig. 4) corresponding to 1 μg of protein were incubated for 1 h at 37 °C with 40 mM L-proline in 0.2 M NaOAc, pH 6.0. Optical rotation was measured, and % of racemization was determined as described under "Experimental Procedures."

Fractions	A15	B1	B2	B3	B4	B5	B6	B7
% racemization	1.3	35.5	62.9	42.8	0.7	0	0	0

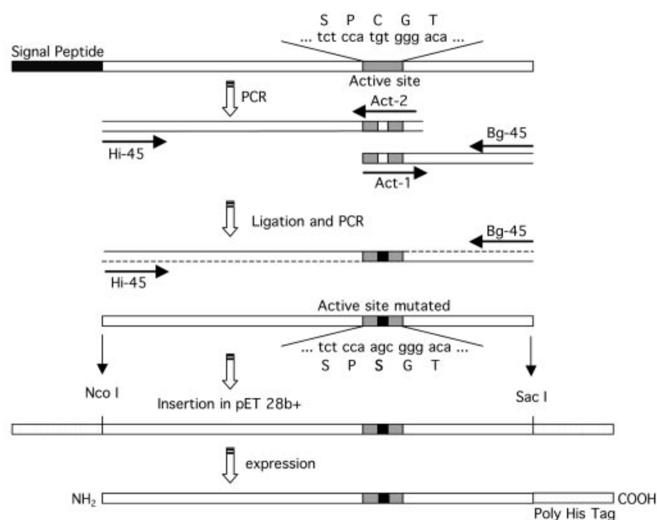


FIG. 5. **Site-directed mutagenesis of *Tc*PRACA proline racemase**. Schematic representation of the active site mutagenesis of proline racemase of *TcPRACA* gene.

residues between parasite and bacterial proline racemases prompted us to search for similarities between *Tc*PRAC and other protein sequences in SWISS-PROT and TrEMBL databases. Twenty-one protein sequences yielded significant homologies, from 11 organisms, such as several proteobacteria of the α-subdivision (*Agrobacterium*, *Brucella*, *Rhizobium*) and γ-subdivision (*Xanthomonas* and *Pseudomonas*), as well as of the fermicutes (*Streptomyces* and *Clostridium*). Within the eukaryota, besides in *T. cruzi*, homologous genes were detected in the human and mouse genomes, where predicted proteins show overall similarities with proline racemase. Except for *C. sticklandii* and *Xantomonas campestri*, each other organism encodes two paralogs, and *Agrobacterium tumefaciens* contains three genes. The multiple alignment also allowed for the definition of three signatures of proline racemase, which are described here in PROSITE format. As can be seen in Table IV, when using a minimal motif of proline racemase protein (M I), [IVL][GD]*X*H*XX*G[ENM]*XX*[RD]*X*[VI]*XX*G, located immediately after the start codon at position 79, we obtained nine hits. A second motif (M II), consisting of [NSM][VA][EP][AS][FY]*X*-(13, 14)[GK]*X*[IVL]*XX*D[IV][AS][YWF]GG*X*[FWY], starting at position 218, gave 14 hits; however, the first or the second half of this motif is not sufficiently stringent to be restrictive for putative proline racemases but gives hits for different protein families. A third motif (M III), from positions 326 to 339, namely DRSP*X*G*X*[GA]*XX*A*XX*A, was considered as a minimal pattern. Note that in position 330, the cysteine of the active site was replaced by an *X*. As shown in Table IV, this minimal

Table III
Loss of racemase enzymatic activity in the site-directed ^{C330S}TcPRACA

After purification, 5 μg of r*Tc*PRACA or ^{C330S}r*Tc*PRACA were incubated at 37 °C with 40 mM L-proline in NaOAc buffer, pH 6.0. Optical rotation was measured at different times and % of racemization was determined as described under "Experimental Procedures."

Data set	r*Tc*PRACA				^{C330S}r*Tc*PRACA			
Time (*min*)	0	10	30	60	0	10	30	60
Optical rotation	−0.385	−0.300	−0.162	−0.088	−0.385	−0.382	−0.391	−0.387
% Racemization	0	22	58	77	0	0	0	0

Table IV
SWISS-PROT and TrEMBL databases screening using PROSITE motifs

SWISS-PROT and TrEMBL databases were screened using motifs I to III (M I, M II and M III). M I corresponds to [IVL][GD]*X*H*XX*G[ENM]-*XX*[RD]*X*[VI]*XX*G, M II to of [NSM][VA][EP][AS][FY]*X*(13, 14)[GK]*X*[IVL]*XX*D[IV][AS][YWF]GG*X*[FWY], M III to DRSP*X*G*X*G*XX*A*XX*A, and M III* to DRSPCG*X*G*XX*A*XX*A. Access. nb, SWISS-PROT accession number of the sequence; Seq. sequence number according to Fig. 6; + and −, presence or absence, respectively, of hit using the corresponding motif.

Organism	Seq	Access. nb	Motif			
			M I	M II	M III	M III*
Agrobacterium tumefaciens	1	Q8UIA0	+	+	+	+
Agrobacterium tumefaciens	2	Q8U6X2	−	−	+	−
Agrobacterium tumefaciens	3	Q8U8Y5	−	−	+	−
Brucella melitensis	1	Q8YJ29	−	+	+	+
Brucella melitensis	2	Q8YFD6	+	−	+	−
Clostridium stickilandii		Q9L4Q3	−	+	+	+
Homo sapiens	1	Q96EM0	+	+	+	−
Homo sapiens	2	Q96LJ5	+	+	+	−
Mus musculus	1	Q9CXA2	+	+	+	−
Mus musculus	2	Q99KB5	+	+	+	−
Pseudomonas aeruginosa	1	Q9I476	−	+	+	+
Pseudomonas aeruginosa	2	Q9I489	−	−	+	+
Rhizobium loti	1	Q98F20	−	+	+	+
Rhizobium loti	2	Q988B5	+	+	+	−
Rhizobium meliloti	1	Q92WR9	−	−	+	−
Rhizobium meliloti	2	Q92WS1	−	+	+	+
Streptomyces coelicolor		Q93RX9	+	−	+	+
Trypanosoma cruzi	1	Q9NCP4	+	+	+	+
Trypanosoma cruzi	2	[a]	+	+	+	+
Xanthomonas axonopodis	1	Q8PJI1	−	+	+	+
Xanthomonas axonopodis	2	Q8PKE4	−	−	+	+
Xanthomonas campestris		Q8P833	−	+	+	+

[a] GenBank™ accession number AY140947.

pattern yields all 21 hits. Curiously, both genes in human, as well as in mouse, encode threonine instead of cysteine at the *X* position in motif III, whereas in *Brucella*, *Rhizobium*, and *Agrobacterium* species, each encode one protein with C and one with T in this position. We cannot hypothesize the implications of this substitution for the functionality of these putative proteins. If the residue at position 330 is maintained as a cysteine in motif III, a reduced number of 12 hits from nine organisms is thus obtained, which can probably be considered as true proline racemases. The alignment of the 21 protein sequences and derived cladogram are shown in Fig. 6 and Fig. 7, respectively, and the three *boxes* depicted correspond to motifs I, II, and III described above. We thus propose DRSPCG*X*G*XX*A*XX*A as the minimal signature for proline racemases. BLAST searches against unfinished genomes yielded, at present, an additional 13 predicted protein sequences from 9 organisms, with high similarity to proline racemases, all containing motif III. Organisms are *Clostridium difficile*, *Clostridium botulinum*, *Bacillus anthracis*, *Brucella suis*, *Pseudomonas putida*, *Rhodobacter sphaeroides*, *Burkholderia pseudomallei*, *Burkholderia mallei*, and the fungus *Aspergillus fumigatus*. These results indicate that proline racemases might be quite widespread.

DISCUSSION

Proline racemase, an enzyme previously only described in protobacterium *C. sticklandii* (11), was shown to be encoded also by the eukaryote *T. cruzi*, a highly pathogenic protozoan parasite (13). The *Tc*PRAC (*T. cruzi* proline racemase), formerly called TcPA45, is an efficient mitogen for host B cells and is secreted by the metacyclic forms of the parasite upon infection, contributing to its immune evasion and persistence through nonspecific polyclonal lymphocyte activation (13). Our previous results (13) suggested that *Tc*PRAC is encoded by two paralogous genes per haploid genome. Protein localization studies have also indicated that *T. cruzi* can differentially express intracellular and secreted versions of *Tc*PRAC during cell cycle and differentiation, as the protein is found in the cytoplasm of non-infective replicative (epimastigote) forms of the parasite, and bound to the membrane or secreted in the infective, non-replicative (metacyclic trypomastigote) parasites (13). Here we have characterized the two *TcPRAC* paralogs and demonstrated that both *TcPRACA* and *TcPRACB* give rise to functional isoforms of co-factor-independent proline racemases, which display different biochemical properties that may well have important implications in the efficiency of the respective enzymatic activities. As suggested previously (11, 17, 18) by biochemical and theoretical studies for the bacterial proline racemase, our studies reveal that *Tc*PRAC activities rely on two monomeric enzyme subunits that perform interconversion of L- and/or D-proline enantiomers by a two-base mechanism reaction in which the enzyme removes an α-hydrogen from the substrate and donates a proton to the opposite side of the α-carbon. It has been predicted that each subunit of the homodimer furnishes one of the sulfhydryl groups (18). In the present work we showed that *Tc*PRAC enzymatic activities are *bona fide* dependent on the Cys^{330} residue of the active site, as site-specific C330S mutation totally abrogates L- and D-proline racemization, in agreement with our previous demonstration

FIG. 6. **Sequence alignments of proteins (Clustal X) obtained by screening SWISS-PROT and TrEMBL databases using motifs I, II, and III.** Amino acids involved in MI, MII, and MIII are shaded in *dark gray*. The 13–14 unspecific amino acids involved in M II are shaded in *light gray*. *SWISS-PROT* accession numbers of the sequences are in Table IV.

(13) that *Tc*PRAC enzymatic activity is abolished through alkylation with iodoacetate or iodoacetamine, similarly to the *Clostridium* proline racemase, where carboxymethylation was shown to occur specifically with the two cysteines of the active site leading to enzyme inactivation (12). Although gene sequence analysis predicted that by a mechanism of alternative splicing *TcPRACA* could generate both intracellular and secreted versions of parasite proline racemase, our present stud-

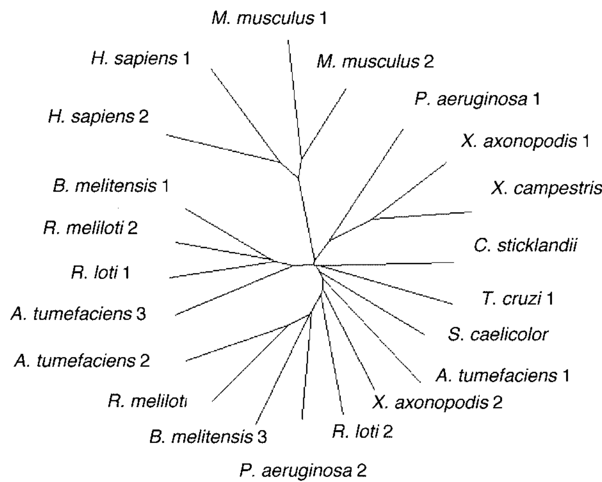


FIG. 7. **Cladogram of protein sequences obtained by T-coffee alignment radial tree.** See Table IV for SWISS-PROT protein accession numbers.

ies demonstrate that *TcPRACB* gene sequence *per se* codes for a protein lacking the amino acids involved in peptide signal formation and an extra N-terminal domain present in TcPRACA protein, resembling more closely the CsPR. Thus, *TcPRACB* can only generate an intracellular version of *Tc*PRAC proline racemase.

Interestingly, the presence of two homologous copies of *TcPRAC* genes in the *T. cruzi* genome, coding for two similar but with distinct specific biochemical properties, could reflect an evolutionary mechanism of gene duplication and a parasite strategy to ensure a better environmental flexibility. This assumption is comforted by the potential of *TcPRACA* gene to generate two related protein isoforms by alternative splicing, a mechanism that is particularly adept for cells that must respond rapidly to environmental stimuli. Primarily, *trans*-splicing appears indeed to be an ancient process that may constitute a selective advantage for split genes in higher organisms (19), and alternative *trans*-splicing was only proven to occur in *T. cruzi* recently (20). As an alternative for promoter selection, the regulated production of intracellular and/or secreted isoforms of proline racemase in *T. cruzi* by alternative *trans*-splicing of *TcPRACA* gene would allow the stringent conservation of a constant protein domain and/or the possibility of acquisition of an additional secretory region domain. As a matter of fact, our recent investigations using RT-PCR based strategy and a common 3′ probe to *TcPRACA* and *TcPRACB* sequences combined to a 5′ spliced leader oligonucleotide followed by cloning and sequencing of the resulting fragments have indeed proved that an intracellular version of *Tc*PRAC may also originate from the *TcPRACA* gene, corroborating this hypothesis.[2,3]

Gene duplication is a relatively common event in *T. cruzi* that adds complexity to parasite genomic studies. Moreover, *TcPRAC* chromosomal mapping revealed two chromosomal bands that possess more than three chromosomes each and that may indicate that proline racemase genes are mapped in size-polymorphic homologous chromosomes, an important finding for proline racemase gene family characterization. Preliminary results in this laboratory have, for instance, revealed that *T. cruzi* DM28c type I strain maps proline racemase genes to the same chromoblot regions identified with *T. cruzi* CL type II strain used in the present work. Other isolates of the parasites are presently under investigation (data not shown).

It is well known that proline constitutes an important source of energy for several organisms, such as several hemoflagellates (21–23), and for flight muscles in insects (24). Furthermore, a proline oxidase system was suggested in trypanosomes (25), and the studies reporting the abundance of proline in triatominae guts (26) have implicated proline in metabolic pathways of *T. cruzi* parasites, as well as in its differentiation in the digestive tract of the insect vector (27). Thus, it is well accepted that *T. cruzi* can use L-proline as a principal source of carbon (25). Moreover, our preliminary results using parasites cultured in defined media indicate that both epimastigotes, found in the vector, and infective metacyclic trypomastigote forms can efficiently metabolize L- or D-proline as the sole source of carbon (not shown). Although certain reports indicate that biosynthesis of proline occurs in trypanosomes, *i.e.* via reduction of glutamate carbon chains or transamination reactions, our results reveal that an additional and direct physiological regulation of proline might exist in the parasite to control amino acid oxidation and its subsequent degradation or yet to allow proline utilization. In fact, a recent report (28) showed two active proline transporter systems in *T. cruzi*. We suggest that *T. cruzi* proline racemase may possibly play a consequential role in the regulation of intracellular proline metabolic pathways, or else it could participate in mechanisms of post-translational addition of D-amino acid to polypeptide chains. On one hand, these hypotheses would allow for an energy gain and, on the other hand, would permit the parasite to evade host responses. In this respect, it was reported that a single D-amino acid addition in the N terminus of a protein is sufficient to confer general resistance to lytic reactions involving host proteolytic enzymes (29). The expression of proteins containing D-amino acids in the parasite membrane would benefit the parasite inside host cell lysosomes, in addition to the contribution to the initiation of polyclonal activation, as described previously (30, 31) for polymers composed of D-enantiomers. Although D-amino acid inclusion in *T. cruzi* proteins would benefit the parasite, this hypothesis remains to be proven, and direct evidences are technically difficult to obtain.

It is worth noting that metacyclogenesis of epimastigotes into infective metacyclic forms involves parasite morphologic changes that include the migration of the kinetoplast, a structure that is physically linked to the parasite flagellum, and many other significant metabolic alterations that combine to confer infectivity/virulence to the parasite (13, 32). Proline racemase was shown to be preferentially localized in the flagellar pocket of infective parasite forms after metacyclogenesis (13), as are many other known proteins secreted and involved in early infection (33). It is also conceivable that parasite proline racemase may function as an early mediator for *T. cruzi* differentiation through intracellular modification of internalized environmental free proline, as suggested above and already observed in some bacterial systems. As an illustration, exogenous alanine has been described as playing an important role in bacterial transcriptional regulation by controlling an operon formed by genes coding for alanine racemase and a smaller subunit of bacterial dehydrogenase (34). In bacteria, membrane alanine receptors are responsible for alanine and proline entry into the bacterial cell (35). We can then hypothesize that the availability of proline in the insect gut milieu associated to a mechanism of environmental sensing by specific receptors in the parasite membrane would stand for parasite proline uptake and its further intracellular racemization. Proline racemase would then play a fundamental role in the reg-

[2] N. Chamond, N. Coatnoan, J. C. Barale, A. Cosson, A. Berneman, W. Degrave, and P. Minoprio, manuscript in preparation.

[3] The proline racemase/B-cell mitogen of *T. cruzi* is a virulence factor whose mRNA is regulated differentially through development by alternative splicing.

ulation of parasite growth and differentiation by its participation in both metabolic energetic pathways and the expression of proteins containing D-proline, as described above, consequently conferring parasite infectivity and its ability to escape host-specific responses.

Thus far, and contrasting to the intracellular isoform of *Tc*PRAC found in epimastigote forms of *T. cruzi*, the ability of metacyclic and bloodstream forms of the parasite to express and secrete proline racemase may have further implications in host/parasite interaction. In fact, the parasite-secreted isoform of proline racemase participates actively in the induction of nonspecific polyclonal B-cell responses upon host infection (13) and favors parasite evasion, thus ensuring its persistence in the host. As described for other mitogens and parasite antigens (36–38), and in addition to its mitogenic property, *Tc*PRAC could also be involved in modifications of host cell targets enabling better parasite attachment to host cell membranes, in turn assuring improved infectivity. Because several reports associate accumulation of L-proline with muscular dysfunction (39) and inhibition of muscle contraction (40), the release of proline racemase by intracellular parasites could alternatively contribute to the maintenance of infection through regulation of L-proline concentration inside host cells, as proline was described as essential for the integrity of muscular cell targets. Therefore, we have demonstrated recently that transgenic parasites hyperexpressing *TcPRACA* or *TcPRACB* genes, but not functional knock-outs, are five to ten times more infective to host target cells, pointing to a critical role of proline racemases in the ongoing of the infectious process.[2,3] Likewise, previous reports (41) demonstrated that genetic inactivation of *Lysteria monocytogenes* alanine racemase and D-amino acid oxidase genes abolishes bacterial pathogenicity, because the presence of D-alanine is required for the synthesis of the mucopeptide component of the cell wall that protects virtually all bacteria from the external milieu.

Present analysis using identified critical conserved residues in *TcPRAC* and *C. sticklandii* proline racemase genes and the screening of SWISS-PROT and TrEMBL databases led us to define of a minimal signature for proline racemases, DRSP*X*G*X*[GA]*XX*A*XX*A, and to confirm the presence of putative proteins in at least ten distinct organisms. Screening of unfinished genome sequences showed highly homologous proline racemase candidate genes in an additional 9 organisms, among which are the fungus *A. fumigatus* and the bacteria *B. anthracis* and *C. botulinum*. This is of particular interest, because racemases, but not proline racemases, are widespread in bacteria and only recently described in more complex organisms such as *T. cruzi* (42, 43). These findings may possibly reflect cell adaptative responses to extracellular stimuli and uncover more general mechanisms for the regulation of gene expression by D-amino acids in eukaryotes. Our finding of similar genes in human and mouse genome databases when we used less stringent signatures for proline racemase is striking. However, the absence of the crucial amino acid cysteine in the putative active site of those predicted proteins suggests a different functionality than that of a proline racemase.

Finally, we described here that *Tc*PRAC isoforms are highly stable and have the capacity to perform their activities across a large spectrum of pH. In addition, the affinity of pyrrolcarboxylic acid, a specific inhibitor of proline racemase, is higher for *Tc*PRAC enzymes than for CsPR. Because proline racemase is a protein involved in the mechanisms leading to *T. cruzi* immune evasion, our work should stimulate further studies on biochemical and molecular characterization of putative proline racemases in other microorganisms and draw attention to their importance as potential targets for drug development.

Acknowledgments—We thank A. Waters for critical discussions, A. Cosson for help with the recombinant proteins, A. Berneman for advice, L. Mulard for access to the polarimeter facilities, and F. Lacoste for constant support of N. C.

REFERENCES

1. Lamzin, V. S., Dauter, Z., and Wilson, K. S. (1995) *Curr. Opin. Struct. Biol.* **5,** 830–836
2. Kleinkauf, H., and von Dohren, H. (1987) *Annu. Rev. Microbiol.* **41,** 259–289
3. Fisher, G. H. (1998) *Exper. Suppl.* **85,** 109–118
4. Nagata, Y., Fujiwara, T., Kawaguchi-Nagata, K., Fukumori, Y., and Yamanaka, T. (1998) *Biochim. Biophys. Acta* **1379,** 76–82
5. Nagata, Y., Tanaka, K., Iida, T., Kera, Y., Yamada, R., Nakajima, Y., Fujiwara, T., Fukumori, Y., Yamanaka, T., Koga, Y., Tsuji, S., and Kawaguchi-Nagata, K. (1999) *Biochim. Biophys. Acta* **1435,** 160–166
6. Oguri, S., Kumazaki, M., Kitou, R., Nonoyama, H., and Tooda, N. (1999) *Biochim. Biophys. Acta* **1472,** 107–114
7. Neidle, A., and Dunlop, D. S. (1990) *Life Sci.* **46,** 1512–1522
8. Schell, M. J., Molliver, M. E., and Snyder, S. H. (1995) *Proc. Nat. Acad. Sci.* **92,** 3948–3952
9. Wolosker, H., Blackshaw, S., and Snyder, S. H. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 13409–13414
10. Nagata, Y., Homma, H., Matsumoto, M., and Imai, K. (1999) *FEBS* **454,** 317–320
11. Cardinale, G. J., and Abeles, R. H. (1968) *Biochemistry* **7,** 3970–3978
12. Rudnick, G., and Abeles, R. H. (1975) *Biochemistry* **14,** 4515–4522
13. Reina-San-Martin, B., Degrave, W., Rougeot, C., Cosson, A., Chamond, N., Cordeiro-da-Silva, A., Arala-Chaves, M., and Minoprio, P. (2000) *Nat. Med.* **6,** 890–897
14. Cano, M. I., Gruber, A., Vazquez, M., Cortés, A., Levin, M. J., González, A., Degrave, W., Rondinelli, E., Zingales, B., Ramirez, J. L., Alonso, C., Requena, J. M., and Silveira, J. F. D. (1995) *Mol. Biochem. Parasitol.* **71,** 273–278
15. Higuchi, R., Krummel, B., and Saiki, K. K. (1988) *Nucleic Acids Res.* **16,** 7351–7367
16. Keenan, M. V., and Alworth, W. L. (1974) *Biochem. Biophys. Res. Commun.* **57,** 500–504
17. Fisher, L. M., Albery, W. J., and Knowles, J. R. (1986) *Biochemistry* **25,** 2529–2537
18. Albery, W. J., and Knowled, J. R. (1986) *Biochemistry* **25,** 2572–2577
19. Breitbart, R. E., Andreadis, A., and Nadal-Ginard, B. (1987) *Annu. Rev. Biochem.* **56,** 467–495
20. Manning-Cela, R., Gonzalez, A., and Swindle, J. (2002) *Infect. Immun.* **70,** 4726–4728
21. Krassner, S. M., and Flory, B. (1972) *J. Protozool.* **19,** 917–920
22. Bowman, I. B. R., Srivastava, H. K., and Flynn, I. W. (1972) in *Comparative Biochemistry of Parasites* (Van den Bossche, H., ed) Academic Press, New York
23. Evans, D. A., and Brown, R. C. (1972) *J. Protozool.* **19,** 686–690
24. Auerswald, L., Schneider, P., and Gade, G. (1998) *J. Exp. Biol.* **201,** 2333–2342
25. Sylvester, D., and Krassner, S. M. (1976) *Comp. Biochem. Physiol.* **55,** 443–447
26. de Isola, E. L., Lammel, E. M., Katzin, V. J., and Gonzalez Cappa, S. M. (1981) *J. Parasitol.* **67,** 53–58
27. Contreras, V. T., Salles, J. M., Thomas, N., Morel, C. M., and Goldenberg, S. (1985) *Mol. Biochem. Parasitol.* **16,** 315–327
28. Silber, A. M., Tonelli, R. R., Martinelli, M., Colli, W., and Alves, M. J. (2002) *J. Eukaryot. Microbiol.* **49,** 441–446
29. Janeway, C. A., and Humphrey, J. H. (1970) *Folia Biol.* (*Prague*) **16,** 156–172
30. Mozes, E., Kohn, L. D., Hakim, F., and Singer, D. S. (1993) *Science* **261,** 91–92
31. Sela, M., and Zisman, E. (1997) *FASEB J.* **11,** 449–456
32. Contreras, V. T., Morel, C. M., and Goldenberg, S. (1985) *Mol. Biochem. Parasitol.* **14,** 83–96
33. Souto-Padron, T., Reyes, M. B., Leguizamon, S., Campetella, O. E., Frash, A. C., and de Souza, W. (1989) *Eur. J. Cell Biol.* **50,** 272–278
34. Janes, B. K., and Bender, R. A. (1998) *J. Bacteriol.* **180,** 563–570
35. de Jong, M. H., van der Drift, C., and Vogels, G. D. (1975) *J. Bacteriol.* **123,** 824–827
36. Shakibaei, M., and Frevert, U. (1996) *J. Exp. Med.* **184,** 1699–1711
37. Burleigh, B. A., and Andrews, N. W. (1998) *Curr. Opin. Microbiol.* **1,** 461–465
38. Gao, W., Wortis, H. H., and Pereira, M. A. (2002) *Int. Immunol.* **14,** 299–308
39. Martin, D., Ault, B., and Nadler, J. V. (1992) *Eur. J. Pharmacol.* **219,** 59–66
40. Van Harreveld, A. (1980) *J. Neurobiol.* **11,** 519–529
41. Thompson, R. J., Bouwer, H. G., Portnoy, D. A., and Frankel, F. R. (1998) *Infect. Immunol.* **66,** 3552–3561
42. Wolosker, H., Sheth, K. N., Takahashi, M., Mothet, J. P., Brady, R. O., Jr., Ferris, C. D., and Snyder, S. H. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 721–725
43. Watanabe, T., Shibata, K., Kera, Y., and Yamada, R. (1998) *Amino Acids* **14,** 353–360