From Genome Screening to Creation of Vaccine Against *Trypanosoma cruzi* by Use of Immunoinformatics

Christian Teh-Poot,1 Evelyn Tzec-Arjona,1 Pedro Martínez-Vega,1 Maria Jesus Ramirez-Sierra,1 Miguel Rosado-Vallado,1 and Eric Dumonteil1,2

1Laboratorio de Parasitología, Centro de Investigaciones Regionales Dr Hideyo Noguchi, Universidad Autónoma de Yucatán, Mérida, Mexico; and 2Department of Tropical Medicine, School of Public Health and Tropical Medicine, Tulane University, New Orleans, Louisiana

(See the editorial commentary by Rodrigues and Ersching on pages 175–7.)

Chagas disease is caused by the protozoan parasite *Trypanosoma cruzi*, and activation of CD8+ T cells is crucial for a protective immune response. Therefore, the identification of antigens with major histocompatibility complex class I epitopes is a crucial step for vaccine development against *T. cruzi*. Our aim was to identify novel antigens and epitopes by immunoinformatics analysis of the parasite proteome (12,969 proteins) and to validate their immunotherapeutic potential in infected mice. We identified 172 predicted epitopes, using NetMHC and RANKPEP. The corresponding protein sequences were reanalyzed to generate a consensus prediction, and 26 epitopes were selected for in vivo validation. The interferon γ (IFN-γ) recall response of splenocytes from *T. cruzi*-infected mice confirmed that 10 of 26 epitopes (38%) induced IFN-γ production. The immunotherapeutic potential of a mixture of all 10 peptides was evaluated in infected mice. The therapeutic vaccine was able to control *T. cruzi* infection, as evidenced by reduced parasitemia, cardiac tissue inflammation, and parasite burden and increased survival. These findings illustrate the benefits of this approach for the rapid development of a vaccine against pathogens with large genomes. The identified peptides and the proteins from which they are derived are excellent candidates for the development of a vaccine against *T. cruzi*.

**Keywords.** Chagas disease; parasite; epitope; MHC; immunotherapy; bioinformatics; protozoan.
protection in dogs [11]. More recently, Nakayasu et al [12] combined a proteomic strategy with a bioinformatics prediction of B-cell and T-cell epitopes to identify new vaccine candidates. This immunoinformatics analysis resulted in thousands of peptides with predicted high-affinity binding to major histocompatibility complex (MHC) class I and II molecules, opening the way to the validation of these antigens as vaccine candidates.

T-cell epitope prediction from protein sequences is indeed emerging as a promising strategy for the rapid identification of vaccine candidates [13]. This approach has been used for the identification of epitopes in single proteins [14–16] and small viral or bacterial genomes [17–19], but it has been used to a much less extent in large protozoan genomes [20, 21]. Thus, these tools have the potential to accelerate antigen discovery and vaccine development [8, 13]. T. cruzi has a large genome of 55 Mb, encoding approximately 12 000 proteins [22], and as mentioned above, genome screening for new antigens has been limited to surface proteins. This is based on the assumption that immunogenic proteins need to be localized on the cell surface of the parasite. However, T. cruzi relies on both classical and alternative protein-trafficking pathways [23], so it is likely that many surface proteins are missed by the conventional analyses described above. In addition, a much greater diversity of proteins can be highly immunogenic and potential vaccine candidates [20].

To overcome these limitations, this study focused on the screening of the full T. cruzi proteome with several T-cell epitope prediction tools, to generate consensus analysis and identify novel antigens. The immunogenicity of a subset of predicted epitopes was validated in recall responses assays with splenocytes from T. cruzi–infected mice, and the immunogenic epitopes were tested as an immunotherapeutic vaccine to evaluate their potential to control a T. cruzi infection in mice.

METHODS

Epitope Prediction From T. cruzi Protein Sequences
Sequences of the complete T. cruzi proteome (version 3.3) were obtained from TriTrypDB (http://tritrypdb.org) [22]. These included 12 969 T. cruzi predicted proteins from the non–Esmeraldo-like haplotype and unassigned sequences. In the first step, the protein sequences were analyzed to predict T-cell epitopes for MHC class I H-2Kd and H-2Dd alleles, using RANKPEP [24] and NetMHC [25], generating MHC binding scores for about 45 million overlapping peptides. We used the normalized scores (calculated as percentages of the maximum binding scores) to generate a first consensus based on the average score of the 2 programs for each peptide analyzed. We selected the peptides predicted to have average scores of ≥60% and ≥70% of maximum binding for H-2Dd and H-2Kd molecules, respectively. Because proteins from the trans-sialidase family have been extensively studied [26] and to ensure the identification of novel antigens, epitopes from this family of proteins were excluded from further analysis.

In a second step, we reanalyzed the protein sequences containing these peptides with 6 and 4 different T-cell epitope prediction programs for H-2Dd and H-2Kd, respectively. These programs included SYFPEITHI [27], BIMAS-HLA [28], ProPred-1 [29], MAPP [30], ComPred [31], SVMHC [32], Predep [33], and IEBD [34]. We combined the results from the different programs to calculate an average rank of prediction for each epitope within the proteins, including the results of RANKPEP and NetMHC, to generate consensus predictions [18]. T-cell epitopes predicted by the highest number of programs and the highest average rank were prioritized for subsequent analysis.

In a third step, we analyzed the predicted epitopes with BLAST to identify potential sequence similarity. Epitopes showing >80% of sequence identity with human or mouse proteins were discarded from further analysis to avoid potential autoimmunity, while epitopes conserved among kinetoplastids were further prioritized for validation.

Validation of the Immunogenicity of the Predicted Epitopes
To validate the immunogenicity of the predicted epitopes, we used a recall response assay, using splenocytes from T. cruzi–infected mice, which were stimulated in vitro with the selected peptides. Briefly, BALB/c mice were infected with 500 trypomastigotes of the H1 strain [35]. Five weeks after infection, animals were euthanized, and spleens were extracted. Spleens were macerated in Roswell Park Memorial Institute (RPMI) 1640 culture medium, and the cells were centrifuged at 200 × g for 10 minutes. The cells were resuspended and incubated with 1 mL of 0.16 M NH4Cl, 0.17 M Tris-HCl (pH 7.2) for 5 minutes at 37°C for hemolysis. Splenocytes were then washed 3 times and resuspended in RPMI 1640 medium supplemented with 100 µ/mL penicillin, 100 µg/mL streptomycin, 20 mM sodium pyruvate, 5 µM β-Mercaptoethanol, 10 mM glutamine, and 10% bovine fetal serum. Synthetic peptides (Peptide 2.0, Chantilly, VA) were dissolved at 5 mg/mL in water, isopropyl alcohol, or 10% dimethyl sulfoxide, depending on their sequence. Pools of 1.25 × 106 cells/mL were stimulated with 5 µg/mL of each individual peptide for 72 hours at 37°C with 5% of CO2. Concanavalin A (10 µg/mL) was used as a positive control. After 72 hours, supernatants were collected, and interferon γ (IFN-γ) levels were measured by enzyme-linked immunosorbent assay according to the manufacturer’s instructions (BioLegend, San Diego, CA) [20].

Evaluation of the Immunotherapeutic Activity of the Immunogenic Peptides
To evaluate the potential of the immunogenic epitopes as vaccine candidates, we tested them as a therapeutic vaccine

From Genome to Vaccine Against T. cruzi • JID 2015:211 (15 January) • 259
administered during the acute phase of a *T. cruzi* infection in mice [35, 36]. BALB/c mice were infected with 500 *T. cruzi* parasites and then treated with 2 doses of a pool of all 10 immunogenic peptides (50 µg of each peptide) via intramuscular injection (with 5 µg of monophosphoryl lipid A [MPLA] as an adjuvant) 7 and 14 days after infection. Parasitemia and survival were monitored for up to 50 days after infection [35, 36]. Mice were then euthanized, and their hearts were collected. Cardiovascular tissue damage was assessed by histopathologic analysis of tissue sections stained with eosin and hematoxylin [35, 36]. Inflammation was quantified by image analysis from 5 digital micrographs per animal, using Multispec 3.2 (Purdue University, West Lafayette, IN) as described before [35, 37]. The parasite burden in the heart was assessed by quantitative polymerase chain reaction [38].

**Statistical Analysis**

Parasitemia and inflammatory cell density data are presented as means and standard errors of the mean, and differences between groups were assessed by the Student t test. Survival curves were analyzed with the log–rank test, and parasite burdens between groups were compared by the Mann–Whitney test.

![Figure 1](image_url)

**Figure 1.** Distribution of the number of predicted epitopes as a function of their average major histocompatibility complex (MHC) binding score. Binding scores normalized to maximum binding scores from RANKPEP and NetMHC were used to calculate the average binding score for H-2Dd (top) and H-2Kd alleles (bottom). To ensure the selection of epitopes predicted with a high binding score by both programs, we selected a binding threshold of >70% and >60% for H-2Kd and H-2Dd, respectively (horizontal dotted line). Eighty peptides were selected for H-2Dd and 92 for H-2Kd.

**Ethics Statement**

The study was approved by the Institutional Bioethics Committee of the Universidad Autonoma de Yucatan, and all animal procedures were performed according to national and international guidelines.

**RESULTS**

**Bioinformatics-Based T-Cell Epitope Prediction**

The 12,969 protein sequences of the *T. cruzi* proteome were first analyzed using RANKPEP [24] and NetMHC [25], to predict T-cell epitopes for MHC class I H-2Kd and H-2Dd alleles from about 45 million overlapping peptides. There were 4407 epitopes predicted with a binding score >50% of the maximum binding score for H-2Kd, using RANKPEP, and 96,450 epitopes, using NetMHC. For H-2Dd, there were 2,251 predicted epitopes, using RANKPEP, and 13,725 epitopes, using NetMHC, with a binding score >50% of the maximum. These corresponded to 0.03%–0.2% of all peptides tested.

We then used these normalized scores (calculated as percentages of the maximum binding scores) to generate a first consensus based on the average score of the 2 programs for each peptide analyzed (Figure 1). A total of 1047 epitopes were...
predicted with an average binding score >50% of maximum binding for H-2K^d, and 232 such epitopes were predicted for H-2D^d (Figure 1). We nonetheless selected a higher binding threshold of >70% and >60% for H-2K^d and H-2D^d, respectively (Figure 1). Thus, we selected 92 peptides with a binding score of >70% for H-2K^d (0.002% of all peptides analyzed) and 80 peptides with an average binding score of >60% for H-2D^d (0.0005% of all peptides analyzed). Because proteins from the trans-sialidase family have been extensively studied [26, 39, 40] and to ensure the identification of novel antigens, epitopes from this family of proteins were excluded from further analysis. Twenty-two potential epitopes of the H-2D^d allele were thus discarded, and none for the H-2K^d allele were discarded.

In the second step of the analysis, the complete protein sequences containing these 150 predicted epitopes (58 for H-2D^d and 92 for H-2K^d) were analyzed with several additional T-cell epitope prediction tools to generate consensus predictions. Predicted epitopes were ranked according to the number of programs that predicted them and their average rank of prediction by each program. For the H-2K^d allele, there was a relatively high consensus among predictions of the different programs, with 16 epitopes predicted by 8 of 8 programs and 60 epitopes predicted by 7 of 8 programs (Figure 2). These epitopes had average ranks of prediction of 1.0–8.4. On the other hand, there were greater discrepancies among predictions for the H-2D^d allele. There was no epitope predicted by all programs and only 5 epitopes predicted by 7 of 8 programs (Figure 2). Most epitopes were predicted by 4 of 8 or 5 of 8 programs. H-2D^d epitopes predicted by several programs had an average rank of 1.0–4.0.

Consensus analysis allowed the selection of 16 epitopes for H-2D^d and 14 epitopes for H-2K^d for further analysis, all predicted by several algorithms and with a high average rank. To further select epitopes for in vivo validation, we used BLAST analysis to discard epitopes with a high similarity to human or mouse sequences and thus with some potential to generate autoimmune responses. Two epitopes for H-2D^d and 2 for H-2K^d presented >88% of sequence identity with human or mouse sequences and were discarded from further analysis. A total of 26 epitopes were retained for in vivo validation of their immunogenicity, 14 for the H-2D^d allele and 12 for H-2K^d allele (Table 1). Interestingly, 6 of the selected epitopes were present in >1 T. cruzi protein, so that these 26 epitopes represented 35 parasite proteins. Also, many epitopes presented some level of conservation in other kinetoplastid parasite species, including Leishmania species and African trypanosomes. Only a few proteins containing the predicted epitopes (11 of 35 [31%]) had a putative function, such as metabolic enzymes or structural proteins of the cytoskeleton, and the large majority (24 of 35 [69%]) were hypothetical proteins (Table 1).

**In Vivo Validation of the Immunogenicity of the Predicted Epitopes**

The immunogenicity of the 26 selected epitopes was investigated using a T-cell recall response assay. Mice were infected with T. cruzi, and their splenocytes were collected after 1 month of infection. We then measured the epitope-specific IFN-γ production in response to stimulation of splenocytes with the individual peptides in vitro. For H-2D^d, 3 of 14 predicted epitopes (P6, P8, and P14) induced a peptide-specific production of IFN-γ, as well as 7 of 12 H-2K^d epitopes (P15, P16, P17, P23, P24, P25, and P26; Figure 3). P15 and P24 were the least immunogenic because they induced low-level IFN-γ production, while P14 and P23 were the most immunogenic, inducing a very high level of IFN-γ production. Overall, 10 of 26 predicted epitopes (38%) were immunogenic in T. cruzi–infected mice and could thus be tested as vaccine candidate. These epitopes belonged to 16 T. cruzi proteins, of which 11 (69%) are hypothetical proteins and 5 (31%) have a putative function. Interestingly,

![Figure 2](https://example.com/figure2.png)

**Figure 2.** Proportion of programs predicting the epitopes. Proteins were analyzed with 8 different T-cell epitope prediction programs, and we determined the number of programs predicting each epitope.
<table>
<thead>
<tr>
<th>ID</th>
<th>Epitope</th>
<th>Rank</th>
<th>Position</th>
<th>Gene ID</th>
<th>Function</th>
<th>Conservation</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>EGPGRNIHL</td>
<td>1.0</td>
<td>26</td>
<td>Tc00.1047053508433.80</td>
<td>Hypothetical protein</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>204</td>
<td>Tc00.1047053506955.198</td>
<td>Hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>EGPSKTIVL</td>
<td>1.0</td>
<td>895</td>
<td>Tc00.1047053506957.140</td>
<td>Phosphatidylinositol-4-phosphate-kinase-like, putative</td>
<td>Leishmania species</td>
</tr>
<tr>
<td>P3</td>
<td>GPPIRSVSL</td>
<td>1.1</td>
<td>429</td>
<td>Tc00.1047053504003.70</td>
<td>S-adenosyl-methyltransferase mraW-like protein, putative</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>GGPYTRVVAL</td>
<td>2.4</td>
<td>138</td>
<td>Tc00.1047053504003.70</td>
<td>S-adenosyl-methyltransferase mraW-like protein, putative</td>
<td>T. brucei, Leishmania species</td>
</tr>
<tr>
<td>P5</td>
<td>FGPSRARKL</td>
<td>1.0</td>
<td>575</td>
<td>Tc00.1047053510975.40</td>
<td>Hypothetical protein, conserved</td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td>FGDPRSVVF</td>
<td>3.2</td>
<td>123</td>
<td>Tc00.1047053511885.10</td>
<td>Hypothetical protein, conserved</td>
<td>T. brucei</td>
</tr>
<tr>
<td>P7</td>
<td>GPPIRSVSL</td>
<td>1.0</td>
<td>199</td>
<td>Tc00.1047053503657.20</td>
<td>Hypothetical protein, conserved</td>
<td></td>
</tr>
<tr>
<td>P8</td>
<td>EGPSKTIVL</td>
<td>4.0</td>
<td>1254</td>
<td>Tc00.1047053504131.140</td>
<td>Hypothetical protein, conserved</td>
<td>T. brucei</td>
</tr>
<tr>
<td>P9</td>
<td>FGPSRARKL</td>
<td>1.0</td>
<td>439</td>
<td>Tc00.1047053506813.170</td>
<td>Hypothetical protein, conserved</td>
<td></td>
</tr>
<tr>
<td>P10</td>
<td>EGPYTRVVAL</td>
<td>1.0</td>
<td>575</td>
<td>Tc00.1047053510975.40</td>
<td>Hypothetical protein, conserved</td>
<td></td>
</tr>
<tr>
<td>P11</td>
<td>EGPYTRVVAL</td>
<td>3.2</td>
<td>123</td>
<td>Tc00.1047053511885.10</td>
<td>Hypothetical protein, conserved</td>
<td>T. brucei</td>
</tr>
<tr>
<td>P12</td>
<td>APPRRIPTI</td>
<td>2.4</td>
<td>199</td>
<td>Tc00.1047053508355.380</td>
<td>Ubiquitin hydrolase, putative cysteine peptidase, ClanCA, familyC19, putative</td>
<td></td>
</tr>
<tr>
<td>P13</td>
<td>GGPERIAVTL</td>
<td>4.0</td>
<td>580</td>
<td>Tc00.1047053511367.170</td>
<td>Vacular protein sorting-associated protein 35, putative</td>
<td>Leishmania species</td>
</tr>
<tr>
<td>P14</td>
<td>KLYTVENAL</td>
<td>1.0</td>
<td>158</td>
<td>Tc00.1047053507681.200</td>
<td>Serine carboxy peptidase S28, putative</td>
<td></td>
</tr>
<tr>
<td>P15</td>
<td>SYTERITGL</td>
<td>1.0</td>
<td>111</td>
<td>Tc00.1047053506469.100</td>
<td>Hypothetical protein, conserved</td>
<td>T. brucei</td>
</tr>
<tr>
<td>P16</td>
<td>SYVTIPTSI</td>
<td>1.0</td>
<td>14</td>
<td>Tc00.1047053444777.10</td>
<td>Drug resistance protein, putative</td>
<td>L. donovani</td>
</tr>
<tr>
<td>P17</td>
<td>KYGGTVSAL</td>
<td>1.2</td>
<td>222</td>
<td>Tc00.1047053503823.130</td>
<td>Hypothetical protein, conserved</td>
<td></td>
</tr>
<tr>
<td>P18</td>
<td>YYYAVNITLSL</td>
<td>1.2</td>
<td>325</td>
<td>Tc00.1047053509179.220</td>
<td>Hypothetical protein, conserved</td>
<td>T. brucei, Leishmania species</td>
</tr>
<tr>
<td>P19</td>
<td>FYITVITFL</td>
<td>1.6</td>
<td>254</td>
<td>Tc00.1047053510729.290</td>
<td>Phosphatidylcholine:ceramide cholinephosphotransferase 2, putative</td>
<td></td>
</tr>
<tr>
<td>P20</td>
<td>KYLAAVSAL</td>
<td>1.6</td>
<td>177</td>
<td>Tc00.1047053503847.200</td>
<td>Myosin heavy chain, putative</td>
<td>Leishmania species</td>
</tr>
<tr>
<td>P21</td>
<td>KYNKLITAM</td>
<td>2.9</td>
<td>137</td>
<td>Tc00.1047053510437.10</td>
<td>Dynene heavy chain, putative</td>
<td>Trypanosoma species</td>
</tr>
<tr>
<td>P22</td>
<td>RYAVTVNI</td>
<td>1.0</td>
<td>210</td>
<td>Tc00.1047053507519.190</td>
<td>Protein kinase (pseudogene), putative</td>
<td>Trypanosoma species</td>
</tr>
<tr>
<td>P23</td>
<td>IYLPLQSAL</td>
<td>1.0</td>
<td>210</td>
<td>Tc00.1047053508479.10</td>
<td>Protein kinase (pseudogene), putative</td>
<td></td>
</tr>
<tr>
<td>P24</td>
<td>SYIGLGSAL</td>
<td>1.4</td>
<td>243</td>
<td>Tc00.1047053508799.280</td>
<td>Hypothetical protein, conserved</td>
<td></td>
</tr>
<tr>
<td>P25</td>
<td>SYIGLGSAL</td>
<td>1.1</td>
<td>159</td>
<td>Tc00.1047053508799.290</td>
<td>Hypothetical protein, conserved</td>
<td>L. infantum</td>
</tr>
<tr>
<td>P26</td>
<td>SYIGLGSAL</td>
<td>1.4</td>
<td>23</td>
<td>Tc00.1047053509431.10</td>
<td>Hypothetical protein, conserved</td>
<td></td>
</tr>
<tr>
<td>P27</td>
<td>SYIGLGSAL</td>
<td>1.4</td>
<td>94</td>
<td>Tc00.1047053509713.10</td>
<td>Hypothetical protein, conserved</td>
<td></td>
</tr>
<tr>
<td>P28</td>
<td>SYIGLGSAL</td>
<td>1.6</td>
<td>159</td>
<td>Tc00.1047053510069.20</td>
<td>Hypothetical protein, conserved</td>
<td></td>
</tr>
<tr>
<td>P29</td>
<td>SYIGLGSAL</td>
<td>1.6</td>
<td>148</td>
<td>Tc00.1047053508799.280</td>
<td>Hypothetical protein, conserved</td>
<td></td>
</tr>
<tr>
<td>P30</td>
<td>SYIGLGSAL</td>
<td>1.9</td>
<td>159</td>
<td>Tc00.1047053508799.270</td>
<td>Hypothetical protein, conserved</td>
<td></td>
</tr>
</tbody>
</table>
Tc00.1047053508799.270 and Tc00.1047053510069.20 both contained 3 immunogenic epitopes (P24, P25, and P26). All of the epitopes identified were novel, and none of the proteins they are derived from had been described as antigens before.

Evaluation of the Immunotherapeutic Potential of the T-Cell Epitopes

To further evaluate the potential of these novel epitopes, we tested them as a therapeutic vaccine [35, 36]. Mice were infected with *T. cruzi* and immunized with 2 doses of a therapeutic vaccine formulated as a pool of all 10 immunogenic peptides (P6, P8, P14, P15, P16, P17, P23, P24, P25, and P26) with MPLA as adjuvant, 7 and 14 days after infection. We monitored the progression of the infection by measuring parasitemia and survival for 50 days after infection, as well as cardiac tissue inflammation and parasite burden after 50 days. While control mice developed a high parasitemia (Figure 4A), and most died during the acute phase of the infection (Figure 4B), mice treated with the peptide mixture had a significantly lower parasitemia, and 85% survived the acute phase of the infection. Analysis of cardiac tissue damage 50 days after infection further indicated that vaccine recipients had a significantly lower density of inflammatory cells (Figure 4C), together with a 100-fold reduction in the *T. cruzi* burden (Figure 4D), compared with untreated mice. Thus, the identified epitopes administered as a therapeutic vaccine were able to control the ongoing *T. cruzi* infection, confirming their potential as a novel vaccine candidate.

DISCUSSION

The activation of CD8+ T cells has been found to be central to the control of *T. cruzi* infection [35, 41]. Thus, the identification of parasite epitopes presented in the context of class I MHC and able to induce the production of cytokines such as IFN-γ by T cells is a critical step for vaccine development against this parasite. We reported here the first unbiased screening of the complete *T. cruzi* proteome (>12 000 sequences) with a combination of immunoinformatics tools for T-cell epitope prediction. The immunogenicity of selected epitopes was confirmed by

---

**Table 1 continued.**

<table>
<thead>
<tr>
<th>ID</th>
<th>Epitope</th>
<th>Rank</th>
<th>Position</th>
<th>Gene ID</th>
<th>Function</th>
<th>Conservation</th>
</tr>
</thead>
<tbody>
<tr>
<td>P6</td>
<td>SYGCICSSL</td>
<td>2.5</td>
<td>31</td>
<td>Tc00.1047053508799.290</td>
<td>Hypothetical protein, conserved</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P8</td>
<td></td>
<td>4.5</td>
<td>31</td>
<td>Tc00.1047053508799.270</td>
<td>Hypothetical protein, conserved</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P14</td>
<td></td>
<td>4.7</td>
<td>31</td>
<td>Tc00.1047053510069.20</td>
<td>Hypothetical protein, conserved</td>
<td>...</td>
</tr>
</tbody>
</table>

P1 to P14 correspond to epitopes for H-2D<sup>d</sup>, and P15 to P26 correspond to epitopes for H-2K<sup>d</sup>. Abbreviations: ID, identification; *L. donovani*, Leishmania donovani; *L. infantum*, Leishmania infantum; *L. major*, Leishmania major; *T. brucei*, Trypanosoma brucei; *T. vivax*, Trypanosoma vivax.

*a* The average rank of prediction for the epitope calculated for each protein (proportion of T-cell epitope prediction tools that predicted the epitope).

*b* First amino acid position of the epitope in the protein.

*c* Gene ID of the protein containing the predicted epitope.

*d* Species in which the epitope sequence is conserved (>77% identity).

---

**Figure 3.** Interferon γ (IFN-γ) production by spleen cells obtained from infected mice and stimulated in vitro. Mice were infected with *Trypanosoma cruzi* for 1 month, their spleen cells were removed, and IFN-γ levels were measured by enzyme-linked immunosorbent assay in the supernatant of cultured spleen cells stimulated with the individual peptides. All measurements were done in triplicates. We used unstimulated cells as negative control and cells stimulated with concanavalin A (ConA) as positive control.
measuring epitope-specific IFN-γ production in vitro, after the stimulation of spleen cells from *T. cruzi*-infected mice with each predicted peptide, as well as by evaluating the immunotherapeutic potential of a mixture of these peptides.

We used several bioinformatics tools predicting epitopes presented in the context of MHC class I H-2D<sup>d</sup> and H-2K<sup>d</sup> molecules from BALB/c mice, to generate consensus predictions of greater accuracy [18]. Overall, there was a greater consensus in the predictions for H-2K<sup>d</sup>, compared with H-2D<sup>d</sup>, with the large majority of epitopes predicted by 7 or 8 different algorithms for H-2K<sup>d</sup>, compared with only 5–6 for H-2D<sup>d</sup>. This suggested a better performance of the algorithms for H-2K<sup>d</sup>, which was confirmed upon validation of the predicted epitopes. Indeed, 53% (7 of 12) of the H-2K<sup>d</sup> epitopes were validated, compared with only 21% (3 of 14) for H-2D<sup>d</sup>. This appears similar to previous studies, suggesting that predictions for H-2K<sup>d</sup> are more reliable than that for H-2D<sup>d</sup> [19, 20]. Nonetheless, this result may have been biased by the removal from our analysis of 22 H-2D<sup>d</sup> predicted epitopes derived from proteins belonging to the trans-sialidase family of surface proteins. Indeed, it is one of the best-characterized antigen families of the parasite [26] and includes immunodominant epitopes [42–44].

From our immunoinformatics analysis, we selected 26 peptides predicted with a good consensus by several algorithms to test their immunogenicity and validate the predictions. We measured IFN-γ production by splenocytes from *T. cruzi*-infected mice after recall stimulation with the peptides, which occurs when a secondary immune response to these peptide sequences is triggered. Thus, a recall response indicates that the epitopes had previously been processed from their corresponding *T. cruzi* proteins to prime T cells following an infection. Overall, we found that 38% (10 of 26) of the tested epitopes induced a recall response in splenocytes from *T. cruzi*-infected mice. This clearly indicated that several epitopes predicted with a very high binding score by many different algorithms were actually not immunogenic and stresses the need for the
experimental validation of immunoinformatics predictions, which have a limited value alone [18, 21]. The performance is similar to that observed in other studies [17, 19, 20] and allows for a faster identification of epitopes, compared with more traditional approaches. Most importantly, it allows screening of large genomes, such as that of protozoan parasites [20]. The false predictions of epitopes may be due to the fact that most programs predict the interaction of peptides with the MHC molecule, not the processing of the peptides, which is the previous step of the antigen-processing pathway. Thus, taking the proteasome cleavage of proteins into account, as well as TAP transport of the peptides to the endoplasmic reticulum, may increase epitope prediction accuracy, and several independent or integrated algorithms can perform this task [45, 46].

Of the 10 epitopes confirmed as immunogenic, most are derived from hypothetical proteins, which may thus be considered as partially validated, confirming the quality of the genome annotation [47, 48]. Further studies should demonstrate the expression of these proteins and explore their functions in T. cruzi biology. The other antigens have putative functions, including 2 kinases, a serine carboxypeptidase, a drug resistance protein, and a vacuolar protein sorting-associated protein, confirming that nonsurface proteins can also be good antigens [20]. The conservation of several of these validated epitopes in other kinetoplastids is particularly interesting since it suggests that they could provide a strong immunity against a broad range of T. cruzi strains and possibly even against other species, such as Leishmania. Similarly, epitopes P24, P25, and P26 are present in several proteins, which are also presenting >1 epitope, suggesting that these may be highly immunogenic and good vaccine candidates.

Testing the immunotherapeutic activity of the identified epitopes clearly indicated that the mixture of peptides provided strong control of T. cruzi infection in treated mice. Indeed, these presented a marked reduction in parasitism, cardiac tissue damage, and parasite burden and a marked increase in survival. This effect was comparable to that obtained with some previous therapeutic vaccines tested against T. cruzi [36, 37], including a vaccine based on a combination of 2 antigens [35], despite the lack of optimization of vaccine formulation and administration. Thus, these peptides represent excellent candidates for further vaccine development, alone or in combination with other antigens, and this result clearly validate the potential of the immunoinformatics approach for the rapid development of vaccines from a large parasite genome.

Because we used a mixture of all 10 identified epitopes for immunization, we could not determine the contribution of the different epitopes for the control of the infection. Future studies of the individual protective potential of the peptides may allow identification of the most protective ones among these immunogenic epitopes and further reduce the number of peptides to include in a vaccine candidate. Strong protection against T. cruzi infection has previously been observed with trans-sialidase antigens containing an immunodominant CD8+ T-cell epitope [42, 43] and even with recombinant adenoviral and vaccinia virus expressing a single immunodominant epitope [49]. However, vaccines based on single or even a small number of epitopes may be of limited relevance because of the considerable genetic and antigenic variability of T. cruzi, and they may lead to strain-specific protective immunity [50]. Indeed, T. cruzi is divided into 6 genetic lineages (ie, discrete typing units [DTUs]), referred to as Tc1-TcVI, which represents an additional challenge to vaccine development. Thus, conserved epitopes from multiple antigens (as in a polyepitope vaccine) are more likely to provide a broad immune response against many parasite strains and DTUs. As mentioned above, the conservation of several of the epitopes identified in this study in other kinetoplastid species suggests a very good potential for the induction of a broad immunity by these epitopes.

In conclusion, we performed the first unbiased screening of the T. cruzi complete genome to predict T-cell epitopes and validated 10 novel T-cell epitopes as immunogenic. In addition, we demonstrated the strong immunotherapeutic potential of these epitopes to control an ongoing T. cruzi infection in mice. These findings illustrate the benefits of this immunoinformatics approach, in combination with the in vivo validation of the immunogenicity and protective potential of the peptides, for the rapid development of vaccine against pathogens with large genomes, such as T. cruzi. The identified peptides and the parasite proteins from which they are derived may be excellent candidates for the further development of a vaccine against T. cruzi.

Notes

Financial support. This work was supported by the Consejo Nacional de Ciencia y Tecnología (CB-2010-01-156513 to E. D.)

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References


46. Doytchinova IA, Flower DR. Class I T-cell epitope prediction: improvements using a combination of proteasome cleavage, TAP affinity, and MHC binding. Mol Immunol 2006; 43:2037–44.


