Eur. J. Biochem. **271**, 3050–3063 (2004) © FEBS 2004 doi:10.1111/j.1432-1033.2004.04237.x

Identification of three proteins that associate *in vitro* with the *Leishmania* (*Leishmania*) *amazonensis* G-rich telomeric strand

Maribel F. Fernández[1], Rafael R. Castellari[2], Fábio F. Conte[1], Fábio C. Gozzo[3], Adão A. Sabino[3], Hildete Pinheiro[4], José C. Novello[2], Marcos N. Eberlin[3] and Maria I. N. Cano[1]

[1]*Departamento de Patologia Clínica, Faculdade de Ciências Médicas;* [2]*Laboratório de Química de Proteínas (LAQUIP), Departamento de Bioquímica, Instituto de Biologia;* [3]*Instituto de Química;* [4]*Departamento de Estatística, Instituto de Matemática, Estatística e Computação Científica, Universidade Estadual de Campinas (UNICAMP), Brazil*

The chromosomal ends of *Leishmania (Leishmania) amazonensis* contain conserved 5′-TTAGGG-3′ telomeric repeats. Protein complexes that associate *in vitro* with these DNA sequences, *Leishmania amazonensis* G-strand telomeric protein (LaGT1-3), were identified and characterized by electrophoretic mobility shift assays and UV cross-linking using protein fractions purified from S100 and nuclear extracts. The three complexes did not form (a) with double-stranded DNA and the C-rich telomeric strand, (b) in competition assays using specific telomeric DNA oligonucleotides, or (c) after pretreatment with proteinase K. LaGT1 was the most specific and did not bind a *Tetrahymena* telomeric sequence. All three LaGTs associated with an RNA sequence cognate to the telomeric G-rich strand and a complex similar to LaGT1 is formed with a double-stranded DNA bearing a 3′ G-overhang tail. The protein components of LaGT2 and LaGT3 were purified by affinity chromatography and identified, after renaturation, as $\approx$35 and $\approx$52 kDa bands, respectively. The $\leq$15 kDa protein component of LaGT1 was gel-purified as a UV cross-linked complex of $\approx$18–20 kDa. Peptides generated from trypsin digestion of the affinity and gel-purified protein bands were analysed by matrix-assisted laser desorption/ionization-time of flight and electrospray ionization tandem mass spectrometry. The fingerprint and amino acid sequence analysis showed that the protein components of LaGT2 and of LaGT3 were, respectively, similar to the kinetoplastid Rbp38p and to the putative subunit 1 of replication protein A of *Leishmania* spp., whereas the $\leq$15 kDa protein component of LaGT1 was probably a novel *Leishmania* protein.

Keywords: affinity purification; EMSA; *Leishmania amazonensis*; mass spectrometry; telomeric proteins.

In almost all eukaryotes, including the pathogenic protozoan *Leishmania (Leishmania) amazonensis*, telomeres are nucleoprotein complexes formed by tandem repeats of conserved DNA sequences associated with proteins [1,2]. One of the telomere strands is G-rich and runs 5′ $\rightarrow$ 3′ towards the end of the chromosomes, where it forms a single-stranded protrusion or 3′ G-overhang [3]. The G-rich strand is the substrate for telomerase and for other telomere binding proteins involved in telomere length regulation and maintenance [4,5]. The length of this G-rich telomere extension appears to be cell cycle regulated in humans and yeast [6–8] and its loss leads to genome instability and chromosomal end fusion through the activation of DNA damage checkpoints [5,9].

Proteins associated with both double-stranded and G-rich single-stranded telomeric DNA and with accessory proteins have been described in many eukaryotes. These proteins form a high order nucleoprotein complex that functions mainly to maintain the genome stability by regulating telomerase activity, the expression of genes positioned at telomeres, and the capping of chromosome ends to protect them from degradation and fusions [10,11]. For example, during the S phase, which is the period of increased single-strand extension in yeast telomeres [7], Cdc13p exhibits high affinity for the G-strand. Cdc13p activity is essential for the protection of chromosome ends and also positively and negatively regulates the replication of telomeres [12–14]. The positive regulatory role involves the formation of a complex with the telomerase-associated protein Est1, resulting in the recruitment of telomerase to telomeres [15,16]. In addition, the interaction of Cdc13p with Stn1p and/or with Ten1p, might negatively regulate telomerase recruitment [17,18]. Cdc13p is also associated with DNA pol α [19], although the relevance of this association has only very recently been clarified. Chandra *et al.* [14] identified mutations of

Correspondence: M. I. N. Cano, Departamento de Patologia Clínica, Faculdade de Ciências Médicas, Universidade Estadual de Campinas (UNICAMP), CP 6109, Campinas, São Paulo, 13083-970, Brazil. Fax:/Tel.: + 55 19 37887370, E-mail: micano@unicamp.br

Abbreviations: LaGT, *Leishmania amazonensis* G-strand telomeric protein; Cdc13, cell division control protein 13; EMSA, electrophoretic mobility shift assays; Est1, ever short telomere 1; NP-40, Nonidet P-40; OB, oligonucleotide/oligosaccharide-binding; *On*Tebp, *Oxythricha nova* telomere binding protein; Pot1, protection of telomere 1; Rpa1, replication protein A subunit 1; Rbp38, RNA binding protein 38; Trf1 and Trf2, telomere repeat factor 1 and 2.

(Received 28 December 2003, revised 23 April 2004, accepted 1 June 2004)

CDC13 which led them to propose that the activities of Cdc13p actually correspond to distinct steps during telomere replication: one that coordinates and the other that regulates the synthesis of both telomere strands. In humans, hPot1 protein binds specifically to the G-rich telomere strand [20] and act as a telomerase-dependent positive regulator of telomere length [21]. Furthermore, it was shown recently that hPot1p interacts with the double-stranded telomeric protein Trf1 and this interaction increases the loading of hPot1p on the single-stranded telomeric DNA, which can provide a role for hPot1p in regulating telomere length [22].

Apart from telomerase activity [23] and the results described below, there are no descriptions of proteins that may interact specifically with *Leishmania* telomeres. Among the Kinetoplastida, a few reports have dealt with the telomeric chromatin of *Trypanosoma brucei* [24,25]. Eid and Sollner-Webb [26,27] described St1p and St2p, which are protein·DNA complexes with a high affinity for subtelomeric sequences of both procyclic and bloodstream forms of *T. brucei*. Three single-stranded protein·DNA complexes (C1, C2 and C3) specific for the G-rich telomeric repeat have been shown to copurify with telomerase activity in *T. brucei* [28]. Two of these complexes (C2 and C3) also bind to an RNA sequence cognate to the telomeric DNA and to a partial duplex that mimics 3′ G-overhangs. Complex C3 also shares features with single-stranded telomeric G-rich proteins described in other eukaryotes [29], and the predictive sequence of C3-associated proteins shows that they are probably novel specific *T. brucei* single-stranded telomere-binding proteins. Other G-strand binding proteins of *Leptomonas* and *T. brucei* have been described but not characterized [30].

Here, we report the partial characterization and the identification of three *L. amazonensis* proteins that bind *in vitro* to the telomeric G-strand (*L. amazonensis* G-strand telomeric proteins; LaGT1, LaGT2 and LaGT3). Binding activities were found in S100 and nuclear extracts of *L. amazonensis* promastigotes after anion-exchange chromatography. Purification of the protein components of LaGT2 and LaGT3 was achieved using single-stranded 5′-biotinated G-telomeric oligonucleotide affinity columns. Two major proteins of approximately 35 kDa and 52 kDa were eluted from the columns and identified as components of LaGT2 and LaGT3, respectively, after renaturation experiments. LaGT1 protein (≤ 15 kDa) was gel-purified as a Coomassie-stained UV-irradiated complex of ≈ 18–20 kDa that migrated in the same position of the radiolabeled LaGT1 UV-irradiated complex. MALDI-TOF MS fingerprint analysis and ESI-MS/MS sequencing of tryptic digested peptides indicated that the ≈ 52 kDa band with LaGT3 activity was similar to subunit 1 of the conserved single-stranded binding protein, replication protein A (Rpa1) of *Leishmania* spp., whereas the ≈ 35 kDa protein with LaGT2 activity was homologous to the RNA-binding protein characterized previously as Rpb38p in *Leishmania tarentolae* and *T. brucei* [31]. The protein component of LaGT1 (≤ 15 kDa) has no homologues in the protein databases indicating that it is probably a novel *Leishmania* protein. The telomere function of the LaGT protein components in *Leishmania* remains to be determined.

Materials and methods

Parasite cultures

Promastigote forms of *L. amazonensis*, strain MHOM/BR/73/M2269, were cultivated in Schneider's medium (Sigma) supplemented with 5% (v/v) heat-inactivated fetal bovine serum (Cultilab) and 1× antibiotic/antimycotic solution (Life Technologies) at 28 °C for 72 h in 25 cm^3 culture flasks. Parasite cultures were maintained in exponential growth and monitored by counting in a hemocytometer.

L. amazonensis S100 and nuclear extracts

S100 extract was obtained in the presence of protease inhibitors as described by Cano *et al.* [23]. Nuclear extracts were prepared using a modification of the protocol reported by Noll *et al.* [32]. Parasite cells were harvested by centrifugation at 11 400 ***g*** for 15 min at 4 °C and washed in 1× $NaCl/P_i$ supplemented with 2% (v/v) glucose. The pellets were resuspended in buffer A (20 mM Tris/HCl, pH 7.5, 1 mM EGTA, pH 8.0, 1 mM EDTA, pH 8.0, 1 mM spermidine, 0.3 M spermine, 5 mM 2-mercaptoethanol), supplemented with a cocktail of protease inhibitor (Set III, Calbiochem) and 0.5% (v/v) Nonidet P-40 (NP-40) at 4 °C. The lysis was checked by reverse phase optical microscopy and fluorescence microscopy after DAPI staining. Nuclei were separated from the cytoplasmic fraction by centrifugation at 11 000 ***g*** for 1 h at 4 °C. The pellet containing intact nuclei was washed twice in 1× TMG [10 mM Tris/HCl, pH 8.0, 1.2 mM $MgCl_2$, 10% glycerol (v/v)] at 17 700 ***g*** for 30 min at 4 °C and resuspended in 1× TMG supplemented with the protease inhibitor cocktail, 1 mM dithiothreitol and 1 mM EGTA, pH 8.0. The lysis was achieved by blending in a mixer in the presence of liquid nitrogen. The protein extract was separated from nuclear debris by centrifugation at 39 800 ***g*** for 20 min at 4 °C followed by ultracentrifugation at 100 000 ***g*** for 90 min at 4 °C. The supernatant (aqueous phase) was aliquoted and frozen in liquid nitrogen. The protein concentrations of the resulting S100 and nuclear extracts were determined by the Bradford method (Bio-Rad). For the binding assays, the extracts were fractionated by anion-exchange DEAE-agarose chromatography (Bio-Gel A, Bio-Rad). The columns were equilibrated with 1× TMG containing 50 mM sodium acetate (NaOAc), pH 8.0, and washed with six volumes of 1× TMG. The proteins were eluted with increasing concentrations of NaOAc, pH 8.0, in 1× TMG. When appropriate, and before testing for binding activity, all fractions were desalted in Microcon-30 filters (Amicon) to a final salt concentration of 50 mM.

Preparation of single-stranded, partial duplex with 3′ G-overhang and double-stranded oligomers

DNA oligonucleotides (Table 1) were purchased from MWG (http://www.mwg-biotech.com) and Operon Technologies (http://www.qiagen.com) and gel purified before and after 5′ end-labeling with [^{32}P]ATP[γP] and T_4 polynucleotide kinase [33]. The partial duplex 3′ G-rich overhang and double-stranded telomeric DNA were obtained by mixing equimolar amounts of radiolabeled

Table 1. Oligonucleotides used in EMSA, UV cross-linking and in affinity chromatography.

Oligonucleotide	Sequence
Tel1	5′-TTAGGGTTAGGGTTAGGG-3′
Tel2	5′-TAGGGTTAGGGTTAGGGT-3′
Tel3	5′-AGGGTTAGGGTTAGGGTT-3′
Tel4	5′-GGGTTAGGGTTAGGGTTA-3′
Tel5	5′-GGTTAGGGTTAGGGTTAG-3′
Tel6	5′-GTTAGGGTTAGGGTTAGG-3′
Tel6-Rev	5′-CCTAACCCTAACCCTAAC-3′
Tel6RNA	5′-GUUAGGGUUAGGGUUAGG-3′
Tet-tel	5′-GTTGGGGTTGGGGTTGG-3′
T3	5′-AATTAACCCTCACTAAAGGG-3′
T7	5′-GTAATACGACTCACTATAGGG-3′
TS	5′- AATCCGTCGAGCAGAGTT-3′
OvhF	5′-CTGGCCGTCGTTTTACTTAGGGTTAGGGTTAGG-3′
OvhR	5′- GTAAAACGACGGCCAG-3′
CSB1	5′-GTACAGTGTACAGTGTACAGT-3′
5′ biotinTel6	5′ biotin-GTAATACGACTCGTTAGGGTTAGGGTTAGG-3′

sense and antisense oligonucleotides, as described by Cano *et al.* [28]. Fully partial duplex and double-stranded DNA (dsDNA) were purified from the residual single-stranded DNA (ssDNA) and quantified [28].

Electrophoretic mobility shift assay

All the conditions used for the binding reactions and the EMSA, including temperature of binding and the concentration of protein fractions and oligoprobes were standardized prior to proceeding with the experiments. Due to the scarcity of telomeric proteins in semipurified S100 and nuclear extracts, the complexes (LaGT1, LaGT2 and LaGT3) were formed when a minimum of 1 μg of protein fractions and 9–25 fmol of labeled telomeric DNA oligoprobe were used in the binding reactions. In most of the assays shown here we used protein fractions (1 μg each) from the S100 and nuclear extracts that were semipurifeid in DEAE-agarose columns. They were incubated individually with 9 fmol of purified 5′ [^{32}P]ATP[γP] end-labeled oligonucleotide in a 20 μL reaction containing 25 mM Hepes, pH 7.5, 5 mM $MgCl_2$, 0.1 mM EDTA, pH 8.0, 100 mM KCl, 10% (v/v) glycerol, 0.1% (v/v) NP-40, 0.5 mM dithiothreitol and 100 ng of poly(dI-dC)·poly(dI-dC) (Amersham Biosciences). Samples were incubated on ice for 30 min before loading onto a 6% native PAGE gel [37.5 : 1, acrylamide/bis-acrylamide (w/w)] in 0.5× TBE (44.5 mM Tris base, 44.5 mM boric acid, 1 mM EDTA, pH 8.0) at 4 °C followed by electrophoresis at 150 V for ≈ 3 h. For autoradiography, wet gels were exposed for 2 h to a Kodak X-Omat film at −80 °C.

Competition assays

For the binding assays, nonradiolabeled oligonucleotide competitors were added in excess relative to the amount of 5′ [^{32}P]ATP[γP] end-labeled Tel6 oligoprobe (Table 1). The concentrations of competitors in these reactions were 0.45, 0.9, 2.25, 4.5, 9, 18 and 36 pmol. As the order of addition of the competitors relative to the probe did not affect the binding activity of the complexes tested (data not shown), the competition assays were done by adding the probe and competitor at the same time.

The shift in the protein·DNA complexes in the absence or presence of a molar excess of unlabeled competitors in two independent EMSA, was assessed quantitatively using SCION IMAGE processing and analysis software (http://www.scioncorp.com) as described in Cano *et al.* [28]. The results plotted in the graphs represent the percentage of the binding activity of a shifted complex (the ratio of the density area in arbitrary scanning units, and the sum of the density areas of all shifted complexes, including unbound oligonucleotide, in each lane, multiplied by 100). The statistical analysis of three independent results was performed using SAS software as described below.

Statistical analysis

The software used for the statistical analysis was SAS (SAS Institute Inc., The SAS System for Windows, Release 8.02 TS Level 02M0, 2001; SAS Institute Inc., Cary, NC, USA). All the analysis used the Mantel–Haenszel test statistic to test the null hypothesis of equal distribution of the density areas of each complex in the absence or presence of salts and or unlabeled competitors. The null hypothesis was rejected for $P < 0.05$, compared to the control.

Proteinase K digestion

To ensure the complexes were formed by the association of proteins and nucleic acids, 1 μg of each protein fraction was treated with 10 μg of proteinase K (Amersham Biosciences) for 15 min at 56 °C before the binding assays.

Effect of salt concentration

Binding assays using the DEAE fractions of S100 and nuclear extracts were done in the presence of a standard concentration of KCl (100 mM) used in normal reactions and varying concentrations of $MgCl_2$ (0–50 mM), or of a standard concentration of $MgCl_2$ (5 mM) used in normal reactions and varying concentrations of KCl (0–800 mM).

UV cross-linking assays

UV cross-linking in solution was performed on ice by exposing the 20 μL binding reaction mixture in siliconized Eppendorf tubes covered with plastic film to 254 nm UV light (Ultra-lum, Inc., Claremont, CA, USA) for 15 min as described previously [28]. After irradiation, the samples were mixed with 5× SDS loading buffer to a final concentration of 1×, boiled for 5 min and loaded onto a 12% polyacrylamide gel [29 : 1, acrylamide/bis-acrylamide (w/w)]. Electrophoresis was carried out in 1× protein running buffer [33] at room temperature. The gel was fixed in 10% methanol/5% glacial acetic acid (v/v) for 30 min at room temperature and exposed for 1–18 h to a Kodak X-Omat film at −80 °C.

UV cross-linking *in situ* was also carried out by exposing a wet 6% mobility shift gel on ice to 254 nm UV light for

30 min; the gel was no more than 5–7 cm from the source. The gel was then exposed to film and the bands corresponding to each complex were excised, eluted overnight at 4 °C in 1× SDS loading buffer, denatured for 5 min and loaded onto a 12% gel. The gel was fixed and exposed for 1–18 h to a Kodak X-Omat film at −80 °C. In both cases, molecular mass markers (Rainbow, Amersham Biosciences) were included to identify the positions of the cross-linked proteins.

SDS/PAGE and Coomassie blue staining

Protein fractionation was done in 12% and 15% gels [29 : 1, acrylamide/bis-acrylamide (v/v)] and electrophoresis was carried out in 1× protein running buffer at room temperature. The protein bands were visualized by Coomassie blue staining, according to a standard protocol [33].

Purification of LaGT2 and LaGT3 activities by G-DNA affinity chromatography

The purification step using anion-exchange chromatography was done at 4 °C [28]. DEAE-agarose fractions (2.98 mg of protein corresponding to $\approx 2.8 \times 10^9$ cells) from S100 extracts containing the activities of all three LaGTs were affinity purified on separate G-DNA columns (0.5 mL each) prepared with modifications of the protocol described by Schnapp *et al.* [34]. For preparation of the column, 1 mL of 50% (v/v) Ultralink Immobilized Neutravidin™ Plus (Pierce) was pre-equilibrated in buffer E (100 mM KCl, 0.1% (v/v) NP-40, 25 mM Hepes, pH 7.5, 5 mM $MgCl_2$, 0.1 mM EDTA, pH 8.0, 10% (v/v) glycerol, 0.5 mM dithiothreitol) for 15 min at 4 °C. Pools of three DEAE fractions (≈ 2.98 mg of protein) enriched for LaGT2 and LaGT3 activities (75 mM, 100 mM and 200 mM) were then mixed with 4 nmol of 5′-biotin-Tel6 oligonucleotide (Table 1) in the presence of buffer E and 10 μg of poly(dI-dC)·poly(dI-dC) for 30 min at 4 °C. These oligonucleotide/extract mixtures were added to 500 μL of pre-equilibrated Neutravidin™ beads and incubated overnight at 4 °C. The mixtures were then poured and packed into a 2 mL disposable column (Bio-Rad) and the unbound proteins were collected in the column flow-through. The column was washed with 10 column volumes of buffer E and the bound proteins were eluted with a stepwise KCl gradient (0.6–2.2 M) in buffer E. Five 1.0 mL fractions were collected, concentrated, desalted in Microcon-30 filters, and tested for LaGT activities in UV cross-linking assays. As a control, mock columns were prepared in the absence of 5′-biotinylated oligonucleotide.

Purification of LaGT1 activity

The protocol used to purify the LaGT1 protein component was a modification of the method described for the purification of *T. brucei* telomeric complex C3 [28]. DEAE fractions from S100 extract enriched for LaGT1 protein were pooled (10 mg) and mixed with 5.0 nmol of unlabeled Tel6 in a preparative binding reaction as described above. As a control, a 20 μL binding reaction was carried out with the pool of the DEAE fractions and a 5′ end-labeled Tel6 oligonucleotide (see above). Both binding reactions were fractionated in the same 6% native polyacrylamide gel, and after running, the complexes were UV cross-linked *in situ* (see above). The gel was then exposed to film to reveal the position of the labeled LaGT1 complex. The labeled and unlabeled complexes were excised from the gel based on the position of the labeled complex and eluted overnight with gentle agitation at 4 °C in 1× protein-loading buffer. The protein-forming LaGT1 complexes were separated by SDS/PAGE in a 15% gel, Coomassie-stained and exposed to Kodak X-Omat film. The unlabeled protein band was further digested with trypsin and submitted to MS analysis (see below).

Peptide mapping and sequencing by mass spectrometry (MALDI-TOF MS and ESI-MS/MS)

Coomassie-stained protein bands containing LaGT2 and LaGT3 activities and the irradiated protein·DNA LaGT1 complex were excised from the gel, in-gel digested with trypsin (sequencing grade porcine trypsin, Promega), according to the University of California, San Francisco (UCSF) Mass Spectrometry Facility in-gel digestion procedure (http://donatello.ucsf.edu/ingel.html), and subjectd to MALDI-TOF MS, using a Voyager-DE PRO mass spectrometer (PerSeptive Biosystems) and a MALDI LR instrument (Micromass). To determine the molecular masses of the predicted peptides, the MALDI-TOF MS fingerprints were compared with the protein sequence databases (NCBInr and Genpept) using the Protein Prospector MS-FIT 4.0 analysis program (P. R. Baker & K. R. Clauser; http://prospector.ucsf.edu/ucsfhtml4.0/msfit.htm) set at a mass tolerance (accuracy) of 50 p.p.m. and calibrated with protein standards (Sequazyme Peptide Mass Standards Kit, Calibration Mixture 1 and 2; Applied Biosystems). The searches were also performed manually using the *Leishmania* protein sequence database (*Leishmania* GeneDB, http://www.ebi.ac.uk/parasites/leish.html).

ESI-MS/MS analysis were performed in a Q-Tof (Micromass) coupled to a CapLC (Waters) chromatographic system. The tryptic peptides were purified using a Waters Opti-Pak C18 trap column. The trapped peptides were eluted using a water/acetonitrile 0.1% (v/v) formic acid gradient and separated by a 75 μm i.d. capillary column home-packed with C18 silica. Data was acquired in data dependent mode, and multiply charged ions were subjected to MS/MS experiments. The MS/MS spectra were processed using MAXENT3 (Micromass) and manually sequenced using the PEPSEQ program (Micromass).

Results

Three protein·DNA complexes interact *in vitro* with the G-rich telomeric strand of promastigotes of *L. amazonensis*

In addition to telomerase activity (data not shown), three protein·DNA complexes that interact *in vitro* with the G-rich telomeric strand were identified in DEAE-agarose fractions of S100 and nuclear extracts from *L. amazonensis* promastigotes. Due to the limiting amount of telomeric proteins present in these extracts, binding reactions were done with a minimum of 1 μg of the semipurified fractions

of S100 and nuclear extracts and varying concentrations of the telomeric DNA oligoprobe (data not shown). Three complexes named LaGT1, LaGT2 and LaGT3, according to their electrophoretic mobility in a 6% nondenaturing gel, were formed with different protein fractions of the S100 and nuclear extracts and the 5′ end-labeled Tel6 oligonucleotide, and detected at 4 °C by EMSA. Using DEAE-agarose fractions from the S100 extract, complex LaGT1, the fastest migrating complex, was formed with fractions that eluted with 75–800 mM sodium acetate (Fig. 1A, lanes 3–10), although it was more abundant in fractions eluted with 100–400 mM sodium acetate (Fig. 1A, lanes 4–7). Complex LaGT2 was formed mainly with fractions eluted with 75–100 mM sodium acetate (Fig. 1A, lanes 3 and 4) and complex LaGT3 (the slowest migrating complex) was formed only with the fraction eluted with 75 mM sodium acetate (Fig. 1A, lane 3). All three complexes were formed when nonpurified S100 extract was used as the protein source in the binding reactions (Fig. 1A, lane 2) and no LaGT1 was formed when the reaction was incubated at temperatures above 4 °C (data not shown), suggesting that *in vitro* it is labile or unstable.

The same three complexes were formed when the DEAE-agarose fractions from nuclear extracts were used for the binding reactions with Tel6 as the oligoprobe. However, there were differences in the concentration and the elution

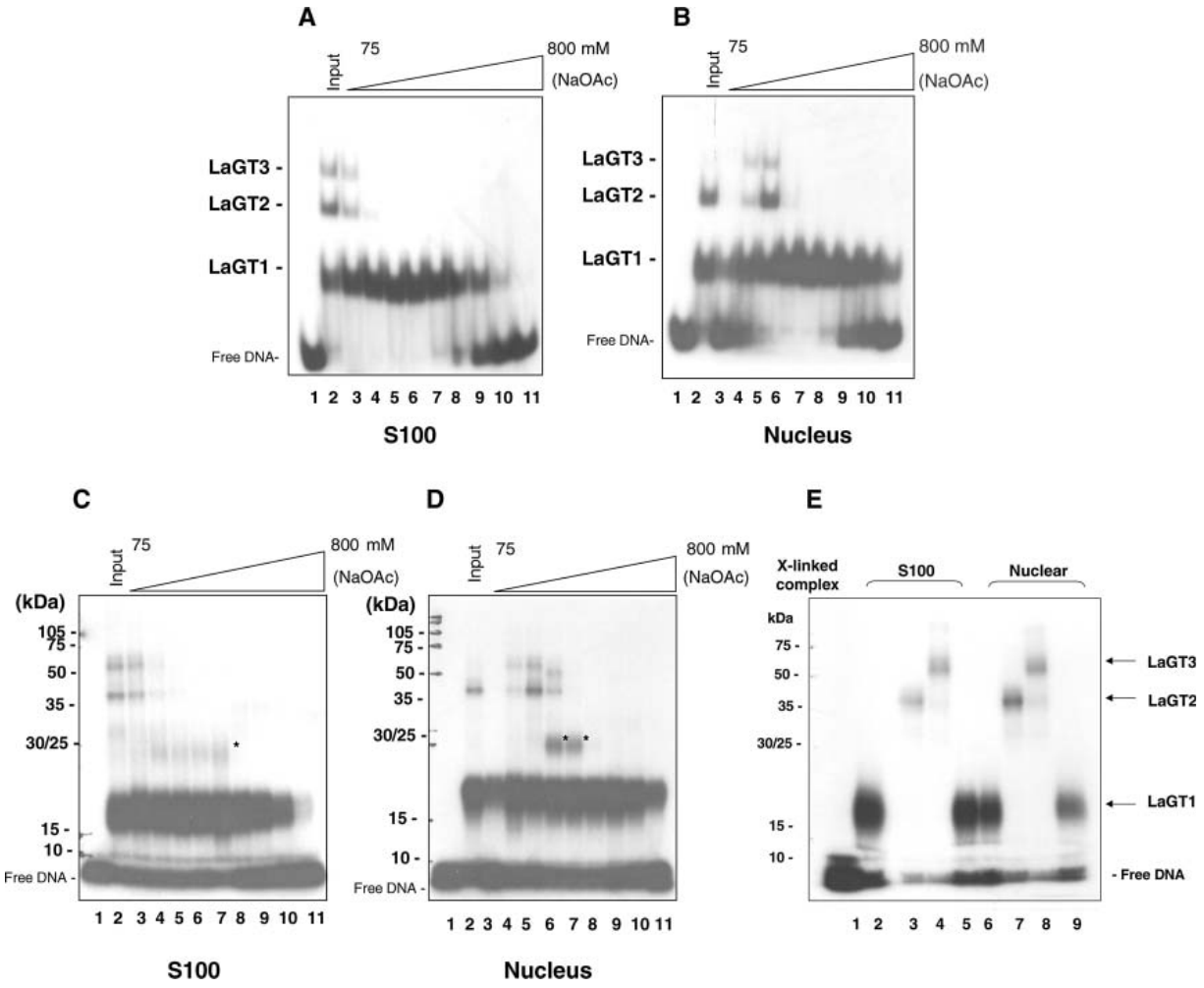

Fig. 1. The protein·DNA complexes that associate *in vitro* with the G-rich telomeric strand of *L. amazonensis*. Assays were carried out with the 5′ end-labeled Tel6 as probe and crude extracts (input) or DEAE fractions eluted with sodium acetate (NaOAc). (A) EMSA of S100 extract (lane 2) and DEAE fractions (lanes 3–11). The shifted bands (LaGT1, LaGT2 and LaGT3) were classified according to their order of migration in the gel. (B) EMSA of nuclear extract (lane 2) and DEAE fractions (lanes 3–11). The shifted complexes were classified as in (A). In lanes 1 of (A-E), the reactions were carried out without extract. (C, D) Binding reactions in (A, B), respectively, were exposed to UV light and the cross-linked proteins then separated by SDS/PAGE in 12% gels (lanes 2–11 in both panels). The arrows (E) indicate the position of the cross-linked complexes. An extra ≈ 24 kDa complex, indicated with an asterisk, appeared only after exposing the binding reactions with S100 DEAE fractions 100–400 mM (lanes 4–7) and nuclear DEAE fractions 300–400 mM (lanes 6 and 7) to UV light. kDa, molecular mass in kilodaltons. (E) UV cross-linking *in situ* with proteins from S100 and nuclear extracts. The bands corresponding to LaGT UV-irradiated complexes that were eluted from the gel matrix and separated in 12% protein gels. For this assay the 75 mM (lanes 2, 3 and 4) and 400 mM (lane 5) DEAE fractions from S100 and the 200 mM (lanes 6, 7 and 8) and 500 mM (lane 9) DEAE fractions from the nuclear extract were used. The positions indicated on the right refer to the UV cross-linked complexes LaGT1, LaGT2 and LaGT3.

profile of some of the protein-forming complexes (Fig. 1B). LaGT1 was the most abundant and formed with all DEAE-agarose fractions eluted with 75–800 mM sodium acetate (Fig. 1B, lanes 3–11), but appeared in high concentration at 300–500 mM sodium acetate fractions (Fig. 1B, lanes 6–8). LaGT2 was formed with fractions eluting at 100–300 mM sodium acetate (Fig. 1B, lanes 4–6), and particularly with fraction 200 mM sodium acetate (Fig. 1B, lane 5). In contrast, LaGT3 was formed only with 100 mM and 200 mM sodium acetate fractions (Fig. 1B, lanes 4 and 5) and it was not visible when the reactions were carried out with nonpurified nuclear extract, probably because of its low concentration (Fig. 1B, lane 2). None of the complexes were formed when S100 and nuclear extracts were pretreated with 10 μg of proteinase K, indicating that they are indeed formed by the interaction of proteins and DNA (data not shown).

UV cross-linking assays were carried out to estimate the size of the proteins responsible for the LaGT1, LaGT2 and LaGT3 activities in both extracts (Fig. 1C,D). For these experiments, the same DEAE-agarose fractions that were used in the binding assays and Tel6 5′ end-labeled assays were used. The irradiated samples were denatured at 95 °C in 1× SDS loading buffer and separated by SDS/PAGE in 12% gels. The gel in Fig. 1C shows four prominent bands of ≈ 18–20 to ≥ 60 kDa formed with the S100 (Fig. 1C, lanes 2–4) and nuclear (Fig. 1D, lanes 2 and 4–6) fractions. All molecular masses included the 18 mer (≈ 5.6 kDa) Tel6 oligonucleotide. The differences in the profile and size of the protein bands between the nonpurified extracts (Fig. 1C,D, lane 2) and the DEAE-agarose fractions may reflect the absence or presence of a specific protein which is able to bind to and cross-link with the Tel6 oligonucleotide.

UV-exposed samples containing the protein extracts and the oligoprobe showed 40–60 kDa bands formed with the fractions 75–100 mM (Fig. 1C, lanes 3 and 4) from S100 extract and fractions 100–300 mM (Fig. 1D, lanes 4–6) from nuclear extract, and ≈ 18–20 kDa band formed with the fractions 75–800 mM (Fig. 1C,D, lanes 3–11) from S100 and nuclear extracts, respectively. Proteins of approximately 24 kDa appeared cross-linked to Tel6 only after exposing the binding reactions to UV light (Fig. 1C, lanes 4–7 and Fig. 1D, lanes 6 and 7). Although this experiment alone was unable to accurately determine which protein bands were part of each individual complex, clearly bands of similar molecular mass were formed with purified and nonpurified S100 or nuclear extracts. UV cross-linking *in situ* was therefore carried out with DEAE fractions of S100 (75 mM and 400 mM) and nuclear (200 mM and 500 mM) extracts containing LaGT1, LaGT2 and LaGT3 activities. The complexes formed with the above fractions were cross-linked in the gel and the bands were then excised and eluted from the gel matrix. The eluted protein-forming complexes were fractionated by SDS/PAGE in 12% gels (Fig. 1E) and exposed to film for further identification. The bands corresponding to LaGT1 from the 75 mM and 400 mM fractions of S100 and 200 mM and 500 mM fractions from nuclear extracts were probably formed by ≈ 18–20 kDa complexed proteins as shown in Fig. 1E (lanes 2, 5, 6 and 9). The proteins that formed complexes LaGT2 and LaGT3 in the S100 (75 mM eluate) and nuclear (200 mM eluate) extracts migrated with molecular masses of approximately 40 kDa (Fig. 1E, lanes 3 and 7) and ≥ 60 kDa (Fig. 1E, lanes 4 and 8). In this experiment, the 'extra' ≈ 24 kDa protein band (Fig. 1C, lanes 4–7 and Fig. 1D, lanes 6 and 7) did not appear, probably because it was not part of any of the three LaGT complexes. The values estimated for the protein masses included the mass of the Tel6 oligonucleotide (≈ 5.6 kDa).

The binding specificity of LaGT protein-forming complexes was further tested with different oligoprobes. EMSA was performed using 5′ end-labeled Tel1–Tel6 (3′ end permutations of the telomeric sequence; Table 1), Tel6-RNA, a *Tetrahymena* telomeric sequence (Tet-tel) and Tel6-Rev (C-strand telomeric sequence) as single-stranded oligoprobes, together with a partial duplex DNA containing a 3′ G-overhang and a double-stranded telomeric DNA (Materials and methods), using the 75 mM (enriched for LaGT2 and LaGT3 activities) and 400 mM (enriched for LaGT1 activity) DEAE fractions from the S100 extract. All three LaGT complexes were formed with Tel1–Tel6 and Tel6-RNA oligonucleotides. A corresponding complex, LaGT1, was also formed when the 3′ G-overhang DNA construct was used as the oligoprobe. Although complexes similar to LaGT2 and LaGT3 were formed with the Tet-tel oligonucleotide, no complex was formed with double-stranded telomeric DNA and with the C-strand (Tel6-Rev), suggesting that all LaGT protein-forming complexes preferably associate to the G-rich *L. amazonensis* telomeric strand. These results are summarized in Table 2.

Salt stability of LaGT complexes

To further test the stability of all three protein·DNA complexes, binding reactions with the 75 mM and 400 mM fractions from S100 extract and with the 100 mM and 500 mM fractions from nuclear extracts and oligonucleotide Tel6 were carried out separately in the presence of increased salt concentration ($MgCl_2$ and KCl). In Fig. 2, the reactions

Table 2. Binding activity of LaGT protein-forming complexes with different oligoprobes. Signs – or + indicate the absence or presence of complex formation with the indicated oligoprobe, respectively. The sequence of each oligoprobe is given in Table 1. Details about the preparation of the partial duplex (with 3′ G-overhang) and the double-stranded telomeric DNA are found in [29] and in Experimental procedures. EMSA was used to identify the binding activity of complexes formed with the oligoprobes. The protein source used for the LaGT1 binding assays were the 400 mM DEAE fraction of S100 extract. The protein source used for the LaGT2 and LaGT3 binding assays were the 75 mM DEAE fraction of S100 extract.

	Binding activity		
Oligoprobes	LaGT1	LaGT2	LaGT3
Tel1-Tel6	+	+	+
Tel6RNA	+	+	+
Tel6-Rev	–	–	–
Tet-tel	–	+	+
Partial duplex (with 3′ G-overhang)	+	–	–
Double-stranded telomeric DNA	–	–	–

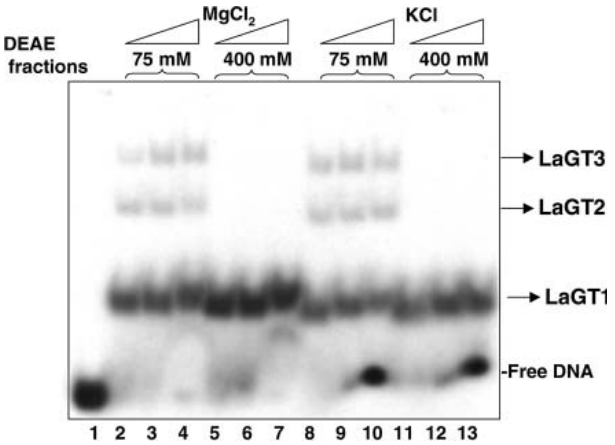

Fig. 2. High concentrations of $MgCl_2$ and KCl do not disturb the formation of the three LaGT complexes. EMSA was carried out with the 75 mM and 400 mM DEAE fractions of S100 extract and the 5′ end-labeled Tel6 oligonucleotide. In lanes 2–7, the reactions were done in the presence of 100 mM KCl and 0 mM (lanes 2 and 5), 5 mM (lanes 3 and 6) and 50 mM (lanes 4 and 7) $MgCl_2$. In subsequent lanes, the reactions were performed in the presence of 5 mM $MgCl_2$ and 0 mM (lanes 8 and 11), 200 mM (lanes 9 and 12) and 800 mM (lanes 10 and 13) KCl. The reaction in lane 1 was carried out in the absence of extract.

were performed with 100 mM KCl (standard concentration used in normal reactions) and varying concentrations of $MgCl_2$ (0–50 mM) although other concentrations were also tested (eluate 75 mM in lanes 2–4, eluate 400 mM in lanes 5–7 and data not shown). The results suggest that regardless of the extract used, high concentrations of $MgCl_2$ did not disturb the binding activity of LaGT proteins. In contrast, binding assays done with 5 mM $MgCl_2$ (standard concentration used in normal reactions) and increased concentrations of KCl (0, 200 mM and 800 mM and others not shown) showed that depending on the extract used, LaGT1–3 activities were partially inhibited (Fig. 2, lanes 10 and 13). Similar results were obtained with the 100 and 500 mM fractions of nuclear extract (data not shown). These results show that the complexes are only slightly unstable in the presence of high concentrations of KCl. This suggests that, under our experimental conditions, the affinity of the proteins to the telomeric sequence may be in part dependent on electrostatic interactions.

LaGT1 is the most abundant and specific G-rich telomeric complex of *L. amazonensis*

The DNA binding specificity of LaGT1, LaGT2 and LaGT3 was also studied by competition assays using the same DEAE fractions (S100 and nuclear extracts) as above. Competition assays were standardized with unlabeled nonspecific oligonucleotides titrated alongside the same amounts of unlabeled telomeric oligonucleotides (in molar excess in relation to the oligoprobe) in the presence of protein extracts and Tel1–Tel6 as the oligoprobes (data not shown). The binding reactions shown in Fig. 3A were carried out with 1 μg of extract and unlabeled telomeric oligonucleotides as specific competitors and in Fig. 3B,C the reactions were done with 1 μg of extract and unlabeled nontelomeric oligonucleotides (Table 1) as nonspecific competitors. The concentration of competitors used in these assays varied from 0.45 to 18 pmol, whereas the probe (labeled Tel6) was used in a fixed concentration of 9 fmol. Figure 3A shows a competition assay in which the 75 mM and 400 mM fractions from S100 (1 μg) were incubated with labeled Tel6 (9 fmol) and increasing concentrations of unlabeled Tel6 as the specific competitor. In assays with the 75 mM fraction 0.45–18 pmol of competitor was used, and in those with the 400 mM fraction 0.9–36 pmol of competitor was used. In lane 1, the reaction was done in the absence of proteins. In subsequent lanes, the reactions were done with 75 mM fraction as the protein source and in the presence of increasing concentration of competitor. All three complexes were completely inhibited [0% binding activity; Fig. 3C, bottom)] by concentrations of unlabeled Tel6 above 9 pmol. Because the LaGT1 activity in the 400 mM fraction was very high, the competition reactions with unlabeled Tel6 were done with 0.9–36 pmol of competitor (Fig. 3A, lanes 9–15). Quantitative analysis of this experiment (Fig. 3A, bottom) showed that LaGT1 activity was almost totally inhibited (96%) only in the presence of 36 pmol of specific competitor. These reactions were also done with the DEAE fractions of nuclear extract with similar results (data not shown).

In Fig. 3B, curves of titration (0.45–9 pmol) by the nonspecific competitors T3, T7 and TS are shown. Binding reactions were done with the 75 mM DEAE fraction from S100 extract as the protein source for all three LaGT activities and labeled Tel6 as probes. The results demonstrate that LaGT2 and LaGT3 binding activities were diminished by 50–80% in the presence of 0.9 pmol of the nonspecific competitors T3, T7 and TS whereas, high concentration of competitors (2.25–9 pmol) increased LaGT1 formation by ≈ 5–23%.

Figure 3C shows assays done with a fixed concentration (9 pmol) of each of the following nonspecific competitors: T3 (Fig. 3C, lanes 3 and 9), T7 (Fig. 3C, lanes 4 and 10), TS (Fig. 3C, lanes 5 and 11), OvhR (Fig. 3C, lanes 6 and 12) and CSB1 (Fig. 3C, lanes 7 and 13), although other concentrations of the these competitors were also tested (Fig. 3B and data not shown). The graph at the bottom of the figure shows that regardless of the protein source used in the assays, LaGT1 activity was not inhibited by any of these nonspecific competitors. In contrast, and as shown in Fig. 3B, increased LaGT1 activity (5–40%) was detected when the assays were done with the 75 mM fraction and the oligonucleotide competitors T3, T7, TS, OvhR and CSB1 (Fig. 3C, lanes 3–7), whereas LaGT2 was inhibited 100% by oligonucleotide T3, and the presence of T7 diminished LaGT3 activity by 99% (Fig. 3C, lanes 3 and 4). LaGT2 activity was also diminished by 51–99% when the competitors used were T7, TS, OvhR and CSB1 (Fig. 3C, lanes 4–7). LaGT3 activity decreased by 87–98% in the presence of unlabeled competitors T3, TS, OvhR and CSB1 (Fig. 3C, lanes 3 and 5–7, respectively). In this case, and as shown in Fig. 3B, the increase in LaGT1 activity probably occurred in detriment to the other complexes, suggesting that more probe became available for LaGT1 binding or that low levels of quadruplex formation in the probes could have changed the effective concentration of the DNA present,

which could be subtle and variable for different competing sequences.

Assays performed with the 400 mM fraction of S100 (Fig. 3C, lanes 8–13 and data not shown) and with a 500 mM fraction of nuclear extract (data not shown), both enriched in LaGT1 activity, showed that LaGT1 was not inhibited by most of the nontelomeric oligonucleotides and was only slightly inhibited (≈ 6%) by the oligonucleotide TS used to detect telomerase activity in TRAP assays [35]. These results indicate that LaGT1 is highly specific for the G-telomeric strand of *L. amazonensis*.

Purification and mass spectrometric identification of the protein-forming LaGT complexes

All three LaGT activities identified in the DEAE-agarose protein fractions were further purified by affinity chromatography on an analytical scale. The 100 mM and 600 mM sodium acetate DEAE fractions from the S100 extract, enriched in LaGT2/LaGT3 and LaGT1 activities, respectively, were loaded in separate affinity columns using a Tel6 5′-biotinylated oligonucleotide with a spacer at the 5′ position as ligand (Table 1). LaGT2 and LaGT3 activities were eluted from the affinity column at 4 °C with increased KCl concentrations (0.6–2.2 M) (Fig. 4A). Size estimation of the affinity-purified proteins was performed in Coomassie-stained gels (Fig. 4A, lanes 6–10). Lanes 2–5 of this gel show the proteins present in the S100 extract, the proteins recovered in DEAE column flow-through, the loaded DEAE fraction (pool of the DEAE fractions 75–200 mM NaOAc) and the proteins that did not associate with the telomeric oligonucleotide in the affinity column (flow-through). Two major protein bands of approximately

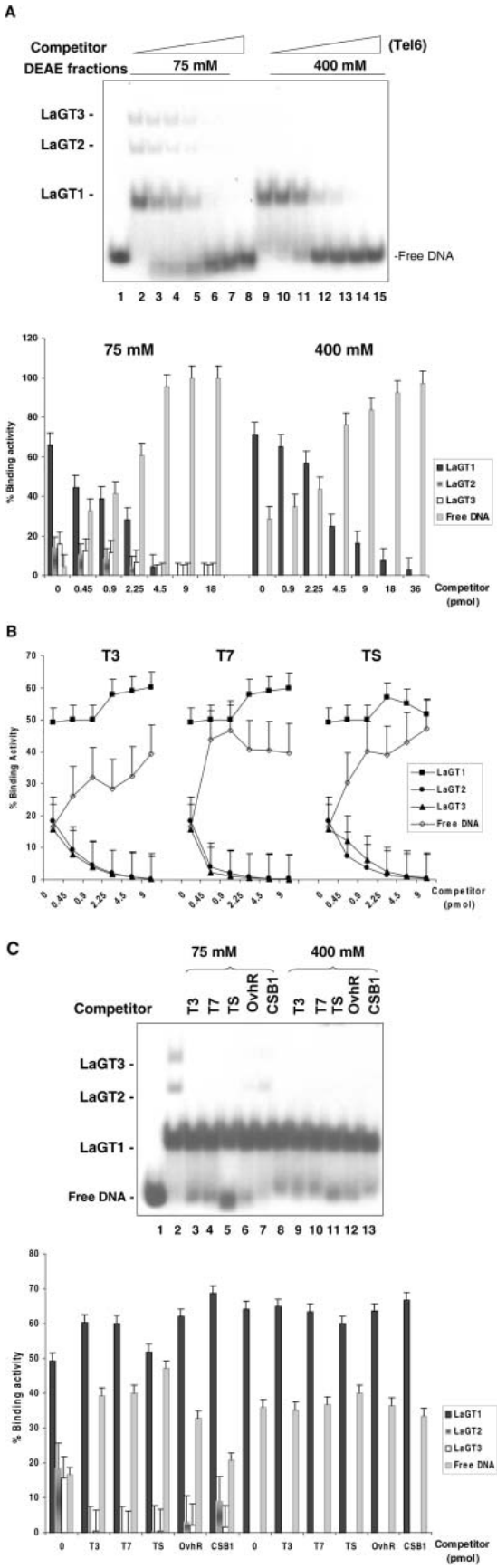

Fig. 3. LaGT1 is highly specific for the *L. amazonensis* G-rich telomeric strand. EMSA using the 75 mM (enriched for LaGT2 and LaGT3 activities) and 400 mM (enriched for LaGT1 activity) fractions from the S100 extract and oligonucleotide Tel6 as probe, under the same conditions as in Figs 1 and 2. (A, top) Unlabeled Tel6 used at concentrations: 0.45 (lane 3), 0.9 (lanes 4 and 10), 2.25 (lanes 5 and 11), 4.5 (lanes 6 and 12), 9 (lanes 7 and 13), 18 (lanes 8 and 14) and 36 pmol (lane 15). The reaction in lane 1, was performed without extract. In lanes 2 and 9 (control reactions), no competitor was added. (A, bottom) The amount of each complex formed in the presence of increased concentrations of unlabeled competitors was expressed as the percentage of binding activity. (B) Titration curves for nontelomeric oligonucleotides T3, T7 and TS in competition assays with labeled Tel6 as probe and the 75 mM DEAE fraction as the protein source. Unlabeled competitors were used at concentrations varying from 0 to 9 pmol. (C, top) Unlabeled nontelomeric oligonucleotides (9 pmol each), T3 (lanes 3 and 9), T7 (lanes 4 and 10), TS (lanes 5 and 11), OvhR (lanes 6 and 12) and CSB1 (lanes 7 and 13) were used as competitors under the same conditions as in (A). Lane 1, reaction performed in the absence of extract; lanes 2 and 8 (control reactions), no competitors were added to the reactions. (C, bottom) The amount of each complex formed in the presence of increased concentrations of unlabeled competitors was expressed as the percentage of binding activity. The graphs show average results of three independent experiments performed in triplicates. Error bars represent the standard error. $P < 0.05$ compared to reactions done in the absence of competitors (control reactions).

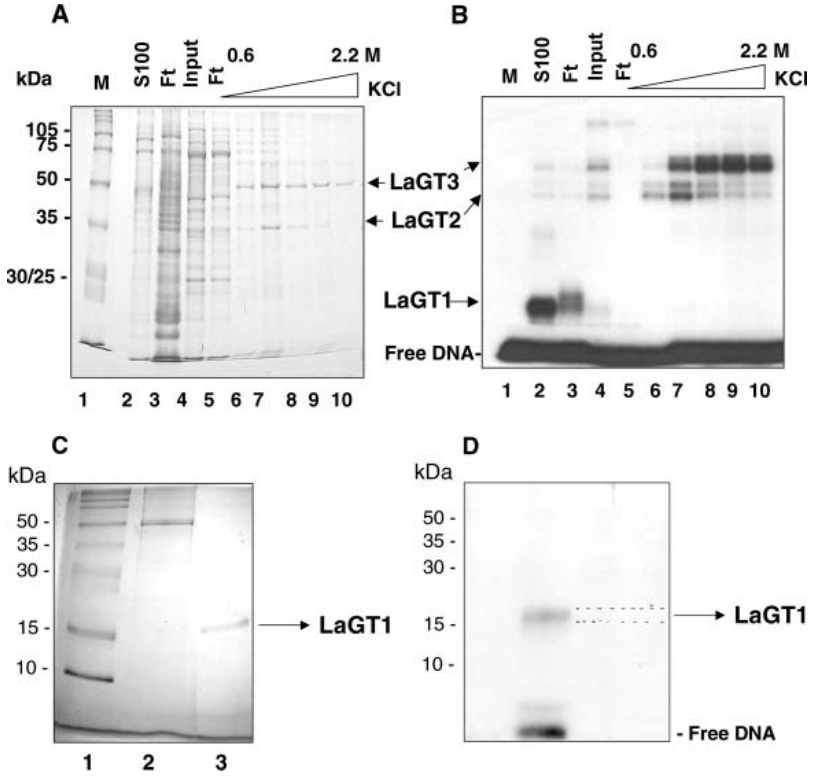

Fig. 4. Purification of LaGT activities. (A) Coomassie-stained SDS/PAGE (12% gel). In lane 1, molecular mass markers; lane 2, total S100 extract; lane 3, flow-through from DEAE-agarose column; lane 4, input (pool of DEAE fractions 75–200 mM); lane 5, flow-through from the affinity column; lanes 6–10, fractions eluted from the affinity column with increasing KCl concentration (0.6–2.2 M). (B) UV cross-linking assay of the protein fractions shown in A. Binding reactions were done with total S100 extract (lane 2), flow-through of DEAE column (lane 3), input (lane 4), flow-through of the affinity column (lane 5), affinity purified fractions (lanes 6–10), and the 5′ end-labeled oligonucleotide Tel6. No extract was added to the assay in lane 1. (C) UV-irradiated LaGT1 complex was gel-purified and fractionated in a Coomassie-stained 15% protein gel. Lane 1, molecular mass marker; lane 2, irradiated LaGT1 complex formed with a labeled Tel6 oligonucleotide; lane 3, a Coomasie-stained ≈ 18–20 kDa band corresponding to the unlabeled LaGT1 irradiated complex. (D) Autoradiogram of the gel in (C).

35 kDa and 52 kDa were eluted in all column fractions with a peak at 1 M KCl (Fig. 4A, lanes 6–10). Protein bands ≥ 65 kDa were also eluted with 0.6 M and 1 M KCl but did not have binding activity (Fig. 4A, lanes 6 and 7). UV cross-linking assays showed that all affinity-purified fractions had LaGT2 and LaGT3 activities with a peak in the 1 M KCl fraction (Fig. 4B, lanes 6–10) that correlated with the protein band patterns shown in Fig. 4A (compare corresponding lanes 6–10). A mock column, to which no biotinylated telomeric oligonucleotide was coupled, was used as a control. In this experiment, all proteins present in the loaded protein fraction were recovered in the column flow-through, indicating that the proteins eluted in the affinity columns associated specifically with the telomeric sequence (data not shown).

Various elution protocols were used to purify LaGT1 activity from affinity columns loaded with the 600 mM DEAE fraction without success (data not shown). LaGT1 remained associated with the telomeric oligonucleotide even at a high salt concentration, a pH gradient (pH 6.0–8.5) and temperatures above 25 °C (data not shown). We only succeed in the purification of LaGT1 after using a modification of the protocol described to purify the protein-forming *T. brucei* telomeric complex C3 [28] (Fig. 4D,E). DEAE fractions enriched for LaGT1 activity were pooled and mixed with an unlabeled Tel6 (preparative reaction) and with a radiolabeled Tel6 oligonucleotide (control reaction), loaded in a preparative 6% native gel and *in situ* UV cross-linked. The irradiated complexes were eluted from the gel matrix and loaded onto a 15% protein gel. A major ≈ 18–20 kDa Coomassie-stained band (Fig. 4D, lane 3) that migrated in the same position as the radiolabeled LaGT1 complex (Fig. 4E, lane 2) was detected.

The affinity purified protein bands of ≈ 35 kDa and ≈ 52 kDa and the UV-irradiated complex of ≈ 18–20 kDa were in-gel digested with trypsin and subjected to MALDI-TOF MS and ESI-MS/MS analysis. The MALDI-TOF spectra obtained for the peptide mixtures produced by tryptic digestion of all proteins are shown in Fig. 5. Comparison of the predicted peptide mass using different databases showed that the ≈ 35 kDa protein shared high similarity with a hypothetical protein of *Leishmania major*, protein L3277.02 or LmRbp38 (Accession no. CAB71224) (the matched peptides cover 52% of the protein), that was identified as a homologue of *L. tarentolae* Rbp38p (Accession no. AAO39844). Rbp38p was recently described by Sbicego *et al.* [31] as an RNA-binding protein that stabilizes mitochondrial RNAs of kinetoplastid protozoa. The gene

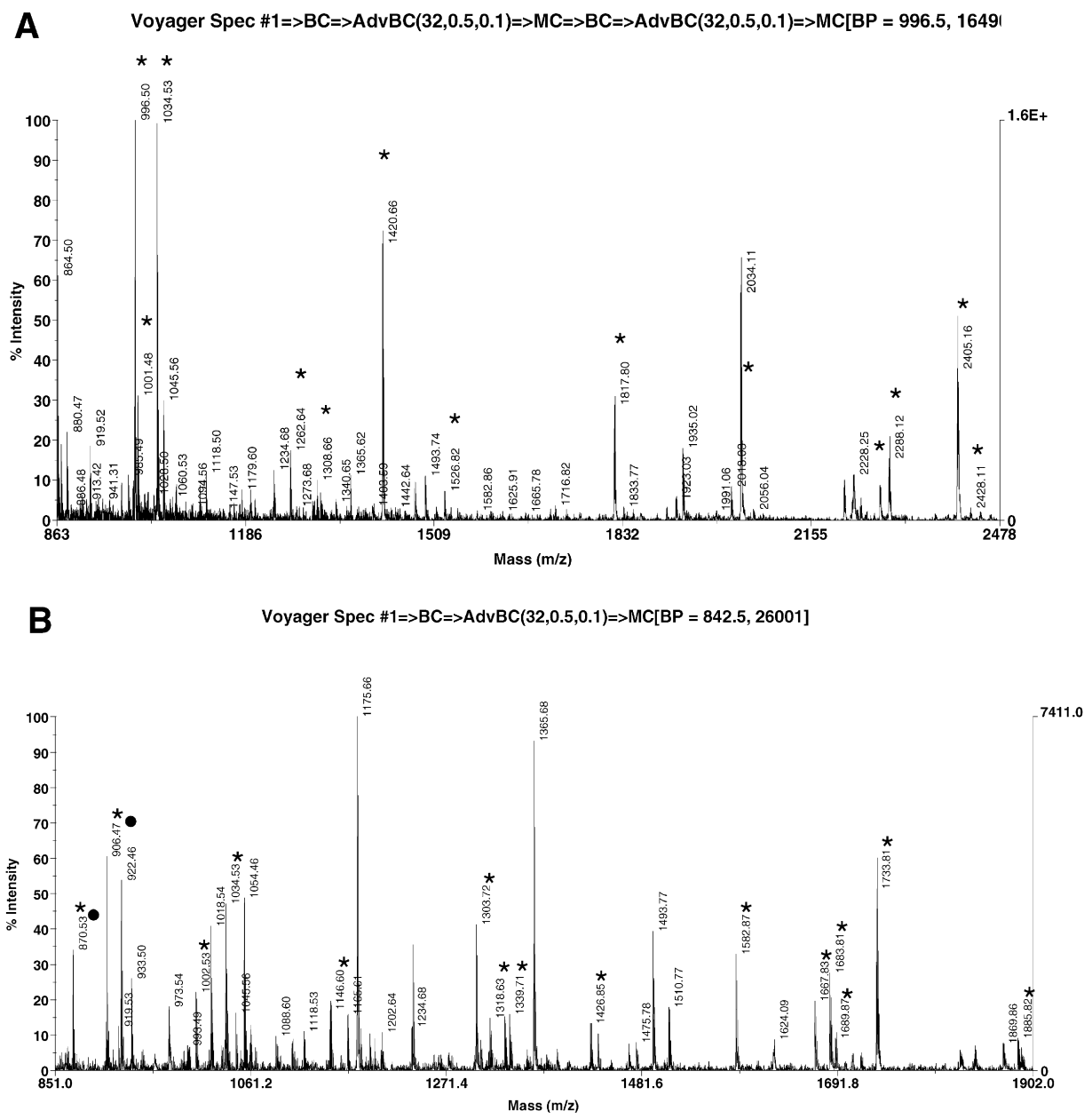

Fig. 5. MS fingerprint analysis of the affinity- and gel-purified protein bands containing LaGT activities. In (A) and (B), the peptides from ions are marked with an asterisk and the correspondent masses (*m*/*z*) were used in the database searches with Protein Prospector MS-FIT 4.0. The peptides from ions *m*/*z* marked with an asterisk can also correspond to trypsin autolysis products. (A) Mass spectrum of the tryptic peptides of the ≈ 35 kDa protein with LaGT2 activity. (B) Mass spectrum of the tryptic peptides of the ≈ 52 kDa protein with LaGT3 activity. (C) Mass spectrum of the LaGT1 UV-irradiated complex band (≈ 18–20 kDa). Peptide standards (Sequazyme Peptide Mass Standards kit, calibration mixture 1 and 2, Applied Biosystems) were used to calibrate the mass scale. (D) Mass spectrum of the unseparated peptide mixture of the LaGT1 UV cross-linked complex band obtained by ESI-MS/MS. The fingerprints shown in (A–C) were obtained by MALDI-TOF MS.

encoding Rbp38p is nuclear and shares high similarity (≈ 72%) with *Tc38* (Accession no. AAQ63938.1), a *Trypanosoma cruzi* gene encoding a ssDNA binding protein [36]. The analysis of the predicted peptide mass from the ≈ 52 kDa protein showed that it was similar to the putative sequences of *Leishmania infantum* and *L. major* replication protein A subunit 1 (*LiRpa-1*, Accession no. AAK84867 and *LmRpa-1*, contig LmjF28-07-20031115_V2.0, respectively) according to the searches in the protein databases (Genpept, NCBInr and *Leishmania* GeneDB, http://www.ebi.ac.uk/parasites/leish.html) (the matched peptides cover 36.4% of the protein). Rpa-1 is a conserved single-stranded binding protein that plays a central role in DNA replication, recombination and repair [37] and is likely to be implicated with telomere maintenance [38]. The analysis of the MALDI-TOF MS spectrum of the trypsin digested LaGT1 UV

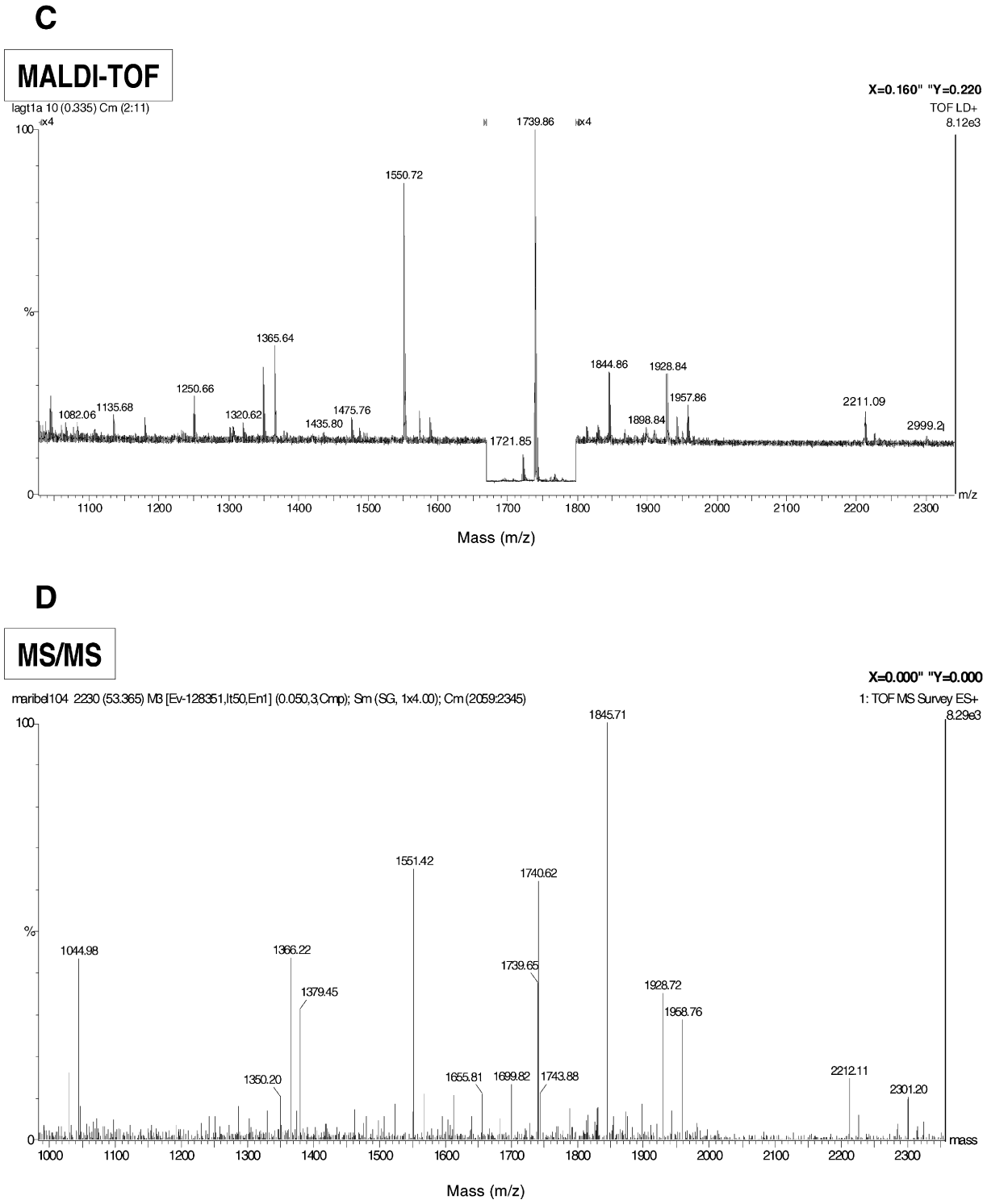

Fig. 5. (Continued).

cross-linked complex (Fig. 5C) shows that there are no homologues for the ≤ 15 kDa protein component in the databases searched, suggesting that it is likely to be a novel *L. amazonensis* protein.

The purification protocols described here permit us to isolate all protein-forming LaGT complexes with a high degree of purity (Table 3, Fig. 4D,E), allowing us to obtain internal peptide sequences of the trypsin digested protein bands using ESI-MS/MS. The sequence tags generated from ESI-MS/MS *de novo* sequencing of LaGT2 and LaGT3 peptides shown in Table 4 corroborate the peptide mass fingerprint identification. However, it was not possible to sequence LaGT1 tryptic peptides, although the analysis of MS/MS mass spectrum of its unseparated peptide

Table 3. Purification scheme for LaGT2 and LaGT3 protein-forming complexes. Specific activity, % binding activity was calculated for 1 μg of protein. The fold purification was calculated based on the specific activity of the fractions assayed after each purification step relative to the S100 extract.

Protein fraction (step)	Fraction volume (mL)	Concentration ($\mu g \cdot \mu L^{-1}$)	Total protein (mg)	Specific activity (% binding·μg protein^{-1})		Purification factor (fold)	
				LaGT2	LaGT3	LaGT2	LaGT3
Whole cell extract (S100)	17.50	2.95	51.62	0.58	0.55	1.00	1.00
Ion exchange chromatography DEAE-agarose (75–200 mM sodium acetate)	1.27	2.35	2.98	6.02	8.37	10.38	15.22
Afinity chromatography (0.6–2.2 M KCl)	0.98	0.16	0.16	26.06	41.80	44.93	76.00

mixture, performed by MAXENT3, undoubtedly confirmed the results of the MALDI-TOF peptide mass fingerprint (compare the MS spectra in Fig. 5C,D).

We are currently trying to identify the LaGT1 protein component and clone the genes encoding LaGT2 and LaGT3 proteins for further functional analysis.

Discussion

In most organisms, including yeast, humans and the protozoa parasite *Leishmania* spp., the telomeric DNA is double-stranded but the 3′ ends are single-stranded and G-rich [1,3,6,39]. This G-rich strand is critical because it is the substrate for telomere replication by telomerase and for the association of proteins that are responsible for chromosome end capping and thus, for regulating telomere length and genome stability [5,11].

Three protein·DNA complexes that associate *in vitro* with the *L. amazonenis* G-rich telomeric strand (LaGT1, LaGT2 and LaGT3) were identified in DEAE column fractions from S100 and nuclear extracts. All complexes did not bind to the C-rich or to the double-stranded form of telomeric DNA, indicating that all LaGT proteins have a preference for the G-rich telomeric sequence. In addition, complex LaGT1 besides being the most specific, interacted with a duplex DNA with a 3′ G-overhang, a feature shared with other single-stranded telomere binding proteins [20,40]. However, the binding activities of LaGT2 and LaGT3 complexes were inhibited by some of the nontelomeric competitors studied. This suggests that, different from LaGT1, they may associate with a variety of sequence targets or recognize a specific DNA structure, such as the β subunit of *On*Tebp [40]. Similarly to the telomeric proteins described in yeast (Est1p), *Clamidomonas reinhardtii* (Gbp1p) and *T. brucei* (complex C3) [28,41,42], LaGT proteins also associated with an RNA cognate sequence of the telomeric DNA. In addition, all LaGT complexes as the *Oxythricha* telomere binding proteins [40] were stable in high salt concentrations, suggesting that these protein–telomeric DNA complexes can be maintained by hydrophobic interactions. Although contacts between DNA and proteins can not be generalized, and each example has its own unique features, hydrophobic interactions could play an important role in the features that govern the *in vitro* interaction of LaGT1 and the G-rich telomeric strand as the complex is formed only at 4 °C. Hydrophobic interactions are favored at lower temperatures contributing to the specificity and affinity of binding [43].

Table 4. LaRbp38 (protein-forming complex LaGT2) and LaRpa-1 (protein-forming complex LaGT3) peptide sequences obtained by ESI-MS/MS.

Protein	*m/z*	Sequence
LaRpa-1	1339.7	IDINPTDLPDVK
	1426.8	EVGSLVDVLGVVLK
LaRbp38	761.4	SSILLTK
	778.4	ENLQFK
	996.5	SNYWLTGR
	1034.5	LQFELNDR
	1262.6	DAELYQWPIK
	1420.6	SNDFNSGLYFTR
	1526.8	KGEAFKLLQDHIK
	1817.8	NSPSNVWIEDWEADR
	2018.0	HLIYNVDQLEDPHLALK
	2405.1	LFHSSQLSGGEALQQYPVSGGSR

Under the conditions used, two proteins of ≈ 35 kDa and ≈ 52 kDa were found in a highly purified form in most of the affinity fractions. Fingerprinting analysis and *de novo* sequencing of the ≈ 35 kDa protein that contained LaGT2 activity showed that it shared identity with the putative amino acid sequence of LmRbp38, a protein first described by Sbicego *et al.* [31] as a novel mitochondrial double-stranded and ssRNA-binding protein that is conserved among kinetoplastid. According to these authors, the gene that encodes Rbp38p is nuclear in *L. tarentolae* and *T. brucei* (Accession no. AAO39843) and the protein does not contain any known RNA-binding motifs. In addition, LaGT2 activity was found in nuclear and S100 extracts of *L. amazonensis* (Fig. 1), and bound single-stranded telomeric DNA and RNA, with higher affinity for RNA. Moreover, the predicted amino acid sequence of the putative *LmRbp38* shares 70% identity with Tc38 (data not shown), a DNA-binding protein that recognizes specifically the motif poly[dT-dG] present in *T. cruzi* intergenic regions [36]. According to Duhagon *et al.* [36], Tc38p may

have nuclear and mitochondrial localization. Thus, it will be interesting to address if LaRbp38p plays a dual role and if they can be differentiated between a possible telomeric and or a mitochondrial function. Coincidently, in *Saccharomyces cerevisiae* the mitochondrial helicase Pif1p, has a nuclear form that coimmunoprecipitates with telomeric DNA *in vivo* and inhibits telomerase activity [44].

MS analysis and the sequence tags m/z 1339.7 and m/z 1426.8 obtained by MS/MS *de novo* sequencing (Fig. 5B, Table 4) of the $\approx$ 52 kDa protein in LaGT3 revealed that it is probably the *L. amazonensis* homologue of *L. infantum* and *L. major* Rpa-1p as the sequence similarities between the LaRpa-1p tags and LiRpa-1p and LmRpa-1p are $\approx$ 100% (data not shown). Rpa-1p, is one of the three subunits of the eukaryotic heterotrimeric complex Rpa that binds to ssDNA mainly by two of the three structural DNA-binding domains located in subunit 1 [45]. BLAST2 pairwise sequence analysis and CLUSTAL W multialignment showed that the putative *L. major* sequence, like all other kinetoplastid sequences annotated as *Rpa1*, lacked the N-terminal domain (data not shown) that in other eukaryotes is involved only in Rpa–protein interactions and has no function in binding DNA [45]. In addition, at the N-terminal of LiRpa-1p and LmRpa-1p there is a region comprising amino acids 23–104, that shares 98% similarity with an oligonucleotide/oligosaccharide-binding (OB) fold structural domain that binds to nucleic acids [46,47]. OB folds were also found in proteins that cap the G-rich telomeric strand and protect the chromosome ends in ciliate protozoa [40,48], human [20] and yeast [49]. This suggests that, in the absence of *in vivo* studies, one can not exclude the possibility that the kinetoplastid Rpa-1p is most likely a novel protozoan single-stranded telomere binding protein related to a large class of proteins that contains the structural OB fold domains. However, there are some evidences suggesting that Rpa plays a role in telomere maintenance. For example, *S. cerevisiae POL12/RPA-1* double mutants show reduced telomere length and decreased viability [38]. In addition, the interaction of the yeast telomere-binding protein Cdc13p with the catalytic subunit of pol α and Rpa directly interferes in telomere metabolism as the disruption of this association in *pol1* alleles induces telomere length [19]. Very recently, Schramke *et al.* [50] showed that in yeast, Rpa is present at the telomeres and activates telomerase by loading Est1p onto telomeres during the S phase. Moreover, hRpa-1p was shown to bind ssDNA and when associated with the Werner and Bloom syndrome helicases, actively unwind long telomeric duplex regions that are pre-bound by Trf2p [51]. Whether the *L. amazonensis* $\approx$ 52 kDa protein identified here is truly a homologue of Rpa-1p or a novel kinetoplastid protein that plays a role in telomere regulation and end capping remains to be determined.

It was not possible to obtain sequence tags from the UV cross-linked LaGT1 complex ($\approx$ 18–20 kDa band), although the mass spectra analysis of the tryptic peptides obtained by MALDI-TOF MS and ESI-MS/MS (Fig. 5C,D) were similar and revealed that its $\leq$ 15 kDa protein component do not have homologues in the protein databases. ESI-MS/MS analyses of covalent protein–nucleic acid complexes can be problematic as the two components possess conflicting requirements for ionization. The combined use of photoactivatable base analogs, such as 5-IdU, and appropriate ESI-MS protocols that have resulted in the successful characterization of different photocrosslinked protein–nucleic acid complexes, may help to identify the LaGT1 protein component [52].

Acknowledgements

The authors thank colleagues in Dr M. I. Cano's laboratory for helpful discussions during the experiments described here, A. C. B. Montes de Oca (Institute of Chemistry, UNICAMP) for helping to determine the protein purification factor and S. Hyslop (School of Medical Sciences, UNICAMP) for critically reading and reviewing the English of the manuscript. This investigation received support from the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR, ID A00753) and from FAPESP (grant 00/01138-6). M. F. F. was supported by a doctoral fellowship from CAPES, Brazil.

References

1. Fu, G. & Barker, D. (1998) Characterisation of *Leishmania* telomeres reveals unusual telomeric repeats and conserved telomere-associated sequence. *Nucleic Acids Res.* **26**, 2161–2167.
2. Henderson, E. (1995) Telomere DNA Structure. In *Telomeres* (Blackburn, E. & Greider, C., eds), pp. 11–34. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
3. Henderson, E. & Blackburn, E.H. (1989) An overhanging-3′ terminus is a conserved feature of telomeres. *Mol. Cell. Biol.* **9**, 345–348.
4. Greider, C. (1996) Telomere length regulation. *Annu. Rev. Biochem.* **65**, 337–365.
5. Chan, S.W.-L. & Blackburn, E.H. (2002) New ways not to make ends meet: telomerase, DNA damage proteins and heterochromatin. *Oncogene* **21**, 553–563.
6. McElligott, R. & Wellinger, R.J. (1997) The terminal DNA structure of mammalian chromosomes. *EMBO J.* **16**, 3705–3714.
7. Wellinger, R.J., Wolf, A.J. & Zakian, V. (1993) *Saccharomyces* telomeres acquire single-strand TG_{1-3} tails late in s phase. *Cell* **72**, 51–60.
8. Wellinger, R.J., Ethier, K., Labrecque, P. & Zakian, V.A. (1996) Evidence for a new step in telomere maintenance. *Cell* **85**, 423–433.
9. de Lange, T. (2002) Protection of mammalian telomeres. *Oncogene* **21**, 532–540.
10. Blackburn, E.H. (2000) Telomere states and cell fates. *Nature* **408**, 53–56.
11. Blackburn, E.H. (2001) Switching and signaling at the telomere. *Cell* **106**, 661–673.
12. Lin, J.J. & Zakian, V.A. (1996) The *Saccharomyces* CDC13 protein is a single-strand TG_{1-3} telomeric DNA-binding protein *in vitro* that affects telomere behavior *in vivo*. *Proc. Natl Acad. Sci. USA* **93**, 13760–13765.
13. Nugent, C., Hughes, T., Lue, N. & Lundblad, V.J. (1996) Cdc13p: a single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science* **274**, 249–252.
14. Chandra, A., Hughes, T.R., Nugent, C.I. & Lundblad, V. (2001) Cdc13 both positively and negatively regulates telomere replication. *Genes Dev.* **15**, 404–414.
15. Evans, S.K. & Lundblad, V.J. (1999) Est1 and Cdc13 as comediators of telomerase access. *Science* **286**, 117–120.
16. Evans, S.K. & Lundblad, V.J. (2000) Positive and negative regulation of telomerase access to the telomere. *J. Cell Sci.* **19**, 3357–3364.

17. Grandin, N., Reed, S.I. & Charbonneau, M. (1997) Stn1, a new *Saccharomyces cerevisiae* protein, is implicated in telomere size regulation in association with Cdc13. *Genes Dev.* **11**, 512–527.
18. Grandin, N., Damon, C. & Charbonneau, M. (2001) Ten1 functions in telomere end protection and length regulation in association with Stn1 and Cdc13. *EMBO J.* **20**, 1173–1183.
19. Qi, H. & Zakian, V.A. (2000) The *Saccharomyces* telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase α and the telomerase-associated Est1 protein. *Genes Dev.* **14**, 1777–1788.
20. Baumann, P. & Cech, T.R. (2001) Pot1, the putative telomere end-binding protein in fission yeast and human. *Science* **292**, 1171–1175.
21. Colgin, L.M., Baran, K., Baumann, P., Cech, T.R. & Reddel, R.R. (2003) Human POT1 facilitates telomere elongation by telomerase. *Curr. Biol.* **13**, 942–946.
22. Loayza, D. & de Lange, T. (2003) POT1 as a terminal transducer of TRF1 telomere length control. *Nature* **26**, 1013–1018.
23. Cano, M.I.N., Dungan, J., Agabian, N. & Blackburn, E.H. (1999) Telomerase in kinetoplastid parasitic protozoa. *Proc. Natl Acad. Sci. USA* **96**, 3616–3621.
24. Cano, M.I.N. (2001) Telomere biology of trypanosomatids: more questions than answers. *Trends Parasitol.* **17**, 425–429.
25. Barry, J.D., Ginger, M.L., Burton, P. & McCulloch, R. (2003) Why are parasite contingency genes often associated with telomeres? *Int. J. Parasitol.* **33**, 29–45.
26. Eid, J. & Sollner-Webb, B. (1995) ST-1 a 39-kilodalton protein in *Trypanosoma brucei*, exhibist a dual affinity for the duplex form of the 29-base-pair subtelomeric repeat and its C-rich strand. *Mol. Cell. Biol.* **15**, 389–397.
27. Eid, J. & Sollner-Webb, B. (1997) ST-2, a Telomere and subtelomere duplex and G-strand binding protein activity in *Trypanosoma brucei*. *J. Biol. Chem.* **272**, 14927–14936.
28. Cano, M.I.N., Blake, J.J., Blackburn, E.H. & Agabian, N. (2002) A *Trypanosoma brucei* protein complex that binds G-overhangs and co-purifies with telomerase activity. *J. Biol. Chem.* **277**, 896–906.
29. Bryan, T.M. & Cech, T.R. (1999) Telomerase and the maintenance of chromosome ends. *Curr. Opin. Cell Biol.* **11**, 318–324.
30. Field, H. & Field, M. (1996) *Leptomonas seymouri, Trypanosoma brucei*: a method for isolating trypanosomatid nuclear factors which bind *T. brucei* single-stranded G-rich telomere sequence. *Exp. Parasitol.* **83**, 155–158.
31. Sbicego, S., Alfonzo, J.D., Estevez, A.M., Rubio, M.A., Kang, X., Turck, C.W., Peris, M. & Simpson, L. (2003) RBP38, a novel RNA-binding protein from trypanosomatid mitochondria, modulates RNA stability. *Eukaryot. Cell* **2**, 560–568.
32. Noll, T.M., Desponds, C., Belli, S.I., Glaser, T.A. & Fasel, N.J. (1997) Histone H1 expression varies during the *Leishmania major* life cycle. *Mol. Biochem. Parasitol.* **84**, 215–227.
33. Sambrook, J. & Russel, D.W. (2001) *Molecular Cloning: a Laboratory Manual*, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
34. Schnapp, G., Rodi, H.-P., Rettig, W.J., Schnapp, A. & Damm, K. (1998) One-step affinity purification protocol for human telomerase. *Nucleic Acids Res.* **26**, 3311–3313.
35. Kim, N., Piatyszek, M., Prowse, K., Harley, C., West, M., Ho, P., Coviello, G., Wright, W., Weinrich, S. & Shay, J. (1994) Specific association of human telomerase activity with immortal cell and cancer. *Science* **266**, 2011–2015.
36. Duhagon, M.A., Dallagiovanna, B., Ciganda, M., Ruyechan, W., Williams, N. & Garat, B. (2003) A novel type of single-stranded nucleic acid binding protein recognizing a highly frequent motif in the intergenic regions of *Trypanosoma cruzi*. *Biochem. Biophys. Res. Commun.* **309**, 183–188.
37. Wold, M.S. (1997) Replication Protein A: a heterotrimeric, ssDNA-binding protein required for eukaryotic DNA metabolism. *Annu. Rev. Biochem.* **66**, 61–92.
38. Smith, J., Zou, H. & Rothstein, R. (2000) Characterization of genetic interactions with RFA1: the role of RPA in DNA replication and telomere maintenance. *Biochimie* **82**, 71–78.
39. Chiurillo, M.A., Cano, M.I., Franco, J.S. & Ramirez, J.L. (1999) Organization of telomeric and sub-telomeric regions of chromosome from the protozoan parasite *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* **100**, 173–183.
40. Price, C.M. & Cech, T.R. (1987) Telomeric DNA–protein interactions of *Oxyticha* macronuclear DNA. *Genes Dev.* **1**, 783–793.
41. Virta-Pearlman, V., Morris, D.K. & Lundblad, V. (1996) Est1 has the properties of a single-stranded telomere end-binding protein. *Genes Dev.* **10**, 3094–3104.
42. Johnston, S., Lew, J. & Berman, J. (1999) Gbp1, a protein with RNA recognition motifs, binds single-stranded telomeric DNA and changes its binding specificity upon dimerization. *Mol. Cell. Biol.* **19**, 923–933.
43. Privalov, P.L. & Gill, S.J. (1988) Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.* **39**, 191–234.
44. Bessler, J.B., Torredagger, J.Z. & Zakian, V.A. (2001) The Pif1p subfamily of helicases: region-specific DNA helicases? *Trends Cell. Biol.* **11**, 60–65.
45. Bochkareva, E., Korolev, S., Less-Miller, S.P. & Bochkarev, A. (2002) Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO J.* **21**, 1855–1863.
46. Murzin, A.G. (1993) OB (oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* **12**, 861–867.
47. Bochkareva, A., Pfuetzner, R.A., Edwards, A.M. & Frappier, L. (1997) Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature* **385**, 176–181.
48. Wang, W., Skopp, R., Scofield, M. & Price, C. (1992) *Euplotes crassus* has genes encoding telomere-binding proteins and telomere-binding protein homologs. *Nucleic Acids Res.* **20**, 6621–6629.
49. Mitton-Fry, R.M., Anderson, E.M., Hughes, T.R., Lundblad, V. & Wuttke, D.S. (2002) Conserved structure for single-stranded telomeric DNA recognition. *Science* **296**, 145–147.
50. Schramke, V., Luciano, P., Brevet, P., Guillot, S., Corda, Y., Longhese, M.P., Gilson, E. & Géli, V. (2004) RPA regulates telomerase action by providing Est1p access to chromosome ends. *Nature Genet.* **36**, 46–54.
51. Opresko, P.L., von Kobbe, C., Laine, J.-P., Harrigan, J., Hickson, I.D. & Bohr, V. (2002) Telomere-binding protein TRF2 binds to and stimulates the Werner and Bloom Syndrome helicases. *J. Biol. Chem.* **25**, 41110–41119.
52. Wong, D., Pavlovich, J.G. & Reich, N.O. (1998) Electrospray ionization mass spectrometric characterization of photocrosslinked DNA-EcoRI DNA methyltransferase complexes. *Nucleic Acids Res.* **26**, 645–649.