Published in final edited form as:

J Proteome Res. 2012 January 1; 11(1): 237–246. doi:10.1021/pr200806s.

Improved proteomic approach for the discovery of potential vaccine targets in *Trypanosoma cruzi*

Ernesto S. Nakayasu^{a,*,#}, Tiago J. P. Sobreira^b, Rafael Torres Jr.^a, Luciane Ganiko^a, Paulo S. L. Oliveira^b, Alexandre F. Marques^a, and Igor C. Almeida^{a,*}

^aDepartment of Biological Sciences, University of Texas at El Paso, 500 W. University Ave., El Paso TX, 79902, USA

^bNational Laboratory for Biosciences (LNBio), National Center for Research in Energy and Materials, Campinas, SP 13083-970, Brazil

Abstract

Chagas disease, caused by *Trypanosoma cruzi*, is a devastating parasitic infection affecting millions of people. Although many efforts have been made for the development of immunotherapies, there is no available vaccine against this deadly infection. One major hurdle for the rational approach to develop a *T. cruzi* vaccine is the limited information about the proteins produced by different phylogenetic lineages, strains, and stages of the parasite. Here, we have adapted a 1D nanoHPLC system to perform online 2D LC-MS/MS, using the autosampler to inject the eluting salt solutions in the first dimension separation. The application of this methodology for the proteomic analysis of the infective trypomastigote stage of *T. cruzi* led to the identification of 1,448 non-redundant proteins. Furthermore, about 14% of the identified sequences comprise surface proteins, most of them glycosylphosphatidylinositol (GPI)-anchored and related to parasite pathogenesis. Immunoinformatic analysis revealed thousands of potential peptides with high affinity for major histocompatibility complex (MHC) class I and II presentation. The high diversity of proteins expressed on trypomastigote surface may have many implications for host-cell invasion and immunoevasion mechanisms triggered by the parasite. Finally, we performed a rational approach to filter potential epitopes that could be further tested and validated for development of a Chagas disease vaccine.

1. Introduction

Chagas disease, or American trypanosomiasis, is considered a neglected infectious disease with an estimated 11 million cases in Latin America, and about 50,000 deaths annually.^{1–3} Chagas disease has become a public health concern in the U.S. and Europe due to the large number of asymptomatic infected people migrating from endemic areas.^{4, 5} In addition, autochthonous cases have already been reported in the U.S.^{4, 6} The etiologic agent of Chagas disease is the protozoan parasite *Trypanosoma cruzi*, which has four life-cycle stages or forms, namely epimastigotes, metacyclic trypomastigotes, amastigotes, and bloodstream trypomastigotes, in two distinct hosts, i.e., triatomine insect and mammal.^{1, 3} In

*To whom correspondence should be addressed: Igor C. Almeida, Phone: 915-747-6086, Fax: 915-747-5808, icalmeida@utep.edu; Ernesto S. Nakayasu, Phone: 915-747-6086, Fax: 915-747-5808, esnakayasu@miners.utep.edu.

[#]Present address: Pacific Northwest National Laboratory, Richland, WA, 99352, USA.

Supporting Information Available: Supplementary table containing the complete set of results. This information is available free of charge via the Internet at <http://pubs.acs.org>.

Data availability: All LC-MS/MS raw datafiles were deposited on Tranche database (www.proteomecommons.org) under hash: mafwayVlhH2YnPPZkxBxJJH+K6od+8sIWqxhCQ5j4tFT/Ua/RmFaWv1PZNNGkWd7eG2 E8Twt5Dg/kPWYU64JbQqr3yO0AAAAAAAACZA==

the natural cycle, infective bloodstream trypomastigote forms are ingested by the triatomine insect vector (popularly known as the *kissing bug*) during a blood meal. In the insect's midgut, trypomastigotes differentiate into noninfective epimastigote forms, which reproduce by binary fission. In the posterior region of the gut (hindgut), upon nutritional stress epimastigotes transform into infective metacyclic trypomastigote forms, which are then released with the excreta during a blood meal. These parasites then penetrate into the host bloodstream through the bite wound or exposed mucosa and infect a variety of nucleated cells. Once inside the cell, the metacyclic trypomastigote immediately escapes the parasitophorous vacuole and transforms into proliferative amastigote form, which replicates by asexual binary division. After several division cycles, amastigotes differentiate into trypomastigotes, which are then released to the extracellular milieu, eventually reaching the bloodstream to infect new cells or be ingested by the kissing bug, completing the natural life cycle.^{1, 3}

Current therapy against Chagas disease is based only in two drugs: benznidazole and nifurtimox. Both these compounds have partial efficacy in the chronic phase of the disease and severe side effects.^{7, 8} Moreover, thus far there is no available vaccine for the treatment and/or prevention of human Chagas disease.^{9–11} In recent years, several studies have focused in the host immune response against *T. cruzi*.^{12–14} Although some promising progress have been made, the difficulty of using a more rational approach to develop a vaccine against *T. cruzi* is the restricted information about proteins expressed on the cell surface of different parasite stages, strains, and phylogenetic lineages.¹⁵

T. cruzi is covered by a thick layer of glycosylphosphatidylinositol (GPI)-anchored molecules, which are encoded by multigene families, such as mucins (TcMUC), mucin-associated surface proteins (MASP), *trans*-sialidase (TS)/gp85, and gp63 glycoproteins, each one of them with hundreds of members.^{16–19} However, the lack of positive correlation between the levels of transcripts and proteins makes difficult the use of approaches such as microarray and cDNA library to map the expressed sequences.²⁰ Thus, the use of proteomic approaches is an interesting and reliable alternative to analyze gene expression in *T. cruzi*. For instance, a large-scale proteomic study of *T. cruzi* Brazil strain identified 1,168 non-redundant proteins (2,784 individual proteins) combining the data from all four stages of the parasite.²¹

Multidimensional liquid chromatography (MDLC) coupled to tandem mass spectrometry (MS/MS) for proteomic analysis was first introduced by John Yates' group and became one of the most popular approaches to analyze large-scale proteomic samples.²² In this approach, proteins from complex mixture samples are chemically or enzymatically digested, resulting in thousands of peptides, which are then analyzed by MDLC-MS/MS. In the first dimension, the separation is generally done in a strong-cation exchange (SCX), strong-anion exchange (SAX), reverse phase (RP), hydrophilic interaction (HILIC), or even mixed bed (SCX and SAX) chromatography capillary column, whereas in the second dimension RP chromatography is widely used for its compatibility with MS solvents.²³

Here, we have adapted a 1D nanoHPLC in a simple way to perform online two-dimensional liquid chromatography coupled to tandem mass spectrometry (2D LC-MS/MS) analysis, using autosampler injections to elute the peptides from the SCX capillary column. We applied this new method for the analysis of the proteome of infective host cell-derived trypomastigote forms of the highly virulent Y strain of *T. cruzi*. Our results showed a substantial (about 2.5-fold) increase in the trypomastigote proteome coverage compared to the previous study.²¹ Moreover, using bioinformatic analysis of proteomic data of all identified proteins and putative surface proteins we have predicted numerous potential

epitopes with theoretical high-binding affinity for class I and II major histocompatibility complex (MHC).

2. Experimental Section

2.1. Cell cultures and protein extraction

Tissue culture-derived trypomastigote forms of *T. cruzi* (Y strain) were obtained 5 to 9 days after infection of green-monkey kidney epithelial cells (LLC-MK₂, ATCC, Manassas, VA), maintained at 37°C, under 5% CO₂ atmosphere, in DMEM supplemented with 10% fetal bovine serum, as described.²⁴ Trypomastigotes were collected by centrifugation (3,000×g for 10 min at 4°C), washed 3 times with phosphate-buffered saline solution (PBS, pH 7.2) and resuspended in PBS-containing protease inhibitor cocktail (Sigma Aldrich). Parasite lysis was performed by sonicating the sample 3× for 15 sec with 30% amplitude, using a Vibra-Cell sonicator (Sonics, Newtown, CT). After the lysis, the sample was centrifuged at 16,000×g for 20 min at 4°C. The soluble protein fraction was separated from the insoluble proteins and quantified with bicinchoninic acid (BCA) kit (Pierce), according to the manufacturer's protocol.

2.2. Protein digestion

After quantification, one mg of soluble proteins and the equivalent number of parasites from the pellet (note: it was not possible to quantify protein in this fraction because its insolubility) were precipitated with 10% trichloroacetic acid for 20 min in an ice bath, centrifuged at 16,000×g for 20 min at 4°C, washed with ice-cold acetone, and dried in a vacuum centrifuge (Vacufuge, Eppendorf). Proteins were redissolved in 200 µL 0.4 M NH₄HCO₃ containing 8 M urea, and disulfide bounds were reduced with 5 mM dithiothreitol for 15 min at 50°C. Then, cysteine residues were alkylated with 10 mM iodoacetamide for 30 min at room temperature, protected from light. Proteins were digested with 20 µg sequencing-grade trypsin (Promega) for 24 h at 37°C. The reaction was stopped by the addition of 10 µL formic acid (FA). Peptides were then desalted in a C18-solid phase extraction cartridge (1 mL, Discovery DSC-18, Supelco, Sigma Aldrich). The cartridge was activated with 4 mL methanol and equilibrated with 4 mL 0.05% trifluoroacetic acid (TFA). After loading and washing with 4 mL 0.05% TFA, the sample was eluted with 2 mL 80% ACN/0.05% TFA and dried in a vacuum centrifuge (Eppendorf).

2.3. Offline peptide fractionation

An equivalent of 50 µg of each sample (dissolved in 25% ACN/0.5% FA) was loaded onto a strong-cation exchange (SCX) microcolumn manufactured with 25 µL POROS 50 HS resin (Applied Biosystems) in 200-µL micropipette tip (Axygen Biosciences), previously equilibrated with 200 µL 25% ACN/0.5% FA. The elution was carried out with 100 µL of increasing NaCl concentrations (0, 10, 20, 40, 60, 80, 100, 150, 200, and 500 mM NaCl in 25% ACN/0.5% FA, pH 2.5; followed by 500 mM NaCl in 20% ACN/10 mM ammonium acetate, pH 7.0). Ensuing fractions were then dried in a vacuum centrifuge and desalted using in-house reverse-phase microcolumns. These were manufactured with 20 µL POROS 50 R2 resin (40 mg/mL suspension in isopropanol) (Applied Biosystems) in 200-µL micropipette tip, and washed with 100 µL methanol. After equilibrating the microcolumn with 200 µL 0.05% TFA, the sample was loaded and washed with 200 µL 0.05% TFA. Peptides were then eluted with 100 µL 80% ACN/0.05% TFA and dried in a vacuum centrifuge (Eppendorf).

2.4. Two-dimensional liquid chromatography-tandem mass spectrometry (2D LC-MS/MS)

Online and offline 2D LC-MS/MS were performed in a nanoLC-1D Plus system (Eksigent, Dublin, CA) connected to an LTQ XL/ETD ESI-MS (Thermo Fisher Scientific, San Jose,

CA). The system comprises one 20- μ L loop, one SCX trap column (5 μ L, Opti-Pak, Optimize Technologies), two C18 trap columns (0.25 μ L, Opti-Pak), and one capillary C18 column (Proteopep II, 10 cm $\times$ 75 μ m, 3 μ m, New Objective) (Figure 1). For the online SCX fractionation, in the first dimension peptides were loaded into the SCX trap column and the peptides were eluted using the autosampler by injecting 20 μ L of increasing salt concentrations (0, 10, 20, 40, 60, 80, 100, 150, 200, and 500 mM NaCl in 5% ACN/0.5% FA, pH 2.5; followed by one injection of 500 mM NaCl in 20% ACN/10mM ammonium acetate, pH 7.0). Eluting peptides were automatically loaded into the C18-trap column, which was then washed with solvent A (solvent A = 2% ACN/0.1% FA) at 2.5 μ L/min flow rate for 130 min. The offline-fractionated peptides were loaded directly onto the C18-trap column, which was then washed with solvent A at 2.5 μ L/min flow rate for 130 min. In the second dimension, the peptides were separated in the capillary C18 column at 300 nL/min flow rate using a linear gradient from 5–40% solvent B over 100 min (solvent B = 80% ACN/0.1% FA), followed by 10 min wash with 100% solvent B and 20 min with 5% solvent B. The electrospray was set to 1.9 kV, and the ion trap was set for a maximum injection time of 100 ms for full scan and 150 ms for MS/MS scan. The 10 most abundant ions were selected for collision-induced dissociation (CID) with an isolation width of 3.0 a.m.u. and normalized collision energy of 35%. Dynamic exclusion was set for fragmenting each ion twice and then excluding it for 1 min.

2.5. Bioinformatic analysis

The MS/MS spectra from peptides from 600 to 3500 Da, at least 15 ions and a minimum of 10 counts were converted to DTA files using Bioworks (v3.3.1, Thermo Fisher Scientific). Then database search was performed using the TurboSequest algorithm against collected sequences from *T. cruzi*, bovine, human keratin, and porcine trypsin (downloaded on March 17th, 2008, from GenBank), in the forward and reverse orientations, forming a dataset of 191,762 sequences. The parameters for database search were as follows: trypsin cleavage at both termini and one missed cleavage allowed; 2 Da for peptide-mass tolerance; 1 Da for fragment-mass tolerance; and cysteine carbamidomethylation and methionine oxidation as fixed and variable modifications, respectively. The datasets were filtered with $dCN \geq 0.085$, protein probability $\leq 1e-3$, and Xcorr of 1.5, 2.2, and 2.7 for singly-, doubly-, and triply-charged peptides, respectively. The datasets were then converted to XML files, which were combined using an in-house written Perl (www.perl.org) script. The algorithm basically assembles the peptides into protein sequences by searching the peptide sequences using Blast algorithm²⁵ and considering only those with 100% identity. The assembly was performed in an Intel Core 2 duo/2.53 MHz computer with 2 GB RAM memory and installed with Debian/Linux (www.debian.org) and Postgresql RDMS (www.postgresql.org/). The redundant protein hits were assembled into protein groups, since it was impossible to determine from which those proteins were actually produced. The abundance of proteins was estimated by exponentially modified protein abundance index (emPAI), as previously described.²⁶

All top hits of each protein group were annotated by GO using GOBlet algorithm (<http://goblet.molgen.mpg.de/cgi-bin/goblet/webapp-goblet.cgi>).²⁷ GO analysis was carried out on October 15th, 2008, against the invertebrate taxonomy and with an *e*-value threshold $\leq 1e-10$. The GPI prediction analysis was performed using FragAnchor (<http://navet.ics.hawaii.edu/~fraganchor/NNHMM/NNHMM.html>)²⁸, as described,²⁹ and only the sequences with high probability were accepted as potential GPI-anchored. The signal peptide prediction was done using SignalP (<http://www.cbs.dtu.dk/services/SignalP/>), using Hidden Markov model.³⁰ The signal peptides and GPI-anchoring sequences of the surface proteins were removed using an in-house written Perl script based on their positions obtained from FragAnchor and SignalP softwares. The epitope analysis was done using the

average prediction binding mode on the Immune Epitope Database and Analysis Resource (IEDB) prediction tool (<http://tools.immuneepitope.org/>).³¹ MHC-I prediction was performed for H-2D^b allele (C57BL/6 mouse) against the whole proteome dataset, whereas the MHC-II, for H-2IA^b allele (C57BL/6 mouse) against the surface protein dataset. Only epitopes with high binding affinity ($IC_{50} \leq 100$ nM) were accepted in our analysis.

3. Results and Discussion

3.1. Implementation of an online 2D LC-MS/MS approach using a 1D LC system

Online 2D LC-MS/MS is widely used for proteomic applications.²³ In this case, the HPLC is composed of at least two binary pumps. One pump is set for the first dimension gradient (i.e., SCX, SAX), whereas the other pump is set for the gradient of the second dimension (RP-LC-MS/MS). When the separation is done in three dimensions, an HPLC with a quaternary pump is required for the first two dimensions of separation.³² Here, we adapted our 1D nanoHPLC in a simple way to perform online 2D LC-MS/MS separations, without the need of having an HPLC with multiple pumps. In our design, the first dimension is driven by a primary HPLC pump set to a 2.5 μ L/min constant flow of 0.1% FA/2% ACN, which basically loads and washes the sample. The peptide mixture is first loaded into an SCX-trap column (5 μ L) and the stepwise salt-gradient is performed by injecting 20 μ L of different concentrations of NaCl solution (4 $\times$ column volume) using the autosampler. The eluted peptides are captured online into one of the two RP-trap columns (0.25 μ L) and washed. In this setup, while the salt gradient is running into the first RP-trap column, a RP gradient is simultaneously being run in the second RP-trap column (Figure 1). In the next run, the RP gradient runs through the first RP-trap column, while the SCX is eluting in the second RP-trap column, thus eliminating time gaps between runs (Figure 1). Moreover, in our approach the number of different elution solvents is not limited by the number of pumps in the HPLC, thus it could be easily adapted to run three-dimensional (3D) LC-MS/MS. To test the possibility of using different elution solvents, the last step of SCX was performed with 20% ACN instead of 5% ACN, and 10 mM ammonium acetate (pH 7.0) instead 0.5% FA (pH 2.5), without noticeable changes in column pressure (data not shown), indicating that the system might be easily adaptable for 3D separations. Next, to estimate the repeatability of the system, we ran the same sample twice (Table 1, sample C and D) and the difference in the number of detected peptides or protein groups were about 10%, which is similar to previous reported by other groups for MDLC-MS/MS.^{32, 33}

Recently, Taylor and colleagues have developed a similar approach to perform 2D LC-MS using a 1D LC-MS system, based on autosampler injections to perform the first dimension elution.³⁴ However, their system differs from ours, since it is based on a vented column separation. In their setup, the column is divided in two parts, separated by a splitter, being the proximal part composed of two trap columns (i.e., SCX followed by RP), and the distal part composed by the analytical RP column. During loading, SCX separation and washing, the flow is directed to the waste by the splitter, then to the separation column and from there to the MS. Thus, both SCX and RP separations cannot be performed simultaneously. We have also tested this vented column setup to perform 2D LC-MS but we found that it was not 100% efficient in washing the salt out (data not shown). This might be due to the fact that during the washing step part of the flow and, consequently, some of the salt went to the analytical column. Another advantage of the system described here is that there are no time gaps between LC-MS runs, maximizing the use of the equipment. On the other hand, the vented column approach might provide a better separation resolution since the dead volumes are smaller. Taken together, we can conclude that the autosampler-based gradient is a simple, robust, and inexpensive alternative to perform online 2D separations. It is also worth pointing out that our approach requires only minimum modifications in the 1D LC-MS/MS system.

3.2. Comparison between offline and online 2D LC-MS/MS

In order to evaluate the performance of our autosampler-based gradient, we compared this methodology with an offline separation, using of *T. cruzi* trypomastigote tryptic digests as the testing sample. Briefly, trypomastigotes were lysed by sonication and centrifuged. The resulting pellet (insoluble proteins) and supernatant (soluble proteins) were digested with trypsin and the equivalent of 20 µg was used for online or offline separations. The offline separation was carried out in SCX microcolumns, whereas the online was performed using the autosampler-based gradient (Figure 2). We could notice that the online separation led to the identification of more protein groups compared to the offline separation (Table 1). In the sample of insoluble proteins, online separation increased by more than 30% the number of identified protein groups, whereas in the soluble protein samples this number increased more than 70% (Table 1).

To investigate why the online separation had a better performance, we summed the number of peptides from all fractions, which includes the peptides eluted in more than one fraction. We obtained similar numbers from the analysis of insoluble proteins (3,970 and 4,098 peptides for online and offline separations, respectively), but a lower number of peptides were detected in the offline separation of the soluble proteins (3,347 peptides against 4,514 and 4,811 peptides for the online separation) (Table 1). However, the number of unique peptides, which does not count the repeated peptides eluted in multiple fractions, was significantly different between distinct fractionation methods. By comparing the number of unique peptides and the sum of peptides from all fractions, which contains redundant peptides eluted in more than one fraction, we can estimate the efficacy of each separation method. The online separations showed that only 13.0–16.7% of the peptides were eluted in more than one fraction, whereas in the offline separations this number was 46.6–58.3% (Table 1). Taken together these results clearly show the superior performance of the autosampler-based online separation in comparison with the offline separation. Interestingly, it was recently shown that the autosampler-based gradient has a superior performance also compared to conventional 2D LC-MS systems.³⁴ This superior performance was proposed to be due to a more efficient delivery of the salt solutions in the autosampler-gradient.³⁴

3.3. Proteomic analysis of trypomastigote forms of *T. cruzi*

Next, to study the proteins produced by trypomastigotes, we combined all five 2D LC-MS/MS runs and assembled the peptides into proteins and the redundant proteins into protein groups. The dataset was further filtered with protein probability ($<5e-4$) and sum of cross-correlation scores ($Xcorr \geq 3.0$) from distinct peptides, resulting in a false-discovery rate of only 1.9% at the protein level. A total of 7,370 individual proteins were identified, which corresponded to 1,616 protein groups. Of this total, 6,154 proteins and 1,448 protein groups were from *T. cruzi*, whereas 1,216 proteins and 168 protein groups were contaminants from the medium and sample preparation (Table 2 and Tables S1–3).

To functionally categorize the identified protein groups, we performed a gene ontology (GO) annotation. Seventy-one percent of the sequences were annotated for molecular function. The most abundant activity was the enzymatic/catalytic activity (43.9%), with emphasis in hydrolases (22.4%), transferases (10.3%), and oxidoreductases (6.3%). Other well-represented categories were transport (7.5%) and binding activities (37.7%) (Figure 3, Table S4). From the cellular component categories (28.5%), we noticed that most of the proteins were intracellular (24.1%), associated to membranes (8.1%) and organelles (16.7%) (Figure 3, Table S4). Among the biological process categories, over half (55.9%) of the proteins were involved in metabolism (Figure 3, Table S4). Interestingly, 147 (10.2%) protein groups were annotated to be involved in the “interaction between organisms” category, particularly in pathogenesis (Figure 3, Table S4). Most of these proteins were from

the TS/gp85 superfamily (141 proteins), four were hypothetical proteins, and one, the 140/116 kDa antigen. TS/gp85 members have been shown to be essential for parasite virulence.^{16, 35, 36} Some members of this family have *trans*-sialidase activity, being able therefore to transfer sialic acid from the host to parasite glycoconjugates such as mucins, which are involved in the host-cell invasion and escaping of host immunity.^{16, 18, 35} Most members of the TS/gp85 superfamily, nevertheless, do not have *trans*-sialidase activity but play key role in the adhesion and invasion of host cells and modulation of host immune response.^{12–14, 16, 19, 35}

Although not detected by GO annotation, other superfamilies of surface proteins, such as mucins and gp63, are known to be involved in the pathogenesis of Chagas disease; thus, increasing the number of protein groups implicated in the pathogenic process to 170 (11.7%) (Figure 3, Table S4). Mucins and their GPI anchors have been shown to be strong proinflammatory agents that activate Toll-like receptor 2 (TLR2).^{37, 38} Recent data from our lab have shown that the activation of TLR2-mediated pathways leads to significant increase in host-cell invasion by the parasite (Torrecilhas, Nakayasu *et al.*, unpublished data). gp63 are GPI-anchored metalloproteinases that also seem to be implicated in host-cell invasion.³⁹ In sum, our current study identified a high number of surface glycoproteins (i.e., TS/gp85, mucins, and gp63) involved in virulence that are very attractive targets for the development of immunotherapies against *T. cruzi*.

3.4. MHC class I epitope prediction

The treatment of Chagas disease is currently limited to only two drugs that exhibit high toxicity with a medium efficacy at the chronic stage.⁷ Furthermore, there is no effective human vaccine available for the prevention or treatment of Chagas disease.^{9–11} Since *T. cruzi* resides at least in part of its life cycle inside the host cell, a vaccine based on cytotoxic CD8+ T lymphocytes (CTLs) seems to be a very attractive approach.^{13, 14} CTLs play a central role in protective immunity to protozoa leading to long-term resistance to many parasites.⁴⁰ Inside the cells, peptides derived from the pathogen are loaded onto MHC class I molecules in the endoplasmic reticulum and transported to the cell surface to be presented to cytotoxic T cells. This peptide-MHC I complex is recognized by T-cell receptors (TCRs) and is essential for the activation of CTLs. Recently, various tools have been developed to predict optimum peptides that could be presented by each MHC allele.⁴¹ Despite its limitations, epitope prediction analysis by immunoinformatics or computational immunology thus could help finding or narrowing the best candidates for vaccine development.^{42–44} Several recent studies using immunoinformatics have shown very promising results for advancing new immunotherapies against a variety of pathogens, including protozoan parasites (i.e., *T. cruzi* and *Plasmodium vivax*),^{45, 46} bacteria,^{47, 48} and viruses.^{49–51}

Despite its usefulness, thus far most of the tested potential *T. cruzi* immunogens have not been based on the systematic proteomic and MHC prediction analysis. Thus, next we performed a prediction analysis for MHC-I presentation of all identified proteins and protein groups identified from *T. cruzi* trypomastigotes. The prediction was done for presentation by the allele H-2D^b of C57Bl/6 inbred mouse strain, one of the most used for testing experimental vaccines. The analysis resulted in the prediction of 2,133 (from protein groups) and 8,367 (from individual proteins) putative epitopes with high binding affinity (IC₅₀ < 100 nM) (Figure 4).

Since a very high number of predicted epitopes were found, we investigated other ways to further filter the data. Fortier et al. have carried out a detailed peptidomic analysis of cancer cells and showed that the most abundantly expressed genes are the ones that had their respective peptides more frequently presented by MHC class I.⁵² Thus, we decided to filter

out our data and consider only peptide predictions from highly abundant proteins. We used a cutoff of 25% of the most abundant proteins calculated by emPAI. After this filtering step, only 310 predicted peptides remained (Figure 4). We then further investigated whether there were peptides with some identity to mouse or human proteins. All peptides with $\geq 50\%$ identity to human or mouse sequences were also removed from the analysis, resulting in 296 predicted epitopes (Figure 4 and Table S5).

A prediction for MHC class I epitopes of *T. cruzi* TS sequences was previously carried out and showed that immunodominant epitopes vary according to the parasite strain.⁵³ Thus, Martin *et al.* proposed that *T. cruzi* evades cytotoxic T cell response by having a different subset of TS proteins expressed on the cell surface for each parasite cell or strain.⁵³ Of those epitopes we predicted, only one sequence fully matched with the prediction made for Brazil and CL Brenner strains,^{46, 53} supporting the idea that different parasite strains express unique subsets of surface proteins, or even more, genes from different strains might also present significant variations. Thus, this new proteomic dataset of the *T. cruzi* Y strain will also help in the discovery of protein subsets commonly expressed by several strains of the parasite.

3.5. Analysis of surface proteins and MHC class II epitope mapping

By predicting the peptides that may bind to MHC-II molecules, we could drastically reduce the number of experiments required to identifying potential CD4⁺ T cell epitopes. Therefore, it could accelerate the discovery of new vaccine candidates. The MHC II presents peptides derived from endocytosed proteins to helper CD4⁺ T lymphocytes (HTLs), which stimulate cellular and humoral immunity against pathogens.⁵⁴ The surface proteins are generally the primary targets for antibody-based vaccines, since they are exposed and in direct contact with the immune system. Thus, we next focused in the analysis of *T. cruzi* surface proteins and prediction of MHC II presentation. Although the centrifugation is not a high-performance separation technique, as expected we noticed a two- to three-fold enrichment of the surface proteins in the insoluble protein fraction (Table 1). A total of 208 (14.4%) protein groups (1,087 individual proteins, 17.7%) were annotated as surface proteins. The identified surface protein families were TS/gp85 (140 protein groups, 789 proteins), MASP (37 protein groups, 136 proteins), mucin TcMUC II (12 protein groups, 42 proteins), gp63 (11 protein groups, 83 proteins), TolT (4 protein groups, 16 proteins), mucin-like (3 protein groups, 12 proteins), and YASP (recently renamed as TcTASV, for *T. cruzi* trypomastigote, alanine, serine and valine-rich proteins)⁵⁵ (1 protein group, 5 proteins) (Table 2 and Tables S1–3). Interestingly, we found a high number of MASP and TcMUC II mucin sequences. These proteins are difficult to be studied by conventional proteomic analysis, since they are likely to be highly glycosylated and several of their genes seem to be expressed at the same time, decreasing the stoichiometry of individual proteins.^{16, 18, 56} Thus, it is possible that the number of expressed MASPs and TcMUC II mucins, found in our study, is still underestimated.

To determine the number of proteins that might be GPI-anchored, we performed a prediction analysis using the FragAnchor algorithm.²⁸ We found 178 protein groups (12.3%) and 890 individual proteins (14.5%) with high probability to be GPI-anchored (Table 2). Of these entries, only 4 protein groups and 15 proteins were not previously annotated as surface proteins (Table 2). We recently performed a genome-wide GPI-anchoring prediction analysis and found that about 11.9% of the *T. cruzi* genes encodes possible GPI-anchored proteins.²⁹ These numbers are much higher as compared to those in other eukaryotes, which have from 0.5 to 2% predicted GPI-anchored proteins in their expected proteomes.²⁸ The expression of several GPI-anchored surface protein isoforms by the parasite may act as a “backup” escaping mechanism. Thus, even if some of the isoforms are targeted by antibodies, others can be functionally promoting the parasite adhesion and invasion of host cells.

Next, we performed the prediction analysis of putative epitopes for MHC class II presentation using annotated membrane and predicted GPI-anchored proteins. The first analysis led to the identification of ~7,000 and ~36,000 epitopes with very high binding affinity ($IC_{50} \leq 10$ nM) in protein groups and proteins, respectively (Figure 4). Looking carefully at the predicted epitopes, we found a bias that most of these predictions were localized in hydrophobic regions of signal peptides and GPI-anchoring sequences. However, these domains are processed in the endoplasmic reticulum and, therefore, not present in the mature protein. To circumvent this issue, the sequences were cut based on the prediction of signal peptides using SignalP, and GPI-anchoring prediction using FragAnchor. After we performed the MHC class II-binding prediction with the truncated version of the sequences, 1,732 and 8,771 possible epitopes with $IC_{50} < 100$ nM were found in protein groups and individual proteins, respectively (Figure 4).

This high number of predicted epitopes can be biologically relevant; however, it is detrimental for choosing potential targets for vaccine development. Therefore, we filtered our data based on the protein abundance, in the same way done for MHC-I predictions. After filtering out low abundance proteins ($emPAI < 0.5$), only 46 peptides were predicted to be presented by MHC-II molecules (Figure 4). After further filtering out peptides with similarity to human or mouse sequences ($\geq 50\%$ identity), 45 epitopes were found to be the best candidates for vaccine targets (Figure 4).

Here we show that *T. cruzi* trypomastigotes concomitantly express hundreds of GPI-anchored glycoproteins on the cell surface. Our data corroborate and significantly expand previous observations of the trypomastigote proteome.²¹ Furthermore, our in silico prediction analysis indicates that thousands of sequences are potential MHC class I epitopes. We thus hypothesize that the extremely high number of proteins expressed on the parasite cell surface may lead to a “dilution” of epitopes being presented to host CD8⁺ T cells, thus leading to ineffective cytotoxic T cell response, which facilitates the parasite perpetuation within the host. A recent study also showed that antibodies against one specific TS sequence failed to inhibit the enzymatic activity of the whole cell lysate, supporting the idea that many sequences are expressed at the same time.⁵⁷ Moreover, antibodies against gp63 or Tc85 only partially inhibit host cell invasion, probably because they recognize only a small subset of these molecules.^{39, 58}

Taking into consideration that: i) with very rare exceptions, patients with acute or chronic Chagas disease, from diverse geographical locations in Latin America, have very high titers of trypanolytic anti- α -galactosyl (anti- α -Gal) antibodies;^{59–67} ii) protective anti- α -Gal antibodies are thought to be one of the major host immune mechanisms responsible for controlling the parasitemia at both acute and chronic stages of Chagas disease;^{59–61, 68} and iii) highly immunogenic α -Gal-containing epitopes are abundantly expressed on major immunodominant GPI-anchored glycoproteins of the mammal-dwelling *T. cruzi* trypomastigote forms^{56, 59, 69, 70} and normally absent in human tissues,⁷¹ we propose that an efficient human vaccine against *T. cruzi* should include these immunodominant α -Gal-containing B-cell epitopes as well as T cell epitopes, universally expressed in different parasite strains and lineages.

3.6. Conclusion

We have developed a simple and robust approach to perform online 2D LC-MS/MS analysis, using autosampler injections to elute peptides from the first dimension column. The application of this methodology to analyze the whole cell lysate of *T. cruzi* trypomastigotes led to the identification of 1,448 non-redundant proteins from *T. cruzi*. Our results show that the trypomastigote stage of *T. cruzi* is rich in cell surface virulence factors, such TS/gp85 and mucins. The expression of several isoforms of these virulence factors may

play a central role during the process of host cell invasion and immunomodulation by the parasite. Furthermore, a rational immunoinformatic prediction analysis of putative MHC epitopes identified various peptides that could be used for vaccine targets. We hope that these findings could help in the design and development of more effective immunotherapeutics for preventing and/or treating Chagas disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Dr. Daniel M. Lorenzini for the critical reading of the manuscript. This work was funded by grants 1R01AI070655-04, 3R01AI070655-04S1, 5S06GM08012-37, 2G12RR008124-16A1, and 2G12RR008124-16A1S1, from the National Institutes of Health. T.J.P.S was supported by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). E.S.N. was partially supported by the George A. Krutilek memorial graduate scholarship from Graduate School, UTEP. We thank the Biomolecule Analysis Core Facility at the Border Biomedical Research Center/Biology/UTEP (NIH grants 2G12RR008124-16A1, and 2G12RR008124-16A1S1) for the access to the LC-MS instruments.

References

- Barrett MP, Burchmore RJ, Stich A, Lazzari JO, Frasch AC, Cazzulo JJ, Krishna S. The trypanosomiasis. *Lancet*. 2003; 362(9394):1469–1480. [PubMed: 14602444]
- Dias JC, Silveira AC, Schofield CJ. The impact of Chagas disease control in Latin America: a review. *Mem Inst Oswaldo Cruz*. 2002; 97(5):603–612. [PubMed: 12219120]
- Stuart K, Brun R, Croft S, Fairlamb A, Gurtler RE, McKerrow J, Reed S, Tarleton R. Kinetoplastids: related protozoan pathogens, different diseases. *J Clin Invest*. 2008; 118(4):1301–1310. [PubMed: 18382742]
- Bern C, Montgomery SP, Herwaldt BL, Rassi A Jr, Marin-Neto JA, Dantas RO, Maguire JH, Acquatella H, Morillo C, Kirchhoff LV, Gilman RH, Reyes PA, Salvatella R, Moore AC. Evaluation and treatment of chagas disease in the United States: a systematic review. *JAMA*. 2007; 298(18):2171–2181. [PubMed: 18000201]
- Piron M, Verges M, Munoz J, Casamitjana N, Sanz S, Maymo RM, Hernandez JM, Puig L, Portus M, Gascon J, Saulea S. Seroprevalence of *Trypanosoma cruzi* infection in at-risk blood donors in Catalonia (Spain). *Transfusion*. 2008; 48(9):1862–1868. [PubMed: 18522707]
- Tarleton RL, Reithinger R, Urbina JA, Kitron U, Gurtler RE. The challenges of Chagas Disease--grim outlook or glimmer of hope. *PLoS Med*. 2007; 4(12):e332. [PubMed: 18162039]
- Bern C. Antitrypanosomal therapy for chronic Chagas' disease. *N Engl J Med*. 2011; 364(26):2527–2534. [PubMed: 21714649]
- Urbina JA. Specific chemotherapy of Chagas disease: relevance, current limitations and new approaches. *Acta tropica*. 2010; 115(1–2):55–68. [PubMed: 19900395]
- Cazorla SI, Frank FM, Malchiodi EL. Vaccination approaches against *Trypanosoma cruzi* infection. Expert review of vaccines. 2009; 8(7):921–935. [PubMed: 19538117]
- Vazquez-Chagoyan JC, Gupta S, Garg NJ. Vaccine Development Against *Trypanosoma cruzi* and Chagas Disease. *Advances in parasitology*. 2011; 75:121–146. [PubMed: 21820554]
- Dumonteil E. DNA Vaccines against Protozoan Parasites: Advances and Challenges. *J Biomed Biotechnol*. 2007; 2007(6):90520. [PubMed: 17710244]
- Tarleton RL. Immune system recognition of *Trypanosoma cruzi*. *Curr Opin Immunol*. 2007; 19(4):430–434. [PubMed: 17651955]
- Padilla AM, Bustamante JM, Tarleton RL. CD8+ T cells in *Trypanosoma cruzi* infection. *Curr Opin Immunol*. 2009; 21(4):385–390. [PubMed: 19646853]
- Junqueira C, Caetano B, Bartholomeu DC, Melo MB, Ropert C, Rodrigues MM, Gazzinelli RT. The endless race between *Trypanosoma cruzi* and host immunity: lessons for and beyond Chagas disease. *Expert reviews in molecular medicine*. 2010; 12:e29. [PubMed: 20840799]

15. Almeida, IC.; Nakayasu, ES. Open Parasitol J. Vol. Vol. 4. Bentham Open; 2010. Subcellular proteomics and global analysis of posttranslational modifications to study functional roles of *Trypanosoma cruzi* molecules; p. 167-177.
16. Acosta-Serrano, A.; Hutchinson, C.; Nakayasu, ES.; Almeida, IC.; Carrington, M. Comparison and evolution of the surface architecture of trypanosomatid parasites. In: Barry, JD.; Mottram, JC.; McCulloch, R.; Acosta-Serrano, A., editors. *Trypanosomes: After the genome*. Norwich, UK: Horizon Scientific Press; 2007. p. 319-337.
17. El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran AN, Ghedin E, Worthey EA, Delcher AL, Blandin G, Westenberger SJ, Caler E, Cerqueira GC, Branche C, Haas B, Anupama A, Arner E, Aslund L, Attipoe P, Bontempi E, Bringaud F, Burton P, Cadag E, Campbell DA, Carrington M, Crabtree J, Darban H, da Silveira JF, de Jong P, Edwards K, Englund PT, Fazelina G, Feldblyum T, Ferella M, Frasch AC, Gull K, Horn D, Hou L, Huang Y, Kindlund E, Klingbeil M, Kluge S, Koo H, Lacerda D, Levin MJ, Lorenzi H, Louie T, Machado CR, McCulloch R, McKenna A, Mizuno Y, Mottram JC, Nelson S, Ochaya S, Osoegawa K, Pai G, Parsons M, Pentony M, Pettersson U, Pop M, Ramirez JL, Rinta J, Robertson L, Salzberg SL, Sanchez DO, Seyler A, Sharma R, Shetty J, Simpson AJ, Sisk E, Tammi MT, Tarleton R, Teixeira S, Van Aken S, Vogt C, Ward PN, Wickstead B, Wortman J, White O, Fraser CM, Stuart KD, Andersson B. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science*. 2005; 309(5733):409–415. [PubMed: 16020725]
18. Buscaglia CA, Campo VA, Frasch AC, Di Noia JM. *Trypanosoma cruzi* surface mucins: host-dependent coat diversity. *Nat Rev Microbiol*. 2006; 4(3):229–236. [PubMed: 16489349]
19. Eugenia Giorgi M, de Lederkremer RM. Trans-sialidase and mucins of *Trypanosoma cruzi*: an important interplay for the parasite. *Carbohydrate research*. 2011; 346(12):1389–1393. [PubMed: 21645882]
20. Jager AV, De Gaudenzi JG, Cassola A, D'Orso I, Frasch AC. mRNA maturation by two-step trans-splicing/polyadenylation processing in trypanosomes. *Proc Natl Acad Sci U S A*. 2007; 104(7): 2035–2042. [PubMed: 17267594]
21. Atwood JA 3rd, Weatherly DB, Minning TA, Bundy B, Cavola C, Opperdoes FR, Orlando R, Tarleton RL. The *Trypanosoma cruzi* proteome. *Science*. 2005; 309(5733):473–476. [PubMed: 16020736]
22. Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol*. 2001; 19(3):242–247. [PubMed: 11231557]
23. Motoyama A, Yates JR 3rd. Multidimensional LC separations in shotgun proteomics. *Anal Chem*. 2008; 80(19):7187–7193. [PubMed: 18826178]
24. Andrews NW, Colli W. Adhesion and interiorization of *Trypanosoma cruzi* in mammalian cells. *J Protozool*. 1982; 29(2):264–269. [PubMed: 7047731]
25. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997; 25(17):3389–3402. [PubMed: 9254694]
26. Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics*. 2005; 4(9):1265–1272. [PubMed: 15958392]
27. Hennig S, Groth D, Lehrach H. Automated Gene Ontology annotation for anonymous sequence data. *Nucleic Acids Res*. 2003; 31(13):3712–3715. [PubMed: 12824400]
28. Poisson G, Chauve C, Chen X, Bergeron A. FragAnchor: a large-scale predictor of glycosylphosphatidylinositol anchors in eukaryote protein sequences by qualitative scoring. *Genomics Proteomics Bioinformatics*. 2007; 5(2):121–130. [PubMed: 17893077]
29. Nakayasu ES, Yashunsky DV, Nohara LL, Torrecilhas AC, Nikolaev AV, Almeida IC. GPIomics: global analysis of glycosylphosphatidylinositol-anchored molecules of *Trypanosoma cruzi*. *Mol Syst Biol*. 2009; 5:261. [PubMed: 19357640]
30. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol*. 2004; 340(4):783–795. [PubMed: 15223320]

31. Zhang Q, Wang P, Kim Y, Haste-Andersen P, Beaver J, Bourne PE, Bui HH, Buus S, Frankild S, Greenbaum J, Lund O, Lundegaard C, Nielsen M, Ponomarenko J, Sette A, Zhu Z, Peters B. Immune epitope database analysis resource (IEDB-AR). *Nucleic Acids Res.* 2008; 36(Web Server issue):W513–W518. [PubMed: 18515843]
32. Wei J, Sun J, Yu W, Jones A, Oeller P, Keller M, Woodnutt G, Short JM. Global proteome discovery using an online three-dimensional LC-MS/MS. *J Proteome Res.* 2005; 4(3):801–808. [PubMed: 15952726]
33. Faca V, Pitteri SJ, Newcomb L, Glukhova V, Phanstiel D, Krasnoselsky A, Zhang Q, Struthers J, Wang H, Eng J, Fitzgibbon M, McIntosh M, Hanash S. Contribution of protein fractionation to depth of analysis of the serum and plasma proteomes. *J Proteome Res.* 2007; 6(9):3558–3565. [PubMed: 17696519]
34. Taylor P, Nielsen PA, Trelle MB, Horning OB, Andersen MB, Vorm O, Moran MF, Kislinger T. Automated 2D Peptide Separation on a 1D Nano-LC-MS System. *J Proteome Res.* 2009
35. Frasch AC. Functional diversity in the trans-sialidase and mucin families in *Trypanosoma cruzi*. *Parasitol Today.* 2000; 16(7):282–286. [PubMed: 10858646]
36. Schenkman S, Eichinger D, Pereira ME, Nussenzweig V. Structural and functional properties of *Trypanosoma* trans-sialidase. *Annu Rev Microbiol.* 1994; 48:499–523. [PubMed: 7826016]
37. Almeida IC, Camargo MM, Procopio DO, Silva LS, Mehlert A, Travassos LR, Gazzinelli RT, Ferguson MA. Highly purified glycosylphosphatidylinositols from *Trypanosoma cruzi* are potent proinflammatory agents. *EMBO J.* 2000; 19(7):1476–1485. [PubMed: 10747016]
38. Campos MA, Almeida IC, Takeuchi O, Akira S, Valente EP, Procopio DO, Travassos LR, Smith JA, Golenbock DT, Gazzinelli RT. Activation of Toll-like receptor-2 by glycosylphosphatidylinositol anchors from a protozoan parasite. *J Immunol.* 2001; 167(1):416–423. [PubMed: 11418678]
39. Cuevas IC, Cazzulo JJ, Sanchez DO. gp63 homologues in *Trypanosoma cruzi*: surface antigens with metalloprotease activity and a possible role in host cell infection. *Infect Immun.* 2003; 71(10):5739–5749. [PubMed: 14500495]
40. Jordan KA, Hunter CA. Regulation of CD8+ T cell responses to infection with parasitic protozoa. *Experimental parasitology.* 2010; 126(3):318–325. [PubMed: 20493842]
41. Lundegaard C, Lund O, Kesmir C, Brunak S, Nielsen M. Modeling the adaptive immune system: predictions and simulations. *Bioinformatics.* 2007; 23(24):3265–3275. [PubMed: 18045832]
42. Flower DR, Macdonald IK, Ramakrishnan K, Davies MN, Doytchinova IA. Computer aided selection of candidate vaccine antigens. *Immunome research.* 2010; 6 Suppl 2:S1. [PubMed: 21067543]
43. Li Pira G, Ivaldi F, Moretti P, Manca F. High throughput T epitope mapping and vaccine development. *J Biomed Biotechnol.* 2010; 2010 325720.
44. Lafuente EM, Reche PA. Prediction of MHC-peptide binding: a systematic and comprehensive overview. *Current pharmaceutical design.* 2009; 15(28):3209–3220. [PubMed: 19860671]
45. Lima-Junior JC, Banic DM, Tran TM, Meyer VS, De-Simone SG, Santos F, Porto LC, Marques MT, Moreno A, Barnwell JW, Galinski MR, Oliveira-Ferreira J. Promiscuous T-cell epitopes of *Plasmodium* merozoite surface protein 9 (PvMSP9) induces IFN-gamma and IL-4 responses in individuals naturally exposed to malaria in the Brazilian Amazon. *Vaccine.* 2010; 28(18):3185–3191. [PubMed: 20189487]
46. Alvarez MG, Postan M, Weatherly DB, Albareda MC, Sidney J, Sette A, Olivera C, Armenti AH, Tarleton RL, Laucella SA. HLA Class I-T Cell Epitopes from trans-Sialidase Proteins Reveal Functionally Distinct Subsets of CD8 T Cells in Chronic Chagas Disease. *PLoS Negl Trop Dis.* 2008; 2(9):e288. [PubMed: 18846233]
47. Gupta SK, Smita S, Sarangi AN, Srivastava M, Akhoun BA, Rahman Q. In silico CD4+ T-cell epitope prediction and HLA distribution analysis for the potential proteins of *Neisseria meningitidis* Serogroup B—a clue for vaccine development. *Vaccine.* 2010; 28(43):7092–7097. [PubMed: 20716448]
48. Zhu YH, Gao YF, Chen F, Liu W, Zhai MX, Zhai WJ, Qi YM, Ye Y. Identification of novel T cell epitopes from efflux pumps of *Mycobacterium tuberculosis*. *Immunology letters.* 2011

49. Diaz I, Pujols J, Ganges L, Gimeno M, Darwich L, Domingo M, Mateu E. In silico prediction and ex vivo evaluation of potential T-cell epitopes in glycoproteins 4 and 5 and nucleocapsid protein of genotype-I (European) of porcine reproductive and respiratory syndrome virus. *Vaccine*. 2009; 27(41):5603–5611. [PubMed: 19646408]
50. Soam SS, Bhasker B, Mishra BN. Improved prediction of MHC class I binders/non-binders peptides through artificial neural network using variable learning rate: SARS corona virus, a case study. *Advances in experimental medicine and biology*. 2011; 696:223–229. [PubMed: 21431562]
51. Singh AK, Rath SK, Misra K. Identification of epitopes in Indian human papilloma virus 16 E6: A bioinformatics approach. *Journal of virological methods*. 2011
52. Fortier MH, Caron E, Hardy MP, Voisin G, Lemieux S, Perreault C, Thibault P. The MHC class I peptide repertoire is molded by the transcriptome. *J Exp Med*. 2008; 205(3):595–610. [PubMed: 18299400]
53. Martin DL, Weatherly DB, Laucella SA, Cabinian MA, Crim MT, Sullivan S, Heiges M, Craven SH, Rosenberg CS, Collins MH, Sette A, Postan M, Tarleton RL. CD8+ T-Cell responses to *Trypanosoma cruzi* are highly focused on strain-variant trans-sialidase epitopes. *PLoS Pathog*. 2006; 2(8):e77. [PubMed: 16879036]
54. O'Shea JJ, Paul WE. Mechanisms underlying lineage commitment and plasticity of helper CD4+ T cells. *Science*. 2010; 327(5969):1098–1102. [PubMed: 20185720]
55. Garcia EA, Ziliani M, Agüero F, Bernabo G, Sanchez DO, Tekiel V. TcTASV: a novel protein family in *trypanosoma cruzi* identified from a subtractive trypomastigote cDNA library. *PLoS Negl Trop Dis*. 2010; 4(10)
56. Buscaglia CA, Campo VA, Di Noia JM, Torrecilhas AC, De Marchi CR, Ferguson MA, Frasch AC, Almeida IC. The surface coat of the mammal-dwelling infective trypomastigote stage of *Trypanosoma cruzi* is formed by highly diverse immunogenic mucins. *J Biol Chem*. 2004; 279(16):15860–15869. [PubMed: 14749325]
57. Ratier L, Urrutia M, Paris G, Zarebski L, Frasch AC, Goldbaum FA. Relevance of the diversity among members of the *Trypanosoma cruzi* trans-sialidase family analyzed with camelids single-domain antibodies. *PLoS ONE*. 2008; 3(10):e3524. [PubMed: 18949046]
58. Alves MJ, Abuin G, Kuwajima VY, Colli W. Partial inhibition of trypomastigote entry into cultured mammalian cells by monoclonal antibodies against a surface glycoprotein of *Trypanosoma cruzi*. *Mol Biochem Parasitol*. 1986; 21(1):75–82. [PubMed: 3534565]
59. Almeida IC, Ferguson MA, Schenkman S, Travassos LR. Lytic anti- α -galactosyl antibodies from patients with chronic Chagas' disease recognize novel O-linked oligosaccharides on mucin-like glycosyl-phosphatidylinositol-anchored glycoproteins of *Trypanosoma cruzi*. *Biochem J*. 1994; 304(Pt 3):793–802. [PubMed: 7818483]
60. Almeida IC, Milani SR, Gorin PA, Travassos LR. Complement-mediated lysis of *Trypanosoma cruzi* trypomastigotes by human anti- α -galactosyl antibodies. *J Immunol*. 1991; 146(7):2394–2400. [PubMed: 1706399]
61. Gazzinelli RT, Pereira ME, Romanha A, Gazzinelli G, Brener Z. Direct lysis of *Trypanosoma cruzi*: a novel effector mechanism of protection mediated by human anti-gal antibodies. *Parasite Immunol*. 1991; 13(4):345–356. [PubMed: 1717927]
62. Avila JL, Rojas M, Galili U. Immunogenic Gal α 1-3Gal carbohydrate epitopes are present on pathogenic American *Trypanosoma* and *Leishmania*. *J Immunol*. 1989; 142(8):2828–2834. [PubMed: 2467941]
63. Antas PR, Medrano-Mercado N, Torrico F, Ugarte-Fernandez R, Gomez F, Correa Oliveira R, Chaves AC, Romanha AJ, Araujo-Jorge TC. Early, intermediate, and late acute stages in Chagas' disease: a study combining anti-galactose, IgG, specific serodiagnosis, and polymerase chain reaction analysis. *The American journal of tropical medicine and hygiene*. 1999; 61(2):308–314. [PubMed: 10463685]
64. Gonzalez J, Neira I, Gutierrez B, Anaconda D, Manque P, Silva X, Marin S, Sagua H, Vergara U. Serum antibodies to *Trypanosoma cruzi* antigens in Atacamenos patients from highland of northern Chile. *Acta tropica*. 1996; 60(4):225–236. [PubMed: 8659322]

65. Medrano-Mercado N, Luz MR, Torrico F, Tapia G, Van Leuven F, AraujoJorge TC. Acute-phase proteins and serologic profiles of chagasic children from an endemic area in Bolivia. *The American journal of tropical medicine and hygiene*. 1996; 54(2):154–161. [PubMed: 8619440]
66. Almeida IC, Covas DT, Soussumi LM, Travassos LR. A highly sensitive and specific chemiluminescent enzyme-linked immunosorbent assay for diagnosis of active *Trypanosoma cruzi* infection. *Transfusion*. 1997; 37(8):850–857. [PubMed: 9280332]
67. de Andrade AL, Zicker F, de Oliveira RM, Almeida Silva S, Luquetti A, Travassos LR, Almeida IC, de Andrade SS, de Andrade JG, Martelli CM. Randomised trial of efficacy of benznidazole in treatment of early *Trypanosoma cruzi* infection. *Lancet*. 1996; 348(9039):1407–1413. [PubMed: 8937280]
68. Pereira-Chioccia VL, Acosta-Serrano A, Correia de Almeida I, Ferguson MA, Souto-Padron T, Rodrigues MM, Travassos LR, Schenkman S. Mucin-like molecules form a negatively charged coat that protects *Trypanosoma cruzi* trypomastigotes from killing by human anti- α -galactosyl antibodies. *J Cell Sci*. 2000; 113(Pt 7):1299–1307. [PubMed: 10704380]
69. Almeida IC, Krautz GM, Krettli AU, Travassos LR. Glycoconjugates of *Trypanosoma cruzi*: a 74 kD antigen of trypomastigotes specifically reacts with lytic anti- α -galactosyl antibodies from patients with chronic Chagas disease. *J Clin Lab Anal*. 1993; 7(6):307–316. [PubMed: 8277354]
70. Couto AS, Goncalves MF, Colli W, de Lederkremer RM. The N-linked carbohydrate chain of the 85-kilodalton glycoprotein from *Trypanosoma cruzi* trypomastigotes contains sialyl, fucosyl and galactosyl (α 1–3)galactose units. *Mol Biochem Parasitol*. 1990; 39(1):101–107. [PubMed: 2106074]
71. Macher BA, Galili U. The Gal α 1,3Gal β 1,4GlcNAc-R (α -Gal) epitope: a carbohydrate of unique evolution and clinical relevance. *Biochim Biophys Acta*. 2008; 1780(2):75–88. [PubMed: 18047841]

Figure 1.

Autosampler-based gradient in a two-dimensional liquid chromatography tandem mass spectrometry system. In the first cycle, the tryptic digest is loaded into the SCX trap column and the unbound peptides are captured into the RP-trap column 1 (RP1). In the second cycle, the autosampler injects salt solution in the SCX and the eluting peptides are captured in the RP-trap column 2 (RP2), while the acetonitrile gradient is running through the RP1 and analytical column (RP3). In the next cycle, a higher-concentration salt solution is injected into the SCX-trap column and the peptides are trapped into RP2, while the acetonitrile gradient is running through the RP1/RP3. The second and third cycle are alternately repeated until all 11 SCX elution steps are analyzed.

Figure 2.

Workflow of *T. cruzi* trypanomastigote whole-cell lysate analysis by 2D LC-MS/MS. Trypanomastigotes were lysed by sonication and centrifuged for 20 min at $16,000 \times g$ and 4°C. The samples were quantified and digested with trypsin. The resulting peptides were fractionated by SCX chromatography, either online (using the autosampler to perform the elution) or offline (using in-house SCX microcolumn). All the fractions were analyzed by RP-LC-MS/MS and the identified proteins submitted to GO and epitope prediction analyses.

Figure 3.

Gene Ontology analysis. GO analysis was performed using GOblet algorithm. Selected representative categories for “molecular function”, “cellular component”, and “biological process” are shown in the graph. For the complete list, see Table S4.

Figure 4.
Prediction for major compatibility complex-binding peptides. The workflow of prediction and filtering MHC class I (A) and MHC class II (B) peptides is shown.

Table 1

Performance comparison (number of peptides per fraction) of online and offline fractionation of soluble and insoluble *T. cruzi* proteins.

Fraction	Insoluble Proteins			Soluble Proteins		
	Online Sample A	Offline Sample B	Online Sample C	Online Sample D	Offline Sample E	
0.5% FA *	12	21	16	22	18	
10 mM NaCl / 0.5% FA *	18	108	33	43	59	
20 mM NaCl / 0.5% FA *	13	911	15	22	772	
40 mM NaCl / 0.5% FA *	602	919	834	910	820	
60 mM NaCl / 0.5% FA *	935	518	772	839	434	
80 mM NaCl / 0.5% FA *	562	703	591	641	549	
100 mM NaCl / 0.5% FA *	544	378	508	471	323	
150 mM NaCl / 0.5% FA *	299	327	614	687	245	
200 mM NaCl / 0.5% FA *	425	152	530	548	93	
500 mM NaCl / 0.5% FA *	506	57	517	554	31	
500 mM NaCl / 10 mM NH ₄ Ac / 20% ACN	54	4	84	74	3	
Sum of number peptides in all fractions	3,970	4,098	4,514	4,811	3,347	
Number of unique peptides	3,456	2,796	3,867	4,259	2,115	
% of peptides eluted in more than one fraction	14.9	46.6	16.7	13.0	58.3	
Peptide - FDR (%)	2.6	2.7	1.4	2.8	2.7	
Number of protein groups	1,258	940	1,008	1,097	608	
Surface proteins	149	129	85	87	41	

*. The online fractionation was performed with 5% ACN, whereas the offline fractionation was done with 25% ACN.

Table 2Proteins, surface proteins and predicted GPI-anchored proteins identified in *T. cruzi* trypomastigotes.

Total proteins and protein groups				
	Protein groups		Individual proteins	
<i>T. cruzi</i>	1,448		6,154	
Contaminants*	168		1,216	
Total	1,616		7,370	
Surface proteins and protein groups				
	Protein groups		Individual proteins	
Protein/Family	Surface proteins	GPI-anchored proteins	Surface proteins	GPI-anchored proteins
<i>TS/gp85</i>	140	114	789	620
trans-sialidase	130	111	744	604
Tc85	3	0	7	1
85 kDa surface protein	2	1	10	1
c71	1	1	1	1
gp90	1	0	1	0
gp82	1	1	11	11
surface protein	0	0	2	0
surface protein-2	1	0	8	1
surface glycoprotein	1	0	3	0
amastigote surface protein-2	0	0	1	0
Shed-acute-phase-antigen	0	0	1	1
<i>MASP</i>	37	36	136	134
<i>TcMUC II</i>	12	10	46	42
<i>mucin-like</i>	3	1	12	8
<i>putative surface antigen YASP (TcTASV)</i>	1	1	5	5
<i>ToIT</i>	4	3	16	14
<i>gp63</i>	11	9	83	51
<i>hypothetical proteins</i>		4		12
<i>unknown</i>		0		1
<i>amastigote cytoplasmic antigen</i>		0		1
<i>cytochrome b5-like</i>		0		2
Total	208	178	1087	890

* Contaminant proteins from sample preparation (trypsin and human keratin) and cell culture medium (bovine serum proteins).