



## Article

# A Data-Driven Approach to Construct a Molecular Map of *Trypanosoma cruzi* to Identify Drugs and Vaccine Targets

Swarsat Kaushik Nath <sup>1,†</sup>, Preeti Pankajakshan <sup>1</sup>, Trapti Sharma <sup>1</sup>, Priya Kumari <sup>1</sup>, Sweety Shinde <sup>1</sup>, Nikita Garg <sup>1</sup>, Kartavya Mathur <sup>1</sup> , Nevidita Arambam <sup>1</sup>, Divyank Harjani <sup>1</sup>, Manpriya Raj <sup>1</sup>, Garwit Kwatra <sup>1</sup>, Sayantan Venkatesh <sup>1</sup>, Alakto Choudhury <sup>1</sup> , Saima Bano <sup>1</sup>, Prashansa Tayal <sup>1</sup>, Mahek Sharan <sup>1</sup> , Ruchika Arora <sup>1</sup>, Ulrich Strych <sup>2,3</sup> , Peter J. Hotez <sup>2,3,4</sup>, Maria Elena Bottazzi <sup>2,3,4</sup> and Kamal Rawal <sup>1,\*,†</sup>

<sup>1</sup> Centre for Computational Biology and Bioinformatics, Amity Institute of Biotechnology, Amity University, Noida 201303, Uttar Pradesh, India

<sup>2</sup> Texas Children's Hospital Center for Vaccine Development, Departments of Pediatrics and Molecular Virology and Microbiology, Baylor College of Medicine, Houston, TX 77030, USA

<sup>3</sup> National School of Tropical Medicine, Baylor College of Medicine, Houston, TX 77030, USA

<sup>4</sup> Department of Biology, Baylor University, Waco, TX 76798, USA

\* Correspondence: kamal.rawal@gmail.com

† These authors contributed equally to this work.

**Abstract:** Chagas disease (CD) is endemic in large parts of Central and South America, as well as in Texas and the southern regions of the United States. Successful parasites, such as the causative agent of CD, *Trypanosoma cruzi* have adapted to specific hosts during their phylogenesis. In this work, we have assembled an interactive network of the complex relations that occur between molecules within *T. cruzi*. An expert curation strategy was combined with a text-mining approach to screen 10,234 full-length research articles and over 200,000 abstracts relevant to *T. cruzi*. We obtained a scale-free network consisting of 1055 nodes and 874 edges, and composed of 838 proteins, 43 genes, 20 complexes, 9 RNAs, 36 simple molecules, 81 phenotypes, and 37 known pharmaceuticals. Further, we deployed an automated docking pipeline to conduct large-scale docking studies involving several thousand drugs and potential targets to identify network-based binding propensities. These experiments have revealed that the existing FDA-approved drugs benznidazole (Bz) and nifurtimox (Nf) show comparatively high binding energies to the *T. cruzi* network proteins (e.g., PIF1 helicase-like protein, trans-sialidase), when compared with control datasets consisting of proteins from other pathogens. We envisage this work to be of value to those interested in finding new vaccines for CD, as well as drugs against the *T. cruzi* parasite.

**Keywords:** system biology; vaccine targets; pathways; gene ontology; chagas disease; *Trypanosoma cruzi*; drug target



**Citation:** Nath, S.K.; Pankajakshan, P.; Sharma, T.; Kumari, P.; Shinde, S.; Garg, N.; Mathur, K.; Arambam, N.; Harjani, D.; Raj, M.; et al. A Data-Driven Approach to Construct a Molecular Map of *Trypanosoma cruzi* to Identify Drugs and Vaccine Targets. *Vaccines* **2023**, *11*, 267. <https://doi.org/10.3390/vaccines11020267>

Academic Editor: Juan C. De la Torre

Received: 6 September 2022

Revised: 10 January 2023

Accepted: 12 January 2023

Published: 26 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Background

Chagas disease is caused by the protist parasite *Trypanosoma cruzi*. It affects 6–7 million humans and a large number of animal species. The study of CD and *T. cruzi* is challenging, due to the complexity and unique characteristics of the parasite's genome. For instance, 50% of the *T. cruzi* genome is composed of repeated sequences, such as transposable elements, microsatellites, and simple tandem repeats. It also includes surface molecules encoding genes, such as trans-sialidases, mucins, gp63s, and a large novel family (>1300 copies) of the mucin-associated surface protein (MASP) [1–3].

*T. cruzi* presents with a complex life cycle comprising four morphological stages: epimastigotes (EP), metacyclic trypomastigotes (MT), cell-derived trypomastigotes (CDT), and amastigotes (AM). During its life cycle, the parasite changes in morphology, metabolism, and gene expression, as it passes from the epimastigote replicative stage in the insect to the metacyclic trypomastigote form, which infects humans. *T. cruzi* appears to have a

unique approach to evading both parasitocidal antibodies and T-cell recognition. It creates an antigenic variation by generating large families of genes encoding both surface and secreted proteins, many of which are expressed simultaneously. Previously, Nde et al., first elucidated the lamc1 sub-network interactome mobilised in the human coronary artery smooth muscle (HCASM) cells by *T. cruzi*. The authors presented the human extracellular matrix interactome network regulated by *T. cruzi* and its gp83 ligand that facilitates cellular infection [4]. These authors also reported that *T. cruzi* modulates the extracellular matrix (ECM) interactome to gain cell entry while evading the host immune system [4]. *T. cruzi* triggers gp83 receptors in the host cell via ERK1/2 to up-regulate LAMC1 which cross-talks to both *LGALS3* and *THBS1* to enhance cellular infection using selected parasite surface molecules, such as TcCRT and TC45 mucin [4]. Roberts et al. used proteomics and interaction data to generate constraint-based models, focusing on pathways of the central metabolism, such as glycolysis [5]. Researchers have also constructed genome-scale metabolic models of *T. cruzi* to investigate the variation in metabolic functions across the life cycle stages and to predict the stage-specific essential genes and reactions. These reactions are essential for the growth of *T. cruzi* and could potentially be drug targets, such as the cytosolic transketolase reaction, the mitochondrial alanine transport reaction, and the glycosomal glucokinase reaction [6].

Despite these efforts, there remains an urgent need to construct a molecular network of *T. cruzi* to understand the pathophysiology of trypanosome infection. Networks have played an important role in the understanding of complex diseases, such as obesity, atherogenesis, myocardial infarction, heart failure, Charcot–Marie–Tooth disease, and spinal muscular atrophy [7–13]. It has been proposed that pathogens that cause acute infections tend to target central hubs of the host’s cellular signaling networks, whereas pathogens causing chronic disease are more likely to target peripheral nodes [14]. Furthermore, the network pharmacology approach has been used to study “compound-protein/genes-disease” pathways, which can describe complexities among biological systems, drugs, and diseases. Thus, molecular networks provide significant avenues for the identification of drug targets [15–17], vaccine targets [18–20], and biomarkers [21–23].

Here, we present a molecular map of *T. cruzi* constructed from an extensive literature review using deep curation and text-mining techniques. We also conducted large-scale docking studies of benznidazole and nifurtimox against several thousand *T. cruzi* predicted protein structures using our new drug-dock approach to identify the important molecular targets of Bz and Nf in the *T. cruzi* molecular network. This work also identified other FDA-approved drug molecules, in the context of *T. cruzi*, and new strategies to identify drug targets using comparative approaches.

## 2. Methods

A hybrid approach combining deep curation and text-mining strategies was used to curate the relevant information related to the *T. cruzi* pathways and molecules (genes, proteins, drugs, etc.) from published scientific articles. The abstracts were retrieved from literature databases, such as PubMed and Google Scholar using keywords, such as “*T. cruzi*” AND pathways name, “*T. cruzi* AND molecules (Gene, protein, and drug)”. Different networks were constructed using CellDesigner 4.4.2 [24], a modeling tool that enables us to graphically represent interactions using a well-defined and consistent graphical notation in the Systems Biology Markup Language (SBML) [25,26]. The schematic diagram for the complete methodology is shown in Supplementary Figure S1. The standard notation scheme for the graphical representation is provided in Supplementary Figure S2.

### 2.1. Extraction of the Literature Information Related to the *T. cruzi* Pathways

We extracted the names of the *T. cruzi* pathways using the keywords “*T. cruzi*” AND “pathway” on Google Scholar and PubMed. A total of 70,600 hits and 1142 hits were found on Google Scholar and PubMed, respectively. We shortlisted 46 unique pathways of *T. cruzi*

from this total number of hits (Supplementary Table S1). All 46 pathways were classified, based on their function (Supplementary Table S2).

Further, we extracted the abstracts related to these 46 pathways from Google Scholar and PubMed using the keywords “*T. cruzi*” AND “<Name of pathway>” through deep curation strategies. Manual curation was performed using extracted abstracts and the text describing the gene/protein or any other molecules involved in the pathway was retrieved during the curation.

## 2.2. Retrieval of the *T. cruzi* Molecules and Their Function

We retrieved 19,607 genes of *T. cruzi* CL Brener (*Tc*-CLB) from the NCBI database [27] (Supplementary Table S3). A total of 1756 unique genes were identified after removing the duplicate genes (Supplementary Table S4). To find out the function of the genes, an extensive literature search was carried out on PubMed and Google Scholar using the key terms “*T. cruzi*” AND “gene name”. All curated molecular evidence is summarized in Supplementary Table S5. In addition, 86,990 abstracts (Google Scholar-85,600, PubMed-1390, as of February 2021) were collected using the literature mining approach (Supplementary File S1). The 1756 unique genes with their known molecular interactions were used to construct a comprehensive molecular map (Figure 1).



**Figure 1.** The network of the comprehensive molecular map which includes 1756 genes and 3109 proteins related to *T. cruzi*. The network is generated using CellDesigner v.4.4.2. Since the figure is very large in size, we are providing alternative links in the form of PNG on our website for the user to download and view (<https://tinyurl.com/supplementary-data>).

Next, a total of 19,757 *Tc*-CLB proteins were retrieved from the UniProt database [28] (Supplementary Table S6). Out of this list, 3109 unique proteins were filtered manually (Supplementary Table S7). An extensive literature search was carried out in PubMed and Google Scholar using the key terms “*T. cruzi*” AND “protein name”. All curated molecular evidence with its interactions is outlined in Supplementary Table S8. A total of 103,868 abstracts (Google Scholar-96,300, PubMed-7568), as of February 2021, were screened to obtain the molecular information related to *T. cruzi* (Supplementary File S2).

All unique proteins from the UniProt database, with their interaction information, were used for the construction of the molecular map.

### 2.3. Retrieval of the *T. cruzi* Drugs

For the construction of the *T. cruzi* drug network, we screened research articles manually, using the key terms “*T. cruzi*” AND “drug”. The 1693 abstracts with the terms “*T. cruzi*” AND “drug” were downloaded from Google Scholar and PubMed (Supplementary File S3). The molecular evidence describing the functions of the drugs were extracted and then used for the construction of the network using CellDesigner 4.4.2 [24]

#### 2.3.1. Networks Analysis

The network topological parameters were analyzed using a Cytoscape plugin called “Network Analyzer” [29]. It calculates several parameters, such as the number of nodes, number of edges, the average number of neighbors, network diameter, network radius, characteristic path lengths, clustering coefficient, network density, network heterogeneity, network centralization, and connected components.

#### 2.3.2. Gene Ontology (GO) Analysis

The unique *T. cruzi* genes retrieved from various sources were used for the gene ontology analysis. The gene ontology was carried out for four categories i.e., biological process, molecular function, cellular component, and metabolic pathways. The enrichment analysis of the gene ontology (GO) [30], KEGG [31], and MetaCyc pathways [32] was performed for the *T. cruzi* genes using TriTrypDB v44 [33,34]. In TriTrypDB, the data is available for a wide range of kinetoplastids. For this analysis, we selected the *T. cruzi* organism to filter out only the *T. cruzi* genes (see <https://tinyurl.com/TriTrypDB>). Further, the significantly enriched GO terms were clustered, summarized, and visualized using REVIGO [35]. The GO terms and pathways enrichment were considered statistically significant when the false discovery rate or FDR was less or equal to 0.05.

#### 2.3.3. Homology Modelling of the *T. cruzi* Proteins

We retrieved the FASTA sequences of 19,607 *Tc*-CLB proteins from NCBI. Since only 127 crystal structures were available in the PDB for *T. cruzi*, we decided to construct three-dimensional (3-D) structures with the help of homology modelling. As a first step, we were able to find 4905 *Tc*-CLB proteins showing a similarity ( $\geq 35\%$ ) with the sequences available in the PDB. Next, we used the MODELLER tool [36], for the homology modeling of the proteins [37].

#### 2.3.4. Molecular Docking

We developed a new algorithm named drug-dock for the large-scale automated docking using Auto dock Vina [38]. This system is based upon Python, JavaScript, and Open Babel. The system also has a module for the prediction of protein structures using homology modeling [39] in case the target structure is not present in the PDB. We obtained the structural information on the genes/proteins implicated in *Tc*-CLB from the PDB. The structure information for Bz, Nf, orlistat, and aspirin, as well as other drugs was obtained from DrugBank [40] and their side effects were retrieved from the SIDER database. The PDB format of the protein structure was converted to PDBQT format before commencing the docking procedure. For the prediction of the binding site, we have incorporated the P2Rank tool [41] in the pipeline. The system performs the generation of a configuration file for each protein, converts the structure into PDBQT format, changes the ligand from SDF to PDBQT format using Open Babel [42], and performs molecular docking using Auto dock Vina. The system also converts the ligand's 2d-SDF to 3d-SDF before converting it into PDBQT format. The obtained config files for the proteins and ligands in PDBQT format are provided as input to Auto Dock Vina. Auto Dock Vina generates a log file and an output



file. The log file contains the binding energies of proteins and drugs, whereas the output file contains the 3D structures of the 10 highest-scoring poses (orientations) of each drug attached to the protein. During the process of parsing, the final result is generated in the form of a matrix.

### 3. Results

#### 3.1. General Features of the Comprehensive Molecular Map of *T. cruzi*

A total of 86,990 abstracts related to *Trypanosoma* genes (using keywords “*T. cruzi* AND genes”) were collected from Google Scholar (85,600 abstracts) and PubMed (1390 abstracts) (published till February 2021) (Supplementary File S1). Out of 19,607 genes available from NCBI (*Tc*-CLB dataset), only 1756 gene names were found to be unique in nature (Supplementary Tables S3 and S4). Next, we extracted the information about the function, as well as the interaction information for each gene (Supplementary Table S5).

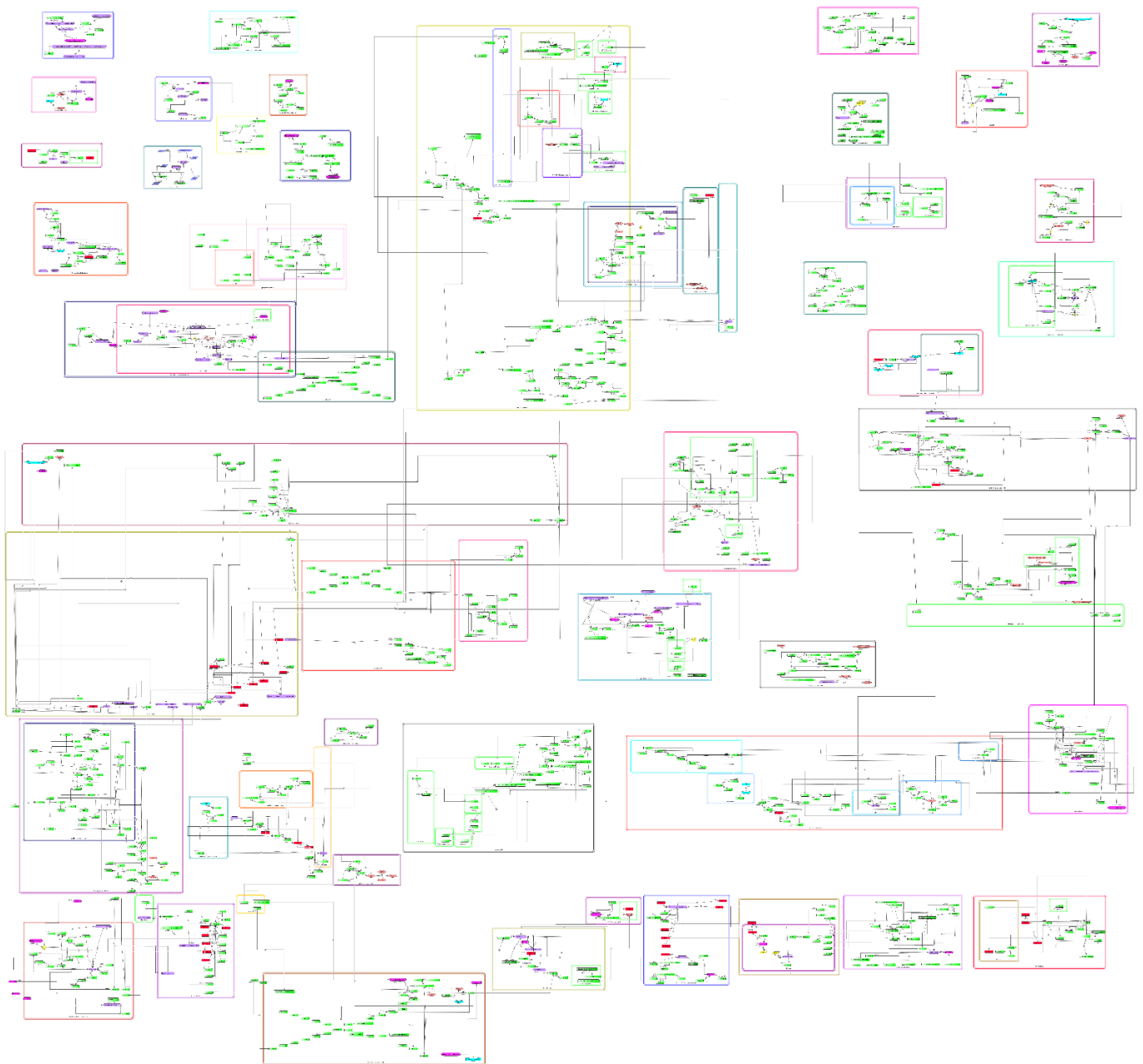
Similarly, we retrieved 103,868 abstracts (Google Scholar and PubMed) using the keyword “*Trypanosoma cruzi* AND Protein” (February 2021) (Supplementary File S2). A list of the *Tc*-CLB proteins was collected from the UniProt database (19,757 proteins after removing the redundant hits [28]) (Supplementary Table S6). Further, we filtered the multi-copy proteins and hypothetical hits from the list to obtain 3109 unique protein names (Supplementary Table S7). Next, we extracted the interaction information for each of these proteins (Supplementary Table S8).

#### 3.2. Features of the Comprehensive Map

The comprehensive molecular map comprises 2415 nodes and 1608 edges. The nodes include 190 genes, 1188 proteins, eight antisense RNAs, 201 complexes, 12 degraded, 48 drugs, 13 ions, 534 phenotypes, 22 RNAs, 42 unknown nodes, and 157 simple molecules (Supplementary Table S9). The edges were categorized into 1458 state transitions, five positive influences, 18 triggers, nine modulations, nine transport-related, seven translation-related, six transcription-related, one physical stimulation, and 92 others (Supplementary Table S10). The comprehensive molecular map is shown in Figure 1. It contains the molecules that are known to participate in important pathways, such as the ERK1/2 mitogen-activated kinase pathway, the glycogen synthase kinase 3 pathway, mTOR pathways, etc. We also found several housekeeping genes on the map, such as *bap1* (BRCA1 associated protein), *coq4* (coenzyme Q4), *dpy30* (Dpy-30 histone methyltransferase complex regulatory subunit), *mpc2* (mitochondrial pyruvate carrier 2), and *ndufv2* (NADH: ubiquinone oxidoreductase core subunit V2). The drugs reported as inhibitors of the *T. cruzi* growth were also included. These include Bz, tipifarnib (R115777), Nf, and leptomycin B.

#### 3.3. Pathways of *T. cruzi*

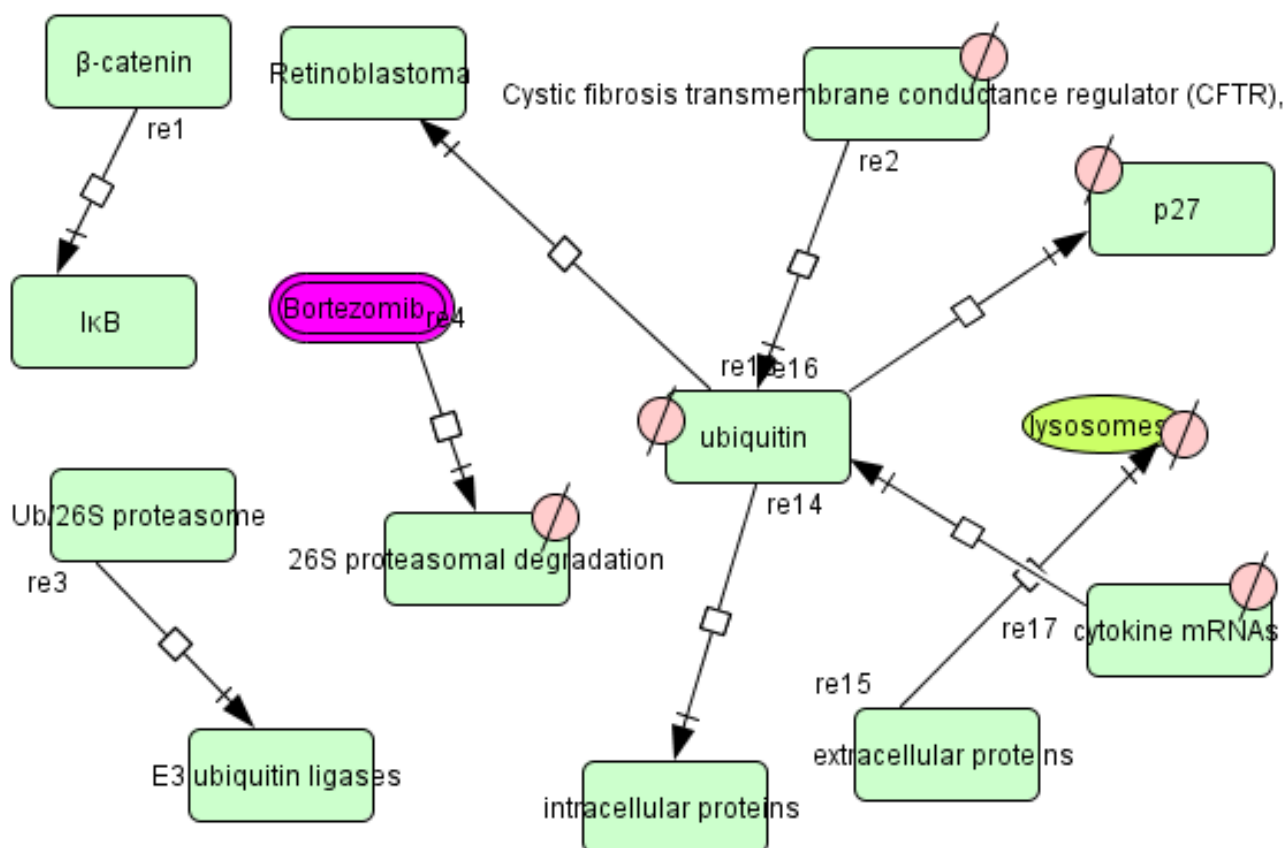
Following the construction of the comprehensive map (see Section 3.2), we started collecting information on the pathways reported in the literature. Using text mining and deep curation, we found 46 unique pathway names in the literature (Supplementary Table S11). We extracted the molecules reported to be associated with these pathways. Further, we also extracted the molecular interaction information among the molecules involved in each pathway. A manually curated resource (database) containing papers or abstracts (with highlighted text to describe the role of the specific molecules) was assembled (Supplementary Table S1). A combined Systems Biology Markup Language (SBML) file was prepared for each pathway (Figure 2). In total, the pathways contained 1055 nodes that connect 43 genes, 838 proteins, one molecule involved in protein degradation, 37 drugs, 10 ions, 81 phenotypes, nine RNAs, and 36 simple molecules (Supplementary Table S12). These nodes are connected by 874 edges (Supplementary Table S13). We categorized these pathways into 17 major classes, including metabolic pathways, signaling pathways, degradative pathways, inflammatory pathways, etc. (Supplementary Table S2). As an example, we shall provide details of one of the constructed pathways (ubiquitin-proteasome pathway) in the next section (Figure 3).



**Figure 2.** The network of molecules (gene, protein, drugs) involved in the 46 pathways. The network is generated in the CellDesigner software v.4.4.2. Since the figure is very large in size, we are providing alternative links in the form of PNG on our website for the user to download and view (<https://tinyurl.com/supplementary-data>).

Ubiquitination is an important process in eukaryotes and the ubiquitin-proteasome pathway enzymes are an important component of the protein degradation machinery. The process of degradation starts with the activation of ubiquitin (Ub), through the Ub-activating enzyme, E1, followed by the transfer of the activated Ub protein to Ub-conjugating enzymes (E2s) by the transacylation reaction. E2 transfers the Ub to the target protein substrates with the aid of the substrate-specific Ub ligases (E3s). The conjugation of a single Ub moiety is termed mono-ubiquitination and the subsequent conjugation of the Ub moieties leads to the formation of a polyubiquitin chain. This cascade of events leads the target substrate protein to the 26S proteasome for elimination [43]. To check the potential targets, we used molecular docking techniques to check the interactions of Bz and Nf (DB11820 and DB11989) against the molecules present in the ubiquitin-

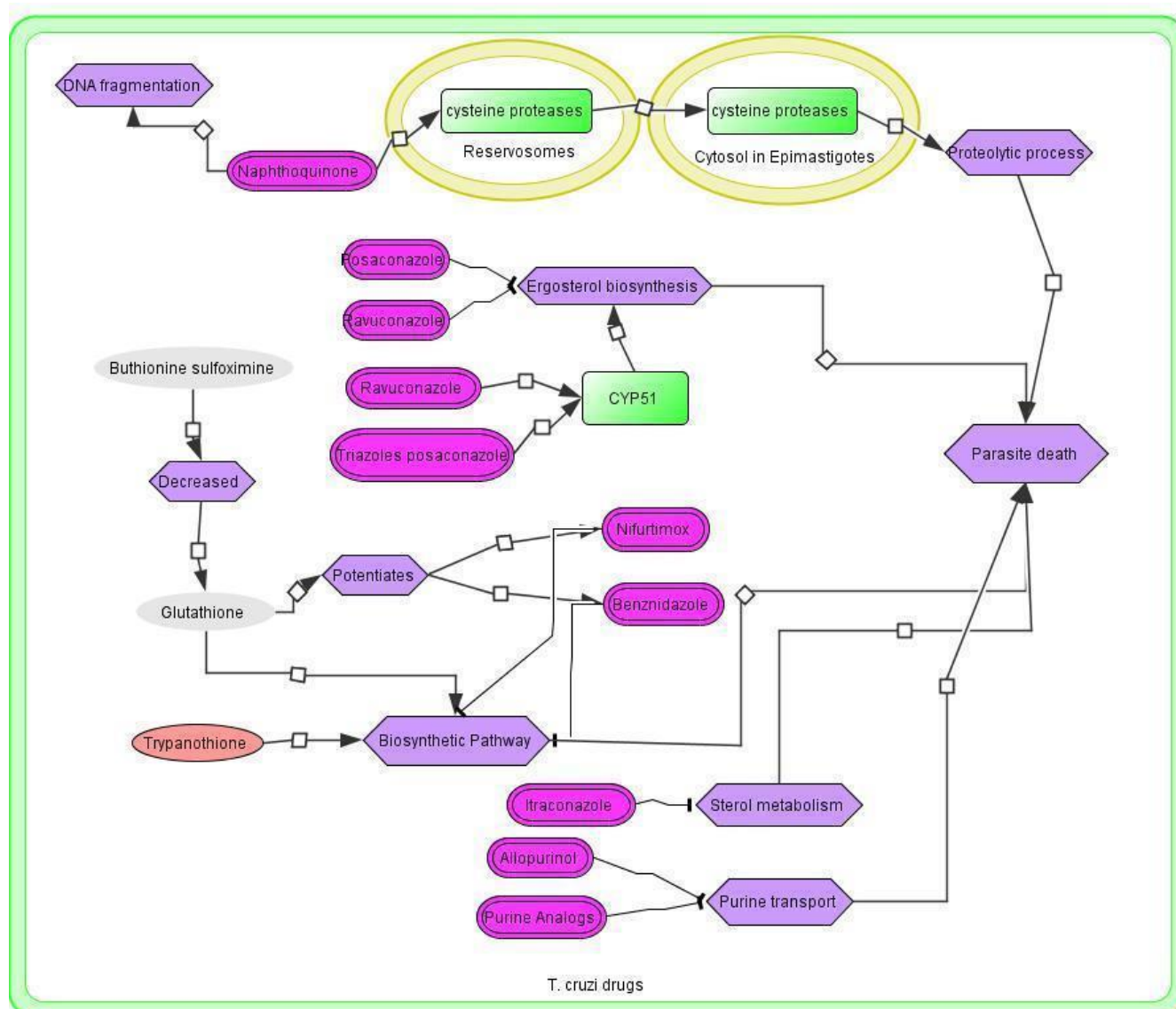
proteasome pathway. The molecules include glutathione peroxidase, ubiquitin/ribosomal protein S27a, ubiquitin-protein ligase, ubiquitin carboxyl-terminal hydrolase, 26S protease regulatory subunit, ubiquitin hydrolase, and ubiquitin-conjugating enzyme [44–46] (Supplementary Table S14).



**Figure 3.** The network of the ubiquitin-proteasome pathway developed using the CellDesigner software (version—4.4.2).

### 3.4. *T. cruzi* Drugs and Network

An extensive literature search was carried out using PubMed and Google Scholar with the key terms “*T. cruzi* AND drug”. A total of 68,293 abstracts (Google Scholar 66,600 and PubMed 1693 hits published up to February 2021) were collected (Supplementary File S3). Each report was curated to obtain the molecular information related to the drug. The relevant line(s) or paragraph(s) was highlighted and used as evidence for building the drug network (Supplementary Table S15 and Figure 4). The drug network comprises 25 nodes including three proteins, 10 drugs, one simple molecule, five ions, nine phenotypes, and two unknown molecules that are connected via 26 edges (Supplementary Table S16). The edges represent interactions between each reactant or node. In this network, the interactions between reactants can be categorized into state transitions (19), and negative influences (Supplementary Table S17). In addition, we found 41 drugs for *T. cruzi* which are in different stages of development. For example, allopurinol, sulfasalazine, and thioridazine (TZD) are undergoing testing in the lab (experimental stages) [47–50] whereas posaconazole, ravuconazole, Nf, and Bz are in different phases of clinical testing [51–54] (Supplementary Table S18).



**Figure 4.** The drug network of *T. cruzi* developed using CellDesigner v.4.4.2.

### 3.5. Application of the *T. cruzi* Drug Network

Treatment of CD is still limited to only two drugs, Bz and Nf [55,56]. Both are orally administered and can cause severe side effects, as well as long-term toxicity. Apart from these two drugs, naphthoquinone derivatives play an important role in DNA fragmentation, as well as in the release of cysteine proteases from reserosomes to the cytosol. This proteolytic process leads to parasite death [57] (Figure 2). Moreover, an antifungal—ravuconazole is a promising drug that is in clinical trials against CD, and used in combined therapy with Bz [58]. CYP51 (Sterol 14 $\alpha$ -demethylase cytochrome P450) is an important enzyme with the trypanocidal activity responsible for ergosterol's biosynthesis, that was identified in 1990. CYP51 inhibits the sterol synthesis, which is lethal to the parasite. Ravuconazole and posaconazole act through the coordination of nitrogen with heme iron into the binding cavity of CYP51 [59]. Studies on animal models found that posaconazole could be used for the treatment of acute and chronic CD [11]. The combination of Bz and itraconazole was shown to decrease the typical lesions (myocardial inflammation and fibrosis) associated with chronic CD and eliminate the parasites from the blood [12].

To check the interactions of various candidate drugs at the network level, we decided to use a drug dock algorithm. For instance, we docked Bz and Nf against the available 127 crystal structures of the *T. cruzi* proteins (Supplementary Table S19). We found that Nf



is predicted to have the highest binding affinity with *T. cruzi* type B ribose 5-phosphate isomerase (TcRpiB) (−8.0 kcal/mol). During the literature curation, we found that the RpiB enzymes are present in the parasite whereas their homologs (RpiA) are absent. Further, TcRpiB turned out to be the only enzyme of the *T. cruzi* PPP (pentose phosphate pathway) which does not have a counterpart in higher eukaryotes. To check the potential impact of Bz and Nf on the PPP, we conducted a docking study against the members of the PPPs. As a comparison, we also studied the interactions of the controls (aspirin and orlistat) against the same targets. We found that the binding energy distributions are higher in the Bz and Nf study group, when compared with aspirin and orlistat. This holds, not only in the PPP of *T. cruzi*, but also in other pathways, suggesting differential binding preferences of Bz/Nf (Supplementary Table S20).

To study the potential side effects of Bz and Nf on humans, we conducted a large-scale docking study on the human proteome. For this, we collected 24,391 human proteins from the AlphaFold database [60] and performed docking with Bz. Due to technical issues, we could only dock 19,523 proteins with Bz. Interestingly, we identified biotin-protein ligase (Supplementary Table S21) as one of the top-ranking interactors of Bz (−9.1 kcal/mol of binding energy). Biotin is a water-soluble vitamin that belongs to the vitamin B complex and is an essential nutrient of all living organisms from bacteria to man [61]. In eukaryotic cells, biotin functions as a prosthetic group of enzymes, collectively known as biotin-dependent carboxylases that catalyze the key reactions in gluconeogenesis, fatty acid synthesis, and amino acid catabolism [61]. Biotin protein ligase (BPL) is required for the covalent attachment of biotin to biotin-dependent enzymes [62]. The clinical features of biotin deficiency include rashes, brittle hair, lethargy, hallucination, sleep disturbances, myalgia, and paraesthesia. The human biotin protein ligase (UniProt ID: P50747) is associated with glutamine deficiency and congenital phenotype [63] (OMIM database). Glutamine also contributes to the normal intestinal barrier function and can become deficient in some intestinal diseases, including Crohn's disease, diarrheal illness, and short gut syndrome [64]. According to Viotti et al., the side effects of Bz vary from person to person [65]. The major side effects of Bz include insomnia, fatigue, anorexia, headache, furred tongue, gastrointestinal disturbances, skin rash, pruritus, erythema multiforme, and toxic epidermolysis [66,67]. We see a significant overlap between the clinical features of biotin deficiency and the side effects of Bz. There is a strong possibility that some of Bz's side effects (i.e., gastrointestinal disturbance, sleep disturbances, etc.) can be linked to the off-target binding with BPL.

We further identified prostaglandin F synthase (PGF) (PDB: 4GIE) from *T. cruzi* bound to NADP, as another top-ranking target (−8.4 kcal/mol with benznidazole). The function of PGF is to catalyze the reduction of aldehydes and ketones to their corresponding alcohols. In humans, these reactions take place mostly in the lungs and the liver [68]. It is pertinent to note that PGF is involved in essential lipid-metabolism pathways in protists. Whereas in humans, PGF (Uniprot P42330) can interconvert active androgens, oestrogens, and progestins with their cognate inactive metabolites [69]. In humans, prostaglandins (PGs) E<sub>2</sub> and PGF<sub>2</sub>α are produced in the endometrium and are important for menstruation and fertility [70]. Prostaglandin F<sub>2</sub>α synthase or old yellow enzyme (OYE), another NAD(P)H flavin oxidoreductase, similar to mitochondrial NADH-dependent type-I nitroreductase (NTR I), has been implicated in the activation pathway of other trypanocidal drugs, such as Nf but not Bz [71]. Different studies have shown that OYE was found to be downregulated in resistant parasites [72,73]. Murta et al. (2006) found that this protein was downregulated in resistant parasites, due to the deletion of three copies of the gene [73]. Likewise, by proteome analysis, it was found that OYE was under-expressed in resistant parasites [72]. Thus, PGF can be considered an important target for Bz. One of the reported side effects of Bz is hepatitis, which could be linked to the unwanted interaction of Bz with human PGF.

Next, we docked 1516 FDA-approved drugs with PGF, intending to find additional candidate drug molecules against CD (Supplementary Table S22). We found that the top three drug molecules were rifabutin, lurbinedectin, and amphotericin B. Further, we predicted 4905 structures of *T. cruzi* using homology modelling, since only 120 experimentally

known structures were available in the PDB. Using this dataset, we found that the binding energy scores of Bz and Nf were significantly different from the binding energy scores of unrelated controls, aspirin, and orlistat (Welch's T-test) (Supplementary Table S23). The top five ranking protein targets for Bz were found to be: hypothetical protein (XP\_813710) (showing a similarity with the putative dynamin family), spermidine synthase (XP\_816871), apurinic/aprimidinic endonuclease (XP\_816327), trans-sialidase (XP\_802286), and peroxisome biogenesis factor 1 (XP\_809676). Next, we used these top five targets to screen FDA-approved drugs to find potential drug candidates.

In a similar approach, the top five protein targets identified using Nf, were PIF1 helicase-like protein (XP\_820585), hypothetical protein (XP\_818590) (showing 99.77% similarity with RNA editing 3' terminal uridylyl transferase 1 (*Trypanosoma cruzi*) (GenBank ID: PWV06812.1)), kinesin (XP\_817032), ras-related protein Rab21 (XP\_804862), and serine/threonine-protein kinase (XP\_804576).

### 3.6. Therapeutic Implications of the *T. cruzi* Network

To check whether Bz/Nf produce their clinical effects, due to the preferential binding to several molecules listed in the *T. cruzi* network (N), rather than the proteomes of other pathogens, we used a dataset of 4905 *T. cruzi* protein structures (P). In addition, we created datasets of randomly selected protein structures from different pathogens (labelled as P1, P2, ... Pn) as controls. We used proteins from *Plasmodium falciparum* (P1), *Mycobacterium tuberculosis* (P2), *Leishmania donovani* (P3), and *Salmonella typhi* (P4). We found that the binding energy distributions of Bz/Nf are significantly higher, when compared with the protein datasets selected from other pathogens (Welch's T-test) (Supplementary Table S24). These results indicate that parasite-specific drugs (Bz/Nf) have specific binding affinities towards the parasite-specific proteins (i.e., *T. cruzi*). The reason could be attributed to the presence of specific types of residues and patterns/motifs in the binding sites of several parasite proteins.

Our research group has a long-standing interest in Tc24 (flagellar calcium-binding protein of 24 kDa). This protein has been proposed as a candidate for an immunotherapeutic vaccine, as well as a drug target [74]. Following the retrieval of the protein structure of Tc24 (accession ID: Q1L1I2\_TRYCR) from AlphaFold [75], we conducted a large-scale docking study against Tc24 using drug molecules derived from the Zinc database [76] and a set of FDA-approved drug molecules. We found the drugs dutasteride, sirolimus, and candicidin to show the highest binding affinities against Tc24 (Supplementary Table S25).

Additionally, we conducted the molecular docking of Bz and Nf against Tc24 using Autodock Vina [38]. The PDB file for Tc24 was obtained from AlphaFold while the ligands (drugs) were obtained from DrugBank. The input PDBQT files for the protein and ligands were obtained from Autodock Vina [38]. The binding affinity for the top-scoring pose of Bz was observed as  $-6.9$  kcal/mol while for Nf, it was observed as  $-6.7$  kcal/mol. We also observed the distance of other modes from the best modes, in terms of the root mean square deviation (RMSD) (Supplementary Table S26 and Supplementary Figure S3).

### 3.7. Quantitative Analysis of the Networks

The topological analysis of all constructed networks was performed using a network analyzer tool [29] (Supplementary Table S27). The details were as follows:

(A) Clustering coefficient—If node A in a network is connected to node B, and B is connected to node C, A likely has a direct connection to C as well. The clustering coefficient can be used to quantify this phenomenon. It determines the average local neighborhood in a network [77], which varies frequently across the network [78]. If the clustering coefficient is near 0, the majority of nodes in the network have less than two neighbors, implying a tree-like topology [79]. The clustering coefficient value on our comprehensive map and modules is close to 0. The average clustering coefficient  $C(k)$ , which indicates the metabolic network's modularity [80], is another significant measure of the network's structure. Using

Network Analyzer [81], we discovered that our network's average clustering coefficient was 0, indicating the presence of a tree-like structure;

(B) Network diameter is the maximum distance between two nodes. If the network is disconnected, the average of the maximum distances between the linked components is used to calculate the diameter. The diameters for the *T. cruzi* pathway network, molecular map network, and drug network were computed to be 28, 22, and 18 units, respectively;

(C) The characteristic path length is measured as an average number of edges dissociating any two nodes in the network. The pathway network has 1929 nodes with a path length of 11.73 units, the comprehensive map has 4016 nodes with a path length of 8.33 units, and the drug network has 51 nodes with a path length of 6.91 units;

(D) The average number of neighbors indicates the average connectivity of a node in the network. We observed 2.18 for our pathway network, 2.02 for the comprehensive molecular network, and 2.03 for the drug network;

(E) The network density determines compactness, which can be simply defined as the ratio of observed edges to the number of possible edges for the given network. The value ranges from 0 to 1, the closer the value, the denser and more cohesive the nodes in the network. We have computed the network density for our network and the average network density for all three networks is 0.027.

### 3.8. Gene Ontology Analysis of the *T. cruzi* Genes

To understand the biological functions of the *T. cruzi* genes used in the network construction, we performed a gene ontology analysis using TriTrypDB [33]. A total of 803 genes were recognized as belonging to the Trypanosomatidae class (Supplementary Table S28). We included only 72 genes that belong to two *T. cruzi* strains (CL Brener Esmeraldo-like (40), and Non-Esmeraldo-like (32)) (Supplementary Table S29). Based on function, the genes were enriched in a cellular component, molecular function, biological processes, and metabolic pathways. The cluster representation was performed using the REVIGO tool [35], which allows the clustering of semantically similar gene ontology terms, and labels each cluster with a single representative gene ontology term. The enriched metabolic pathway was considered statistically significant when the false discovery rate (FDR) was less than or equal to 0.05. The details of the enriched genes are provided in Supplementary Tables S30 and S31.

The *Tc*-CLB Esmeraldo-like genes were enriched in several molecular functions, such as cystathionine beta-synthase activity (GO:0004122), catalytic activity (GO:0003824), and metal ion binding (GO:0046872), different cellular components, such as the intracellular membrane-bounded organelle (GO:0043231), membrane-bounded organelle (GO:0043227), and mitochondrial inner membrane (GO:0005743), and the biological processes, such as alpha-amino acid biosynthesis (GO:1901607) or cellular amino acid biosynthesis (GO:0008652) (Supplementary Figures S4 and S5).

Similarly, the GO of *Tc*-CLB Non-Esmeraldo-like showed that the highest number of enriched genes were involved in different cellular components, such as the kinetochore (GO:0000776), chromosomal region (GO:0098687), and the site of double-strand breaks (GO:0035861), molecular functions, such as cystathionine beta-synthase activity (GO:0004122), catalytic activity (GO:0003824), and hydro-lyase activity (GO:0016836), and in biological processes, such as sulphur compound biosynthetic process (GO:0044272), cellular process (GO:0009987), cysteine metabolic process (GO:0006534), and cysteine biosynthetic process (GO:0019344), (Supplementary Figures S4 and S5). Many pathways, such as UDP-alpha-D-galactose biosynthesis, UDP-alpha-D-glucofuranose biosynthesis, D-galactose detoxification, and puromycin biosynthesis, were enriched in *Tc*-CLB Esmeraldo-like and Non-Esmeraldo-like (Supplementary Figure S6).

## 4. Discussion

This work combines the data from heterogeneous databases, including the literature, structure, and expression to construct a comprehensive map of *T. cruzi* molecules. It

attempts to explain the interactions of the drug molecules in the context of networks. An effective drug molecule is expected to target the key molecules of the pathogen, as well as disrupt the key sections of its molecular network. Historically, molecular docking at a large scale has been deployed sparingly. Gao et al. used ~1100 targets [82]; Hui-Fang et al., used 1714 targets [10,83,84].

The major principle in drug discovery is to design maximally selective ligands to act on individual drug targets. However, several drugs act via the modulation of multiple proteins rather than single targets, suggesting the role of network-based therapeutics [85]. Here we performed the docking of trypanocidal drugs (Bz and Nf), control drugs (orlistat and aspirin), and 1500 FDA-approved drugs with *T. cruzi* network proteins (4905), as well as with the whole human proteome (19,523). We used our new drug repurposing pipeline to conduct large-scale docking [86]. Based on our prediction, we propose that the trypanocidal drugs show a preferential binding with *T. cruzi* proteins, as compared to the control drugs. It has been observed that benznidazole, not only binds to known targets, but also to several other targets of human proteins, which could explain some of the side effects of Bz. The overlap of the Bz side effects and biotin deficiency symptoms suggests a possibility of attenuating the side effects by the biotin administration [87,88]. Here, it is important to mention that another organic compound, “benzimidazole”, was studied by Woolley (1944). Benzimidazole derivatives have been reported to contain the trypanocidal activities [89]. Woolley (1944) reported that the similarity of the symptoms observed in animals receiving benzimidazole to those seen in biotin deficiency, suggested that the action of benzimidazole might be related to its structural similarity to biotin [90]. Both Bz and benzimidazole contain a common chemical nitrogen-containing ring.

We also observed that therapeutically unrelated drugs, i.e., orlistat and aspirin, displayed different binding patterns to the *T. cruzi* network proteins. Further studies are needed on other pathogens, which could explain the therapeutic effect of drugs at the network level. For example, similar studies could be conducted to study the binding of chloroquine against the proteome of malarial parasites. The other interesting potential application of our study is to compare the distribution of the binding energies of Bz/Nf in other related strains and species of *Trypanosoma*.

In our docking study, Nf was predicted to bind with TcRpiB, with a binding affinity of  $-8.0$  kcal/mol. TcRpiB plays an important role in the pentose phosphate pathway (PPP). It is responsible for the production of nucleotide precursors and NADPH, which provide protection to trypanosomatids during oxidative stress [91]. A study performed by Loureiro et al. showed that RpiB silencing in *Trypanosoma brucei* reduced the in vitro growth of the parasites. Furthermore, RpiB silencing in the infected mice, exhibited lower parasitaemia and prolonged survival compared to control mice [92]. The absence of the RpiB enzyme in humans and its pivotal role in PPP makes it a potential chemotherapeutic target for trypanocidal drugs. RpiB is reported to be conserved among different *Trypanosoma* species [92]. Larkin et al. conducted a protein sequence alignment using ClustalW [93] and found a 67% identity for *T. brucei* RpiB versus TcRpiB, and both proteins show no similarity with human ribose 5-phosphate isomerase A. Faria et al. compared the RpiB sequences of *L. infantum* (LiRPIB), *L. major* (LmRPIB), *T. brucei* (TbRPIB), and *T. cruzi* (TcRPIB). They found that LiRPIB displays a 93% sequence identity with LmRPIB and around 50% with RpiB from trypanosomes [94].

Biological systems are robust in the way that they restore the perturbations caused by drug treatments. One of the key avenues for a successful therapeutic drug or vaccine development is to overcome the biological robustness, maintained through positive or negative feedback loops of the drug/vaccine target proteins. The expression of several genes is believed to be altered during the disease phase. The genome-wide transcriptional profiling should enable us to specifically monitor the expression changes of the drug targets induced by their inhibitors or activators. The next version of this network will include the integrated genome-wide expression datasets, to study the perturbation induced by the therapeutic agent (drugs, vaccines, etc.) on *T. cruzi*.



Our research group has a long-standing interest in Tc24 (flagellar calcium-binding protein of 24 kDa). This protein has been proposed as a candidate for an immunotherapeutic vaccine, as well as a drug target. We also conducted a large-scale docking study against Tc24 using drug molecules derived from the Zinc database and a set of FDA-approved drug molecules. We found the drugs dutasteride, sirolimus, and candicidin have shown the best binding affinities against Tc24. (Supplementary Table S25).

Considering the wide variety of factors affecting the CD pathophysiology, we believe that a *T. cruzi* comprehensive map will act as a useful tool to provide information extracted from gene expression experiments, protein-protein interaction data, drug information, and clinical data information. Moreover, this will also advance our research group's effort to use the 'systems vaccinology' approach to develop safe and effective vaccines against neglected tropical diseases, including CD [95]. For instance, Querec et al., 2008 used a systems biology approach to identify the early gene 'signatures' that predicted the immune responses in humans vaccinated with the yellow fever vaccine YF-17D [96]. Similarly, Nakaya et al. (2011) found that in subjects vaccinated with the trivalent inactivated influenza vaccine, early molecular signatures correlated with and could be used to accurately predict later antibody titers in two independent trials [97]. Li et al., (2014) performed a large-scale network integration of publicly available human blood transcriptomes and systems-scale databases in specific biological contexts, and deduced a set of transcription modules in the blood [98]. Those modules revealed distinct transcriptional signatures of antibody responses to different classes of vaccines, which provided key insights into primary viral, protein recall, and anti-polysaccharide responses [98]. These examples demonstrate the power of network-based approaches to predict immunogenicity and provide new mechanistic insights about vaccines.

## 5. Conclusions

A formalized depiction of the biological pathways is increasingly recognized as a crucial requirement for the exchange of the pathway data, modeling of their activity, and systems-level interpretation of biological data. However, there are just a handful of examples of large pathway diagrams constructed using a formalized graphical modeling language, such as SBML. The model of the *T. cruzi* pathways presented here is the most comprehensive pathway of its kind published to date. Although a time-consuming and laborious exercise, the act of converting the literature-derived knowledge into a formalized computational model is essential if we wish to truly gain a systems-level understanding of any cellular system. The *T. cruzi* pathways presented here summarize the results of years of investigations and have allowed the thorough testing of the notation system used to depict it. Furthermore, we performed large-scale molecular docking using our in-house pipeline to identify potential vaccine and drug targets. Our analysis could identify potential targets, such as type B ribose 5-phosphate isomerase, flagellar calcium-binding protein, and prostaglandin F synthase. These proteins are extensively reported in the literature as potential vaccine and drug targets [91,99–101].

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/vaccines11020267/s1>, Supplementary Figure S1: The schematic diagram describing the flowchart to achieve the methodology, Supplementary Figure S2: The standard notations for the graphical representation of species and edges in cell designer v.4.4.2, Supplementary Figure S3: Docking of Tc24 protein with Benznidazole and Nifurtimox using AutoDock Vina. (a. Benznidazole, b. Nifurtimox). The top scoring pose of docking affinity against Tc24 for Benznidazole is  $-6.9$  kcal/mol while the docking affinity for Nifurtimox is  $-6.7$  kcal/mol, Supplementary Figure S4A: *Trypanosoma cruzi* CL Brener Esmeraldo-like. Enriched GO terms of genes (a. Biological process, b. Cellular component, c. Molecular function). The size of the dots and their color gradient (according to the legend) reflects the  $\log_{10}$  value of the  $p$  value of each GO term, larger dots represent higher significance (all  $p$  values  $< 0.05$ ). X and Y axes represent semantic spaces which have no intrinsic meaning. REVIGO uses multi-dimensional scaling to reduce the dimensionality of a matrix of the GO terms' pairwise semantic similarities. This results in semantically similar GO terms remaining close



together in the plot, Supplementary Figure S4B: *Trypanosoma cruzi* CL Brener Non-Esmeraldo-like. Enriched GO terms of genes (a. Biological process, b. Cellular component, c. Molecular function). The size of the dots and their color gradient (according to the legend) reflects the log10 value of the  $p$  value of each GO term, larger dots represent higher significance (all  $p$  values < 0.05). X and Y axes represent semantic spaces which have no intrinsic meaning. REVIGO uses multi-dimensional scaling to reduce the dimensionality of a matrix of the GO terms' pairwise semantic similarities. This results in semantically similar GO terms remaining close together in the plot, Supplementary Figure S5A: *Trypanosoma cruzi* CL Brener Non-Esmeraldo-like. "Interactive graph" view of REVIGO (a. Biological process, b. Cellular component, c. Molecular function). Bubble color indicates the user-provided  $p$ -value. Highly similar GO terms are linked by edges in the graph, where the line width indicates the degree of similarity, Supplementary Figure S5B: *Trypanosoma cruzi* CL Brener Esmeraldo-like. "Interactive graph" view of REVIGO (a. Biological process, b. Cellular component, c. Molecular function). Bubble color indicates the user-provided  $p$ -value. Highly similar GO terms are linked by edges in the graph, where the line width indicates the degree of similarity, Supplementary Figure S6: Gene ontology (GO) classification, KEGG/MetaCyc analysis of the *T. cruzi* genes (a.) *T. cruzi* CL Brener Esmeraldo like (b.) *T. cruzi* CL Brener Non-Esmeraldo like). The percentage regulated genes based on GO annotations and KEGG/MetaCyc pathways. (c) word cloud graph depicting enriched Metabolic Pathway (1) *Trypanosoma cruzi* CL Brener Esmeraldo-like (2) *Trypanosoma cruzi* CL Brener Non-Esmeraldo-like, Supplementary Table S1a: List of all the curated evidence for the construction of *T. cruzi* map from 46 pathways. This table provides information about 46 pathways derived from literature; constituent molecules and the literature evidence. The information is presented in 4 columns: Name; Molecule; Evidence and PMID/LINK. The first column 'Name' provides information about the name of pathways reported in the literature; The second column 'Molecule' lists the specific molecule involved in the given pathway. The third column 'Evidence' shows specific lines/statements/paragraphs from the article/research paper that provides evidence for the role of the molecule in the specific pathway. The last column 'PMID/LINK' provides PMID or hyperlink of the paper from where evidence has been collected, Supplementary Table S1b: List of curated evidence of different biological molecule interactions for the construction of 46 *T. cruzi* pathway map. This table shows a curated list of biological molecule interactions evident to construct the 46 *T. cruzi* pathway map. [Molecule 1 = first molecule of the interaction; Molecule 2 = second molecule of the interaction; Evidence = Statement from the article that provides evidence for the role of the molecule in the specific pathway; Interactions = type of interaction between molecule 1 and molecule 2; PMID/LINK = PMID or link of the paper from where evidence has been collected], Supplementary Table S2a: List of curated pathways and their classifications. This table shows the classification of curated pathways and the sources from where they were extracted. This table has information in 4 columns: Name; Classification; Source 1, and Source 2. The first column 'Name' provides information about the name of pathways reported in the literature about *T. cruzi*; The second column 'Classification' lists the classes of pathways. The third column 'Source 1' is the first source of information from where pathways are extracted. The last column 'Source 2' is the second source of information from where pathway classifications are extracted, Supplementary Table S2b: List of curated pathways and their classifications. This table holds the list of curated pathways and their classifications based on the organism and host interaction. [Pathway name = Name of the pathway; Pathway from Organism = Source organism of the pathway; Type of pathway = The type of a specific pathway], Supplementary Table S3: The list of *Tc*-CLB genes retrieved from the NCBI database. This table has information in 8 columns: Tax\_id, Organism name, GeneID, CurrentID, Status, Symbol, Aliases, and Description. The first column 'Tax\_id' is the taxonomic ID of *T. cruzi*; The second column 'Organism name' is the name of the organism with strain detail. The third column 'GeneID' is the specific ID provided to the gene by NCBI. The fourth column 'CurrentID' is the ID provided by NCBI. 'Status' gives the current status of the gene in NCBI. 'Symbol' is the gene symbol. 'Aliases' gives the available aliases of gene symbols. The last column 'Description' provides the name of the gene, Supplementary Table S4: The list of unique *Tc*-CLB genes used for the construction of the comprehensive molecular map, Supplementary Table S5a: The List of curated molecular evidence of *T. cruzi* genes with their interactions. This table has information in 5 columns: S.No., Molecule Layer 1 (*T. cruzi*), Molecule layer 2 (Human/Mice), Evidence, and PMID/LINK. The first column 'S.No.' is the serial number; The second column 'Molecule layer 1 (*T. cruzi*)' is the first molecule from *T. cruzi* in the interaction. The third column 'Molecule layer 2 (Human/Mice)' is the second molecule from Humans or Mice in the interaction. The fourth column is 'Evidence' which shows

specific lines/statements/paragraphs from the article/research paper that provides evidence for the role of the molecule in the specific pathway. The last column 'PMID/LINK' provides a PMID or hyperlink of the paper from where evidence has been collected, Supplementary Table S5b: The Summary of the source of the information. [Name = Name of the organism whose information is collected; Keyword = phrases used for literature search; PubMed hits = Number of hits found in Pubmed; Google scholar hits; Number of hits found in Google Scholar; Date = till what date data is collected], Supplementary Table S6: The list of *Tc*-CLB proteins retrieved from the Uniprot database. This table has information in 4 columns: Entry, Protein names, Gene names, and Organism. The first column 'Entry' provides UniProt ID for a specific protein entry; The second column 'Protein names' lists the names of the proteins. The third column 'Gene names' provides the names of the genes that encode the protein. The last column 'Organism' provides the name of the organism from which the protein was extracted, Supplementary Table S7: The list of unique *T. cruzi* proteins used for the construction of the comprehensive molecular map [n = 3109], Supplementary Table S8a: The list of curated molecular evidence of *T. cruzi* proteins with their interactions. This table has information in 4 columns: Protein, Molecule 2 (Human/Mice), Evidence, and PMID/Link. The first column 'Protein' provides the name of the protein from *T. cruzi*; The second column 'Molecule2 (Human/Mice)' lists the second molecule from Humans or Mice in the interaction. The third column 'Evidence' shows specific lines/statements/paragraphs from the article/research paper that provides evidence for the role of the molecule in the specific pathway. The last column 'PMID/LINK' provides a PMID or hyperlink of the paper from where evidence has been collected, Supplementary Table S8b: The Summary of the source of the information. [Name = Name of the organism whose information is collected; Keyword = phrases used for literature search; PubMed hits = Number of hits found in Pubmed; Google scholar hits; Number of hits found in Google Scholar; Date = till what date data is collected], Supplementary Table S9: The list of nodes involved in the comprehensive molecular map of *T. cruzi*. This table has information in 12 columns: class, id, name, compartment, positionToCompartment, included, quantity type, initialQuantity, substanceUnits, hasOnlySubstanceUnits, b.c., and constants. The first column 'class' provides molecule type; The second column 'id' is the specific ID assigned to each entry. The third column 'name' lists the name of the molecule. 'compartment' is the compartment where the protein belongs; 'positionToCompartment' provides information on where the molecule is located. 'included' provides information on interactions. 'quantity' provides selected option for quantity an entry in terms of either amount as molecular/item count or concentration as units of substance/units of size; 'initialQuantity' list the value set as initial quantity to run the simulation; 'substanceUnit' provides the unit assigned to each entry. 'hasOnlySubstance' gives a Boolean (True or False) on if species quantity always be as substance or substance/size; 'b.c.' is the boundary condition, which is a boolean (True or False) on should a rate of change equation be constructed for the species based on the system of reactions. The last column 'constants' is a boolean (True or False) on if the species quantity is constant, Supplementary Table S10: List of edges involved in the comprehensive molecular map of *T. cruzi*. This table has 7 columns: type; id; reversible; fast; reactants; products and modifiers. The first column 'type' provides information on the different types/categories of the reactions involved in the molecular map. The second column 'id' is the specific ID assigned to each entry. The third column 'reversible' gives a Boolean (True or False) on if the reaction is reversible or not. The fourth column 'fast' provides a Boolean (True or False) on if the reaction is fast or not. The fifth column 'reactants' lists unique molecule IDs (from Supplementary Table S9) of the reactant molecules in the reaction. The sixth column 'products' lists unique molecule IDs (from Supplementary Table S9) of the product formed in the reaction. The last column 'modifiers' lists unique molecule IDs (from Supplementary Table S9) of molecules that act as modifiers in the reaction, if any, Supplementary Table S11: The list of the total number of curated abstracts retrieved from Google scholar and PubMed against 46 unique *T. cruzi* pathways. This table has 5 columns: Name; Keyword Used; Google scholar hits; PubMed hits and Date. The first column 'Name' provides the list of *T. cruzi* pathways collected from literature and used for the construction of the molecular map. The second column 'Keyword Used' provides information about the key terms that were used for the literature search. The third column 'Google scholar hits' contains the number of abstracts/hits obtained from Google Scholar using the respective keyword. The fourth column 'PubMed hits' contains the number of abstracts/hits obtained from PubMed using the respective keyword. The last column 'Date' list the date till when the literature was extracted, Supplementary Table S12: The list of nodes involved in the network of *T. cruzi* pathway. This table has information in 10 columns: class, id, name, compartment, positionToCompartment, quantity

type, initialQuantity, hasOnlySubstanceUnits, b.c., and constants. The first column 'class' provides molecule type; The second column 'id' is the specific ID assigned to each entry. The third column 'name' lists the name of the molecule. 'compartment' is the compartment where the molecule belongs; 'positionToCompartment' provides information on where the molecule is located. 'quantity type' provides a selected option for the quantity of an entry in terms of either amount as molecular/item count or concentration as units of substance/units of size; 'initialQuantity' list the value set as initial quantity to run the simulation; 'hasOnlySubstance' gives a Boolean (True or False) on if species quantity always be as substance or substance/size; 'b.c.' is the boundary condition, which is a boolean (True or False) on should a rate of change equation be constructed for the species based on the system of reactions. The last column 'constants' is a boolean (True or False) on if the species quantity is constant, Supplementary Table S13: List of edges involved in the network of *T. cruzi* pathways. This table has 7 columns: type; id; reversible; fast; reactants; products and modifiers. The first column 'type' provides information on the different types/categories of the reactions involved in the molecular map. The second column 'id' is the specific ID assigned to each entry. The third column 'reversible' gives a Boolean (True or False) on if the reaction is reversible or not. The fourth column 'fast' provides a Boolean (True or False) on if the reaction is fast or not. The fifth column 'reactants' lists unique molecule IDs (from Supplementary Table S9) of the reactant molecules in the reaction. The sixth column 'products' lists unique molecule IDs (from Supplementary Table S9) of the product formed in the reaction. The seventh column 'modifiers' lists unique molecule IDs (from Supplementary Table S9) of molecules that act as modifiers in the reaction, if any, Supplementary Table S14: The docking of proteins from the Ubiquitin Proteasome pathway with Nifurtimox and Benznidazole. This table has 3 columns: Protein name, DB11820 (Nifurtimox) and DB11989 (Benznidazole). The first column 'Protein name' contains the list of proteins present in the ubiquitin-proteasome pathway which were docked against Nifurtimox and Benznidazole. The second column 'DB11820 (Nifurtimox)' provides the binding affinities of the various pathway proteins when docked against Nifurtimox. The third column 'DB11989 (Benznidazole)' provides the binding affinities of the various proteins when docked against Benznidazole, Supplementary Table S15a: List of curated molecular evidence of *T. cruzi* drugs with its interactions. This table has 6 columns: Name; Keyword; PubMed hits; Google scholar hits; Date; and Total. The first column 'Name' provides information on the name of the molecule. The second column 'Keyword' provides information about the key terms that were used for the literature search. The third column 'PubMed hits' contains the number of abstracts obtained by using PubMed. The fourth column 'Google Scholar hits' contains the number of abstracts obtained by Google Scholar. The fifth column 'Date' provides information about the date on which the literature search was performed/completed. The sixth column 'Total' provides the sum total of hits from PubMed and Google Scholar combined, Supplementary Table S15b: List of curated molecular evidence of *T. cruzi* drugs with its interaction. This table has 4 columns: DRUG; EVIDENCE; LINK/PMID and DATE. The first column 'DRUG' provides the name of the drug. The second column 'EVIDENCE' shows specific lines/statements/paragraphs from the article/research paper that provides evidence of the interaction between the drug and *T. cruzi* molecules. The third column 'LINK/PMID' provides PMID of the paper from where the evidence has been collected. The fourth column 'Date' provides information about the date on which the literature search was performed/completed, Supplementary Table S15c: List of curated molecular evidence of *T. cruzi* drugs with its interaction. This table has 5 columns: DRUG, MOLECULE, EVIDENCE; LINK and DATE. The first column 'DRUG' provides a list of the drug names. The second column 'MOLECULE' provides the name of the molecule that interacts with the drug; The third column 'EVIDENCE' shows specific lines/statements/paragraphs from the article/research paper that provides evidence of the interaction between drugs and molecules. The fourth column 'LINK' provides a hyperlink of the paper from where the evidence has been collected. The fifth column 'Date' provides information about the date on which the literature search was performed/completed, Supplementary Table S16: List of nodes involved in the comprehensive molecular map of drugs. This table has 10 columns: class; id; name; compartment; positionToCompartment; quantity type; initialQuantity; hasOnlySubstanceUnits; b.c. and constants. The first column 'class' provides information on the various categories of nodes present in the network i.e., drug, protein, etc. The second column 'id' is the specific ID assigned to each entry. The third column 'name' contains the names of the different molecules. The fourth column 'compartment' is the compartment where the molecule belongs. The fifth column 'positionToCompartment' provides information on where the molecule is located. The sixth column 'quantity type' provides a selected option for the quantity of an entry in terms of either amount as

molecular/item count or concentration as units of substance/units of size. The seventh column 'initialQuantity' contains the value set as initial quantity to run the simulation. The eighth column 'hasOnlySubstanceUnits' gives a Boolean (True or False) on if species quantity always be as substance or substance/size. The ninth column 'b.c.' is the boundary condition, which is a boolean (True or False) on should a rate of change equation be constructed for the species based on the system of reactions. The last column 'constants' is a boolean (True or False) on if the species quantity is constant, Supplementary Table S17: List of edges involved in the comprehensive molecular map of drugs. This table has 6 columns: type; id; reversible; fast; reactants and products. The first column 'type' provides information on the type of reactions between reactants i.e., state transitions or influences. The second column 'id' is the specific ID assigned to each entry. The third column 'reversible' helps us to know whether the reaction involved is reversible or not. The fourth column 'fast' provides information about the speed of the reaction, whether it's fast or slow. The fifth column 'reactants' contains the IDs given to the reactant(s) involved. The last column 'products' contains the IDs given to the product(s) involved, Supplementary Table S18: The list of drugs used in the network with its experimental status. This table has 5 columns: Drug Name; PMID/link; Experimental status of the drug; Drug Bank ID and FDA-approved. The first column 'Drug Name' provides the list of various *T. cruzi* drugs used in the network and which are in different stages of development; The second column 'PMID/link' contains the link for the paper from which the data was collected; The third column 'Experimental status of the drug' provides the experimental status of the drug; The fourth column 'DrugBank ID' contains the IDs of the drugs that can be used to access the details of the drugs from the DrugBank Database; The last column 'FDA Approved' provides information on the developmental stage of the drugs i.e., approved, investigational, etc., Supplementary Table S19a: The binding affinity of the crystal structure of *T. cruzi* docked with Nifurtimox (DB11820). This table has 3 columns: PDB ID; Binding affinity and Protein name. The first column 'PDB ID' provides the PDB IDs corresponding to the available crystal structures of the 127 proteins of *T. cruzi*; The second column 'Binding affinity (kcal/mol)' provides the binding affinities of the 127 *T. cruzi* proteins with Nifurtimox; The third column 'Protein name' contains the names of the proteins corresponding to the PDB IDs, Supplementary Table S19b: The binding affinity of the crystal structure of *T. cruzi* docked with Benznidazole (DB11989). This table has 3 columns: PDB ID; Binding affinity and Protein name. The first column 'PDB ID' provides the PDB IDs corresponding to the available crystal structures of the 127 proteins of *T. cruzi*; The second column 'Binding affinity (kcal/mol)' provides the binding affinities of the 127 *T. cruzi* proteins with Benznidazole; The third column 'Protein name' contains the names of the proteins corresponding to the PDB IDs, Supplementary Table S19c1: Comparison of binding affinities (kcal/mol) obtained from docking against Nifurtimox (DB11820) with top 10 crystal structure and solved structure. This table has 3 columns: PDB ID; Binding score and Protein name. The first column 'PDB ID' provides the PDB IDs corresponding to the top 10 crystal structures of the *T. cruzi* proteins. The second column 'Binding affinity (kcal/mol)' provides the binding affinities of the proteins with Nifurtimox (DB11820). The last column 'Protein name' contains the names of the proteins corresponding to the PDB IDs, Supplementary Table S19c2: Comparison of binding affinities (kcal/mol) obtained from docking against Benznidazole (DB11989) with top 10 crystal structure and solved structure. This table has 3 columns: PDB ID; Binding score and Protein name. The first column 'PDB ID' provides the PDB IDs corresponding to the top 10 crystal structures of the *T. cruzi* proteins. The second column 'Binding affinity (kcal/mol)' provides the binding affinities of the proteins with Benznidazole (DB11989). The last column 'Protein name' contains the names of the proteins corresponding to the PDB IDs, Supplementary Table S20: The binding affinities of nifurtimox, aspirin, orlistat and benznidazole drugs with proteins involved in the pentose phosphate pathway (PPP). This table has 6 columns: PPP pathway proteins; PMID; DB11820 (Nifurtimox); DB00945 (Aspirin); DB1083 (Orlistat) and DB11989 (Benznidazole). The first column 'PPP pathway proteins' lists the proteins associated with PPP pathways. The second column 'PMID' provides the PMID from PubMed for the paper from where evidence of protein association with the PPP was collected. The third column 'DB11820 (Nifurtimox)' provides the binding energy of the Nifurtimox drug against the different proteins. The fourth column 'DB00945 (Aspirin)' provides the binding energy of Aspirin against the different proteins. The fifth column 'DB1083 (Orlistat)' provides the binding energy of Orlistat against the different proteins. The last column 'DB11989 (Benznidazole)' provides the binding energy of Benznidazole drug against the different proteins, Supplementary Table S21: The binding affinity of benznidazole with 19,523 human proteins. This table has 8 columns: Entry, Entry name, Status, Protein names, Gene names, Organism, Length,

and DB11989. The first column 'Entry' list the protein IDs; the second column 'Entry name' list the name provided for each entry in the Alpha-fold database; the third column 'Status' provides the information on review status of each protein entry; the fourth column 'Protein names' list the names of each protein; the fifth column 'Gene names' list the name of the genes that encode the protein; the sixth column 'Organism' list the source organism of the protein; seventh column 'Length' list the length of each protein. The last column 'DB11989' lists benznidazole's binding affinity against each protein, Supplementary Table S22: The binding affinity of 1516 FDA-approved drugs with the crystal structure of PGF (PDB: 2F38). This table has 2 columns: Ligand and Binding energy. The first column 'Ligand' list the drug IDs and second column 'Binding Affinity' list the binding energy of each drug against prostaglandin F synthase (2F38), Supplementary Table S23a: The binding affinity of *T. cruzi* solved protein structures (targets) with nifurtimox, aspirin, orlistat and benznidazole. This table has information in 5 columns: Protein ID, DB11820 (Nifurtimox), DB00945 (Aspirin), DB01083 (Orlistat) and DB11989 (Benznidazole). The first column 'Protein ID' provides the NCBI accession ID of the protein targets; The second column 'DB11820 (Nifurtimox)' provides the binding affinity (kcal/mol) of the protein target docked against Nifurtimox; The second column 'DB00945 (Aspirin)' provides the binding affinity (kcal/mol) of the protein target docked against Aspirin; The fourth column 'DB01083 (Orlistat)' provides the binding affinity (kcal/mol) of the protein target docked against Orlistat; The fifth column 'DB11989 (Benznidazole)' provides the binding affinity (kcal/mol) of the protein target docked against Benznidazole, Supplementary Table S23b1: The molecular docking statistics of each drug docked against 5004 *T. cruzi* CL Brener solved protein structures. This table has information in 2 columns for each drug: Metric and Values. The first column 'Metric' provides the name of the metric measured to know the overall results of the drug when docked against all the protein structures; The second column provides the measured value of each metric in kcal/mol, Supplementary Table S23b2: List of protein target with highest binding affinity for each drug. This table lists the protein target with the highest binding affinity out of the total 5004 *T. cruzi* CL Brener solved protein structures for each drug. The protein name, binding affinity, residue, NCBI protein accession ID, and NCBI gene accession ID is provided for each protein target, Supplementary Table S23c: The top five ranking *T. cruzi* protein targets for nifurtimox, aspirin, orlistat and benznidazole. This table has information in 3 columns for each drug: Protein ID, Score, and Protein name. The first column 'Protein ID' provides the NCBI accession ID of the protein targets; The second column 'Score' provides the binding affinity (kcal/mol) of the protein target when docked against the respective drug; The third column 'Protein name' provides the name of the top 5 protein targets for each drug, Supplementary Table S23d: Results of Welch's T-Test applied to the binding affinity for different pairs of drug molecules. This table has information in 5 columns: Student T-Test, T-value, 95 percent confidence interval, Sample Estimates and Mean of the Differences. The first column 'Student T-Test' indicates the names of the drugs (drug pair) for which the binding affinities are tested; The second column 'T-value' provides the T-value on the basis of the binding affinities of the two drugs to the *T. cruzi* protein structures; The third column '95 percent confidence interval' provides the 95 percent confidence interval on the basis of the binding affinities of the two drugs to the *T. cruzi* protein structures; The third column 'Sample Estimates' provides the sample estimates on the basis of the binding affinities of the two drugs to the *T. cruzi* protein structures, Supplementary Table S24a: The binding affinity of *Plasmodium falciparum* with Benznidazole (DB11989). This table has 3 columns. The first column 'Protein ID of *P. falciparum* proteins' provides the PDB ID of *P. falciparum* proteins; The second column 'Binding affinity (kcal/mol)' is the binding affinity of *P. falciparum* proteins when docked against Benznidazole (DB11989) drug; The third column 'Name of the protein' provides the name of the *P. falciparum* protein, Supplementary Table S24b: The binding affinity of *Mycobacterium tuberculosis* with Benznidazole (DB11989). This table has 3 columns. The first column 'Protein ID of *M. tuberculosis* proteins' provides the PDB ID of *M. tuberculosis* proteins; The second column 'Binding affinity (kcal/mol)' is the binding affinity of *M. tuberculosis* proteins when docked against Benznidazole (DB11989) drug; The third column 'Name of the protein' provides the name of the *M. tuberculosis* protein, Supplementary Table S24c: The binding affinity of *Leishmania donovani* with Benznidazole (DB11989). This table has 3 columns. The first column 'Protein ID of *Leishmania donovani* proteins' provides the PDB ID of *L. donovani* proteins; The second column 'Binding affinity (kcal/mol)' is the binding affinity of *L. donovani* proteins when docked against Benznidazole (DB11989) drug; The third column 'Name of the protein' provides the name of the *L. donovani* protein, Supplementary Table S24d: The binding affinity of *Salmonella typhi* with Benznidazole (DB11989). This table has 3 columns. The first column 'Protein ID of *Salmonella typhi* proteins' provides the PDB



ID of *S. typhi* proteins; The second column 'Binding affinity (kcal/mol)' is the binding affinity of *S. typhi* proteins when docked against Benznidazole (DB11989) drug; The third column 'Name of the protein' provides the name of the *S. typhi* protein, Supplementary Table S24e: Results of Welch's T-Test applied to the binding affinity of *T. cruzi* solved structures and other pathogens. This table has information in 4 columns: Student T-Test, T-value, 95 percent confidence interval, and Sample Estimates. The first column 'Student T-Test' indicates the name of the pathogen for which the binding affinities are tested against *T. cruzi* solved structures; The second column 'T-value' provides the T-value on the basis of the binding affinities of *T. cruzi* solved structures and the pathogen; The third column '95 percent confidence interval' provides the 95 percent confidence interval on the basis of the binding affinities of *T. cruzi* solved structures and the pathogen; The fourth column 'Sample Estimates' provides the sample estimates on the basis of the binding affinities of *T. cruzi* solved structures and the pathogen, Supplementary Table S25: The binding affinity of FDA-approved drug molecules docked against Tc24. This table has information in 2 columns: Ligand, and Affinity (kcal/mol). The first column 'Ligand' provides the DrugBank ID of the drug molecules docked against Tc24; The second column 'Affinity (kcal/mol)' provides the strength of the interaction between the drug molecule and Tc24 that bind reversibly, Supplementary Table S26: Docking results of Benznidazole and Nifurtimox against Tc24 protein using AutoDock Vina. This table has information in 5 columns: Drug name, Mode (Top 30), Affinity (kcal/mol), Distance from the best mode (rmsd l.b.), and Distance from the best mode (rmsd u.b.). The first column 'Drug name' provides the name of the drug docked against Tc24; The second column Mode (Top 3) provides the orientation of the ligand relative to the receptor as well as the conformation of the ligand and receptor when bound to each other; The third column Affinity (kcal/mol) provides the strength of the interaction between the drug molecule and Tc24 that bind reversibly; The fourth and fifth columns 'Distance from the best mode (rmsd l.b.)' and 'Distance from best mode (rmsd u.b.)' provides the distance of that particular mode from the best mode in terms of Root Mean Square Deviation (RMSD)-upper bound and lower bound, Supplementary Table S27a: Simple parameters of *T. cruzi* pathways map determined using the network analyzer Cytoscape plugin. This table has information in 2 columns: Type and Statistics. The first column 'Type' provides the type of the parameter; The second column 'Statistics' provides the value of the parameter measured using Cytoscape, Supplementary Table S27b: Simple parameters of *T. cruzi* molecular map determined using the network analyzer Cytoscape plugin. This table has information in 2 columns: Type and Statistics. The first column 'Type' provides the type of the parameter; The second column 'Statistics' provides the value of the parameter measured using Cytoscape, Supplementary Table S27c: Simple parameters of *T. cruzi* drug map determined using the network analyzer Cytoscape plugin. This table has information in 2 columns: Type and Statistics. The first column 'Type' provides the type of the parameter; The second column 'Statistics' provides the value of the parameter measured using Cytoscape, Supplementary Table S28: List of *T. cruzi* genes (class: Trypanosomatidae) found from the TritypDB database during Gene Ontology Analysis. This table has information in 7 columns: Gene ID, source\_id, organism, Genomic Location (Gene), Description, Gene Type and Input ID. The first column 'Gene ID' provides the ID of the enriched gene; The second column 'source\_id' provides the source ID of the enriched gene; The third column 'Organism' provides the name of the organism to which the enriched gene belongs; The fourth column 'Gene Location (Gene)' provides the location of the enriched gene; The fifth column 'Description' describes the enriched gene; The sixth column 'Gene Type' provides the type of the enriched gene; The seventh column 'Input ID' provides the input ID of the enriched gene, Supplementary Table S29: Total list of enriched genes that belong to *Trypanosoma cruzi* species (*Trypanosoma cruzi* CL Brener Esmeraldo-like, and *Trypanosoma cruzi* CL Brener Non-Esmeraldo-like). This table has information in 4 columns: Gene ID, source\_id, organism, Genomic Location (Gene), Description, Gene Type, and Input ID. The first column 'Gene ID' provides the ID of the enriched gene; The second column 'source\_id' provides the source ID of the enriched gene; The third column 'Organism' provides the name of the organism to which the enriched gene belongs; The fourth column 'Gene Location (Gene)' provides the location of the enriched gene; The fifth column 'Description' describes the enriched gene; The sixth column 'Gene Type' provides the type of the enriched gene; The seventh column 'Input ID' provides the input ID of the enriched gene, Supplementary Table S30a: List of enriched genes in Biological processes related to *Trypanosoma cruzi* CL Brener Non-Esmeraldo-like. This table has information in 11 columns: ID; Name; Bgd count; Result count; Result gene list; Pct of bgd; Fold enrichment; Odds ratio; *p*-value; Benjamini; and Bonferroni. The first column 'ID' provides the ID of the BP; The second column 'Name' provides the name of the BP; The third column 'Bgd

count' provides the number of genes with this BP in the genome; The fourth column 'Result count' provides the number of genes with this BP in our analysis; The fifth 'Result gene list' provides the names of the genes with this BP in our analysis; The sixth column 'Pct of Bgd' provides percent of genes with this BP in our analysis divided by the percent of genes with this BP in the genome; The seventh column 'Fold enrichment' provides the percent that are present in your analysis of the genes in the genome with this BP; The eighth column 'Odds ratio' provides -The odds of the BP appearing in the gene list are the same as that for the background list; The ninth column '*p*-value' provides the probability of seeing at least *x* number of genes out of the total *n* genes in the list annotated to the BP, given the proportion of genes in the whole genome that are annotated to that BP; The tenth column 'Benjamini' provides the Benjamini-Hochburg false discovery rate which is a method for controlling false discovery rates for type 1 errors; The eleventh column 'Bonferroni' provides the Bonferroni adjusted *p*-values which is a method for correcting significance based on multiple comparisons, Supplementary Table S30b: List of enriched genes in molecular function related to *Trypanosoma cruzi* CL Brener Non-Esmeraldo-like. This table has information in 11 columns: ID; Name; Bgd count; Result count; Result gene list; Pct of bgd; Fold enrichment; Odds ratio; *p*-value; Benjamini; and Bonferroni. The first column 'ID' provides the ID of the MF; The second column 'Name' provides the name of the MF; The third column 'Bgd count' provides the number of genes with this MF in the genome; The fourth column 'Result count' provides the number of genes with this MF in our analysis; The fifth 'Result gene list' provides the names of the genes with this MF in our analysis; The sixth column 'Pct of Bgd' provides percent of genes with this MF in our analysis divided by the percent of genes with this MF in the genome; The seventh column 'Fold enrichment' provides the percent that are present in your analysis of the genes in the genome with this MF; The eighth column 'Odds ratio' provides -The odds of the MF appearing in the gene list are the same as that for the background list; The ninth column '*p*-value' provides the probability of seeing at least *x* number of genes out of the total *n* genes in the list annotated to the MF, given the proportion of genes in the whole genome that are annotated to that MF; The tenth column 'Benjamini' provides the Benjamini-Hochburg false discovery rate which is a method for controlling false discovery rates for type 1 errors; The eleventh column 'Bonferroni' provides the Bonferroni adjusted *p*-values which is a method for correcting significance based on multiple comparisons, Supplementary Table S30c: List of enriched genes in cellular component related to *Trypanosoma cruzi* CL Brener Non-Esmeraldo-like. This table has information in 11 columns: ID; Name; Bgd count; Result count; Result gene list; Pct of bgd; Fold enrichment; Odds ratio; *p*-value; Benjamini; and Bonferroni. The first column 'ID' provides the ID of the CC; The second column 'Name' provides the name of the CC; The third column 'Bgd count' provides the number of genes with this CC in the genome; The fourth column 'Result count' provides the number of genes with this CC in our analysis; The fifth 'Result gene list' provides the names of the genes with this CC in our analysis; The sixth column 'Pct of Bgd' provides percent of genes with this CC in our analysis divided by the percent of genes with this CC in the genome; The seventh column 'Fold enrichment' provides the percent that are present in your analysis of the genes in the genome with this CC; The eighth column 'Odds ratio' provides -The odds of the CC appearing in the gene list are the same as that for the background list; The ninth column '*p*-value' provides the probability of seeing at least *x* number of genes out of the total *n* genes in the list annotated to the CC, given the proportion of genes in the whole genome that are annotated to that CC; The tenth column 'Benjamini' provides the Benjamini-Hochburg false discovery rate which is a method for controlling false discovery rates for type 1 errors; The eleventh column 'Bonferroni' provides the Bonferroni adjusted *p*-values which is a method for correcting significance based on multiple comparisons, Supplementary Table S30d: List of enriched genes in metabolic pathways related to *Trypanosoma cruzi* CL Brener Non-Esmeraldo-like. This table has information in 11 columns: ID; Name; Bgd count; Result count; Result gene list; Pct of bgd; Fold enrichment; Odds ratio; *p*-value; Benjamini; and Bonferroni. The first column 'ID' provides the ID of the MP; The second column 'Name' provides the name of the MP; The third column 'Bgd count' provides the number of genes with this MP in the genome; The fourth column 'Result count' provides the number of genes with this MP in our analysis; The fifth 'Result gene list' provides the names of the genes with this MP in our analysis; The sixth column 'Pct of Bgd' provides percent of genes with this MP in our analysis divided by the percent of genes with this MP in the genome; The seventh column 'Fold enrichment' provides the percent that are present in your analysis of the genes in the genome with this MP; The eighth column 'Odds ratio' provides -The odds of the MP appearing in the gene list are the same as that for the background list; The ninth column '*p*-value' provides the probability of seeing at least

x number of genes out of the total n genes in the list annotated to the MP, given the proportion of genes in the whole genome that are annotated to that MP; The tenth column 'Benjamini' provides the Benjamini-Hochburg false discovery rate which is a method for controlling false discovery rates for type 1 errors; The eleventh column 'Bonferroni' provides the Bonferroni adjusted *p*-values which is a method for correcting significance based on multiple comparisons, Supplementary Table S31a: List of enriched genes in biological process (BP) related to *Trypanosoma cruzi* CL Brener Esmeraldo-like. This table has information in 11 columns: ID; Name; Bgd count; Result count; Result gene list; Pct of bgd; Fold enrichment; Odds ratio; *p*-value; Benjamini; and Bonferroni. The first column 'ID' provides the ID of the BP; The second column 'Name' provides the name of the BP; The third column 'Bgd count' provides the number of genes with this BP in the genome; The fourth column 'Result count' provides the number of genes with this BP in our analysis; The fifth 'Result gene list' provides the names of the genes with this BP in our analysis; The sixth column 'Pct of Bgd' provides percent of genes with this BP in our analysis divided by the percent of genes with this BP in the genome; The seventh column 'Fold enrichment' provides the percent that are present in your analysis of the genes in the genome with this BP; The eighth column 'Odds ratio' provides -The odds of the BP appearing in the gene list are the same as that for the background list; The ninth column '*p*-value' provides the probability of seeing at least x number of genes out of the total n genes in the list annotated to the BP, given the proportion of genes in the whole genome that are annotated to that BP; The tenth column 'Benjamini' provides the Benjamini-Hochburg false discovery rate which is a method for controlling false discovery rates for type 1 errors; The eleventh column 'Bonferroni' provides the Bonferroni adjusted *p*-values which is a method for correcting significance based on multiple comparisons, Supplementary Table S31b: List of enriched genes in molecular function (MF) related to *Trypanosoma cruzi* CL Brener Esmeraldo-like. This table has information in 11 columns: ID; Name; Bgd count; Result count; Result gene list; Pct of bgd; Fold enrichment; Odds ratio; *p*-value; Benjamini; and Bonferroni. The first column 'ID' provides the ID of the MF; The second column 'Name' provides the name of the MF; The third column 'Bgd count' provides the number of genes with this MF in the genome; The fourth column 'Result count' provides the number of genes with this MF in our analysis; The fifth 'Result gene list' provides the names of the genes with this MF in our analysis; The sixth column 'Pct of Bgd' provides percent of genes with this MF in our analysis divided by the percent of genes with this MF in the genome; The seventh column 'Fold enrichment' provides the percent that are present in your analysis of the genes in the genome with this MF; The eighth column 'Odds ratio' provides -The odds of the MF appearing in the gene list are the same as that for the background list; The ninth column '*p*-value' provides the probability of seeing at least x number of genes out of the total n genes in the list annotated to the MF, given the proportion of genes in the whole genome that are annotated to that MF; The tenth column 'Benjamini' provides the Benjamini-Hochburg false discovery rate which is a method for controlling false discovery rates for type 1 errors; The eleventh column 'Bonferroni' provides the Bonferroni adjusted *p*-values which is a method for correcting significance based on multiple comparisons, Supplementary Table S31c: List of enriched genes in cellular component related to *Trypanosoma cruzi* CL Brener Esmeraldo-like. This table has information in 11 columns: ID; Name; Bgd count; Result count; Result gene list; Pct of bgd; Fold enrichment; Odds ratio; *p*-value; Benjamini; and Bonferroni. The first column 'ID' provides the ID of the CC; The second column 'Name' provides the name of the CC; The third column 'Bgd count' provides the number of genes with this CC in the genome; The fourth column 'Result count' provides the number of genes with this CC in our analysis; The fifth 'Result gene list' provides the names of the genes with this CC in our analysis; The sixth column 'Pct of Bgd' provides percent of genes with this CC in our analysis divided by the percent of genes with this CC in the genome; The seventh column 'Fold enrichment' provides the percent that are present in your analysis of the genes in the genome with this CC; The eighth column 'Odds ratio' provides -The odds of the CC appearing in the gene list are the same as that for the background list; The ninth column '*p*-value' provides the probability of seeing at least x number of genes out of the total n genes in the list annotated to the CC, given the proportion of genes in the whole genome that are annotated to that CC; The tenth column 'Benjamini' provides the Benjamini-Hochburg false discovery rate which is a method for controlling false discovery rates for type 1 errors; The eleventh column 'Bonferroni' provides the Bonferroni adjusted *p*-values which is a method for correcting significance based on multiple comparisons, Supplementary Table S31d: List of enriched genes in molecular function related to *Trypanosoma cruzi* CL Brener Esmeraldo-like. This table has information in 11 columns: ID; Name; Bgd count; Result count; Result gene list; Pct of bgd; Fold enrichment; Odds ratio; *p*-value; Benjamini; and Bonferroni. The first column 'ID' provides

the ID of the MP; The second column 'Name' provides the name of the MP; The third column 'Bgd count' provides the number of genes with this MP in the genome; The fourth column 'Result count' provides the number of genes with this MP in our analysis; The fifth 'Result gene list' provides the names of the genes with this MP in our analysis; The sixth column 'Pct of Bgd' provides percent of genes with this MP in our analysis divided by the percent of genes with this MP in the genome; The seventh column 'Fold enrichment' provides the percent that are present in your analysis of the genes in the genome with this MP; The eighth column 'Odds ratio' provides -The odds of the MP appearing in the gene list are the same as that for the background list; The ninth column '*p*-value' provides the probability of seeing at least *x* number of genes out of the total *n* genes in the list annotated to the MP, given the proportion of genes in the whole genome that are annotated to that MP; The tenth column 'Benjamini' provides the Benjamini-Hochburg false discovery rate which is a method for controlling false discovery rates for type 1 errors; The eleventh column 'Bonferroni' provides the Bonferroni adjusted *p*-values which is a method for correcting significance based on multiple comparisons, Supplementary file S1: The 86,990 abstracts (Google Scholar-85,600, PubMed-1390), as of February 2021 were collected using the literature mining approach, Supplementary file S2: The 103,868 abstracts (Google Scholar-96,300, PubMed-7568), as of February 2021, were screened to obtain the molecular information related to *T. cruzi*, Supplementary file S3: The 1693 abstracts with the terms "*T. cruzi*" AND "drug" were downloaded from Google Scholar and PubMed.

**Author Contributions:** K.R., P.J.H., U.S. and M.E.B. designed the research. P.P., P.K., T.S., S.K.N. and S.S., performed the research. N.G. performed the docking experiments. T.S., M.R., K.M., N.A., M.S., D.H., G.K., S.V., A.C., S.B., P.T. and R.A. performed the data curation and constructed a network. K.R., P.K., P.P. and S.K.N., contributed to writing the manuscript. P.P., T.S. and P.K. compiled the information from many different sources. All authors have read and agreed to the published version of the manuscript.

**Funding:** P.K.: S.S., S.K.N., N.G. and T.S. have received the research fellowship supported by a Grant from the Robert J. Kleberg, Jr. and Helen C. Kleberg Foundation, USA, Texas Children's Hospital, and Baylor College of Medicine, Houston USA. P.P. has received a research fellowship from SERB, Department of Science and Technology (File Number: CVD/2020/000842) and ICMR, Indian Council of Medical Research (BMI/12(66)/2021 2021-6442). We used the computational infrastructure provided by the Department of Biotechnology (DBT), Ministry of Science and Technology Government of India received by P.P. (Grant ID: BT/PRI7252/BID/7/708/2016).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data is available in the supplementary materials. In addition, data can be accessed on a dedicated website (<https://tinyurl.com/Tcruzipathwaymapx>).

**Acknowledgments:** We are thankful to Amity University for the support in providing laboratory facilities to conduct the study. We are thankful to Robin Sinha, Prashant Singh, Amit Chaudhary, Monu Kaushik, Prem Prakash Sharma, Priyanshi Dixit, Priya Sharma, Manshi Rai, Muskan Tanwer, Chhaya Bawa, Sadaf Raza, Aishwarya Dharan, Dhruvi Kakadiya, and Isha Joshi for their contribution to the conduct of this study. We are also thankful to Vrinda Khanna, Gayatri Pradhan, and Akanksha Verma for their assistance during the reviewing process.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. De Pablos, L.M.; Osuna, A. Multigene families in *Trypanosoma cruzi* and their role in infectivity. *Infect. Immun.* **2012**, *80*, 2258–2264. [CrossRef] [PubMed]
2. El-Sayed, N.M.; Myler, P.J.; Bartholomeu, D.C.; Nilsson, D.; Aggarwal, G.; Tran, A.-N.; Ghedin, E.; Worthey, E.A.; Delcher, A.L.; Blandin, G.; et al. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* **2005**, *309*, 409–415. [CrossRef]
3. Weston, D.; Patel, B.; Van Voorhis, W.C. Virulence in *Trypanosoma cruzi* infection correlates with the expression of a distinct family of sialidase superfamily genes. *Mol. Biochem. Parasitol.* **1999**, *98*, 105–116. [CrossRef] [PubMed]
4. Nde, P.N.; Johnson, C.A.; Pratap, S.; Cardenas, T.C.; Kleshchenko, Y.Y.; Furtak, V.A.; Simmons, K.J.; Lima, M.F.; Villalta, F. Gene Network Analysis during Early Infection of Human Coronary Artery Smooth Muscle Cells by *Trypanosoma cruzi* and Its gp83 Ligand. *Chem. Biodivers.* **2010**, *7*, 1051–1064. [CrossRef]

5. Roberts, S.B.; Robichaux, J.L.; Chavali, A.K.; Manque, P.A.; Lee, V.; Lara, A.M.; Papin, J.A.; Buck, G.A. Proteomic and network analysis characterize stage-specific metabolism in *Trypanosoma cruzi*. *BMC Syst. Biol.* **2009**, *3*, 52. [CrossRef]
6. Shiratsubaki, I.S.; Fang, X.; Souza, R.O.O.; Palsson, B.O.; Silber, A.M.; Siqueira-Neto, J.L. Genome-scale metabolic models highlight stage-specific differences in essential metabolic pathways in *Trypanosoma cruzi*. *PLoS Negl. Trop. Dis.* **2020**, *14*, e0008728. [CrossRef]
7. Goh, K.-I.; Cusick, M.E.; Valle, D.; Childs, B.; Vidal, M.; Barabási, A.-L. The human disease network. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8685–8690. [CrossRef]
8. Janjić, V.; Pržulj, N. The Core Diseasesome. *Mol. Biosyst.* **2012**, *8*, 2614–2625. [CrossRef]
9. Lusis, A.J.; Weiss, J.N. Cardiovascular networks: Systems-based approaches to cardiovascular disease. *Circulation* **2010**, *121*, 157–170. [CrossRef] [PubMed]
10. Jagannadham, J.; Jaiswal, H.K.; Agrawal, S.; Rawal, K. Comprehensive Map of Molecules Implicated in Obesity. *PLoS ONE* **2016**, *11*, e0146759. [CrossRef]
11. Ferraz, M.L.; Gazzinelli, R.T.; Alves, R.O.; Urbina, J.A.; Romanha, A.J. The Anti-*Trypanosoma cruzi* Activity of Posaconazole in a Murine Model of Acute Chagas' Disease Is Less Dependent on Gamma Interferon than That of Benznidazole. *Antimicrob. Agents Chemother.* **2007**, *51*, 1359–1364. [CrossRef] [PubMed]
12. Assíria Fontes Martins, T.; de Figueiredo Diniz, L.; Mazzeti, A.L.; da Silva do Nascimento, Á.F.; Caldas, S.; Caldas, I.S.; de Andrade, I.M.; Ribeiro, I.; Bahia, M.T. Benznidazole/Itraconazole Combination Treatment Enhances Anti-*Trypanosoma cruzi* Activity in Experimental Chagas Disease. *PLoS ONE* **2015**, *10*, e0128707. [CrossRef] [PubMed]
13. Molina, I.; Perin, L.; Aviles, A.S.; de Abreu Vieira, P.M.; da Silva Fonseca, K.; Cunha, L.M.; Carneiro, C.M. The effect of benznidazole dose among the efficacy outcome in the murine animal model. *A quantitative integration of the literature. Acta Trop.* **2020**, *201*, 105218. [CrossRef] [PubMed]
14. Brodsky, I.E.; Medzhitov, R. Targeting of immune signalling networks by bacterial pathogens. *Nat. Cell Biol.* **2009**, *11*, 521–526. [CrossRef] [PubMed]
15. Chanumolu, S.K.; Rout, C.; Chauhan, R.S. UniDrug-Target: A Computational Tool to Identify Unique Drug Targets in Pathogenic Bacteria. *PLoS ONE* **2012**, *7*, e32833. [CrossRef] [PubMed]
16. Luo, M.; Jiao, J.; Wang, R. Screening drug target combinations in disease-related molecular networks. *BMC Bioinform.* **2019**, *20*, 198. [CrossRef] [PubMed]
17. Dai, Y.-F.; Zhao, X.-M. A Survey on the Computational Approaches to Identify Drug Targets in the Postgenomic Era. *BioMed Res. Int.* **2015**, *2015*, 1–9. [CrossRef]
18. Rawal, K.; Sinha, R.; Abbasi, B.A.; Chaudhary, A.; Nath, S.K.; Kumari, P.; Preeti, P.; Saraf, D.; Singh, S.; Mishra, K.; et al. Identification of vaccine targets in pathogens and design of a vaccine using computational approaches. *Sci. Rep.* **2021**, *11*, 17626. [CrossRef]
19. Rawal, K.; Sinha, R.; Nath, S.K.; Preeti, P.; Kumari, P.; Gupta, S.; Sharma, T.; Strych, U.; Hotez, P.; Bottazzi, M.E. Vaxi-DL: A web-based deep learning server to identify potential vaccine candidates. *Comput. Biol. Med.* **2022**, *145*, 105401. [CrossRef]
20. Abbasi, B.A.; Saraf, D.; Sharma, T.; Sinha, R.; Singh, S.; Sood, S.; Gupta, P.; Gupta, A.; Mishra, K.; Kumari, P.; et al. Identification of vaccine targets & design of vaccine against SARS-CoV-2 coronavirus using computational and deep learning-based approaches. *PeerJ* **2022**, *10*, e13380. [CrossRef]
21. Qin, G.; Zhao, X.-M. A survey on computational approaches to identifying disease biomarkers based on molecular networks. *J. Theor. Biol.* **2014**, *362*, 9–16. [CrossRef] [PubMed]
22. Agwa, S.H.A.; Elghazaly, H.; Meteini, M.S.E.; Shawky, S.M.; Ali, M.; Abd Elsamee, A.M.; Sayed, S.M.; Sherif, N.; Sharaf, H.M.; Alhadidy, M.A.; et al. In Silico Identification and Clinical Validation of a Novel Long Non-Coding RNA/mRNA/miRNA Molecular Network for Potential Biomarkers for Discriminating SARS CoV-2 Infection Severity. *Cells* **2021**, *10*, 3098. [CrossRef]
23. Gao, Y.; Xu, T.; Zhao, Y.-X.; Ling-Hu, T.; Liu, S.-B.; Tian, J.-S.; Qin, X.-M. A Novel Network Pharmacology Strategy to Decode Metabolic Biomarkers and Targets Interactions for Depression. *Front. Psychiatry* **2020**, *11*, 667. [CrossRef] [PubMed]
24. CellDesigner. Available online: <https://www.celldesigner.org/> (accessed on 18 February 2022).
25. Funahashi, A.; Morohashi, M.; Kitano, H.; Tanimura, N. CellDesigner: A process diagram editor for gene-regulatory and biochemical networks. *Biosilico* **2003**, *5*, 159–162. [CrossRef]
26. Funahashi, A.; Matsuoka, Y.; Jouraku, A.; Morohashi, M.; Kikuchi, N.; Kitano, H. CellDesigner 3.5: A versatile modeling tool for biochemical networks. *Proc. IEEE* **2008**, *96*, 1254–1265. [CrossRef]
27. *Trypanosoma cruzi* CL Brener-Gene-NCBI. Available online: <https://www.ncbi.nlm.nih.gov/gene/?term=Trypanosoma+cruzi+CL+brener> (accessed on 6 January 2022).
28. *Trypanosoma cruzi* cl Brener in UniProtKB Search (19656) | UniProt. Available online: <https://www.uniprot.org/uniprotkb/?query=trypanosoma+cruzi+cl+brener&sort=score> (accessed on 10 January 2022).
29. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [CrossRef]
30. Harris, M.A.; Clark, J.; Ireland, A.; Lomax, J.; Ashburner, M.; Foulger, R.; Eilbeck, K.; Lewis, S.; Marshall, B.; Mungall, C.; et al. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* **2004**, *32*, D258–D261. [CrossRef]
31. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [CrossRef]
32. Karp, P.D.; Riley, M.; Paley, S.M.; Pellegrini-Toole, A. The MetaCyc Database. *Nucleic Acids Res.* **2002**, *30*, 59–61. [CrossRef]



33. Aslett, M.; Aurrecochea, C.; Berriman, M.; Brestelli, J.; Brunk, B.P.; Carrington, M.; Depledge, D.P.; Fischer, S.; Gajria, B.; Gao, X.; et al. TriTrypDB: A functional genomic resource for the Trypanosomatidae. *Nucleic Acids Res.* **2010**, *38*, D457–D462. [CrossRef]
34. TriTrypDB. Available online: <https://tritrypdb.org/tritrypdb/app> (accessed on 13 April 2022).
35. Supek, F.; Bošnjak, M.; Škunca, N.; Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* **2011**, *6*, e21800. [CrossRef] [PubMed]
36. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinform.* **2016**, *54*, 5–6. [CrossRef] [PubMed]
37. Pearson, W.R. An Introduction to Sequence Similarity (“Homology”) Searching. *Curr. Protoc. Bioinform.* **2013**, *42*, 3.1.1–3.1.8. [CrossRef] [PubMed]
38. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [CrossRef]
39. Sali, A.; Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815. [CrossRef]
40. Knox, C.; Law, V.; Jewison, T.; Liu, P.; Ly, S.; Frolkis, A.; Pon, A.; Banco, K.; Mak, C.; Neveu, V.; et al. DrugBank 3.0: A comprehensive resource for “omics” research on drugs. *Nucleic Acids Res.* **2011**, *39*, D1035–D1041. [CrossRef]
41. Krivák, R.; Hoksza, D. P2Rank: Machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure. *J. Cheminformatics* **2018**, *10*, 39. [CrossRef]
42. O’Boyle, N.M.; Banck, M.; James, C.A.; Morley, C.; Vandermeersch, T.; Hutchison, G.R. Open Babel: An open chemical toolbox. *J. Cheminformatics* **2011**, *3*, 33. [CrossRef]
43. Lecker, S.H.; Goldberg, A.L.; Mitch, W.E. Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *J. Am. Soc. Nephrol. JASN* **2006**, *17*, 1807–1819. [CrossRef]
44. Gupta, I.; Aggarwal, S.; Singh, K.; Yadav, A.; Khan, S. Ubiquitin Proteasome pathway proteins as potential drug targets in parasite *Trypanosoma cruzi*. *Sci. Rep.* **2018**, *8*, 8399. [CrossRef]
45. Bijlmakers, M.-J. Ubiquitination and the Proteasome as Drug Targets in Trypanosomatid Diseases. *Front. Chem.* **2020**, *8*, 630888. [CrossRef]
46. de Diego, J.L.; Katz, J.M.; Marshall, P.; Gutiérrez, B.; Manning, J.E.; Nussenzweig, V.; González, J. The ubiquitin-proteasome pathway plays an essential role in proteolysis during *Trypanosoma cruzi* remodeling. *Biochemistry* **2001**, *40*, 1053–1062. [CrossRef]
47. Rassi, A.; Luquetti, A.O.; Rassi, A.; Rassi, G.G.; Rassi, S.G.; DA Silva, I.G.; Rassi, A.G. Specific treatment for *Trypanosoma cruzi*: Lack of efficacy of allopurinol in the human chronic phase of Chagas disease. *Am. J. Trop. Med. Hyg.* **2007**, *76*, 58–61. [CrossRef]
48. Lo Presti, M.S.; Bazán, P.C.; Strauss, M.; Báez, A.L.; Rivarola, H.W.; Paglini-Oliva, P.A. Trypanothione reductase inhibitors: Overview of the action of thioridazine in different stages of Chagas disease. *Acta Trop.* **2015**, *145*, 79–87. [CrossRef]
49. D’Avila, H.; Toledo, D.A.M.; Melo, R.C.N. Lipid Bodies: Inflammatory Organelles Implicated in Host-*Trypanosoma cruzi* Interplay during Innate Immune Responses. *Mediat. Inflamm.* **2012**, *2012*, 1–11. [CrossRef]
50. Buckner, F.S.; Bahia, M.T.; Suryadevara, P.K.; White, K.L.; Shackleford, D.M.; Chennamaneni, N.K.; Hulverson, M.A.; Laydbak, J.U.; Chatelain, E.; Scandale, I.; et al. Pharmacological Characterization, Structural Studies, and In Vivo Activities of Anti-Chagas Disease Lead Compounds Derived from Tipifarnib. *Antimicrob. Agents Chemother.* **2012**, *56*, 4914–4921. [CrossRef]
51. Molina, I.; Gómez i Prat, J.; Salvador, F.; Treviño, B.; Sulleiro, E.; Serre, N.; Pou, D.; Roure, S.; Cabezas, J.; Valerio, L.; et al. Randomized Trial of Posaconazole and Benznidazole for Chronic Chagas’ Disease. *N. Engl. J. Med.* **2014**, *370*, 1899–1908. [CrossRef]
52. Jackson, Y.; Alirol, E.; Getaz, L.; Wolff, H.; Combescure, C.; Chappuis, F. Tolerance and Safety of Nifurtimox in Patients with Chronic Chagas Disease. *Clin. Infect. Dis.* **2010**, *51*, e69–e75. [CrossRef] [PubMed]
53. Pinazo, M.-J.; Muñoz, J.; Posada, E.; López-Chejade, P.; Gállego, M.; Ayala, E.; del Cacho, E.; Soy, D.; Gascon, J. Tolerance of Benznidazole in Treatment of Chagas’ Disease in Adults. *Antimicrob. Agents Chemother.* **2010**, *54*, 4896–4899. [CrossRef]
54. Apt, W.; Arribada, A.; Zulantay, I.; Rodriguez, J.; Saavedra, M.; Munoz, A. Treatment of Chagas’ disease with itraconazole: Electrocardiographic and parasitological conditions after 20 years of follow-up. *J. Antimicrob. Chemother.* **2013**, *68*, 2164–2169. [CrossRef]
55. Coura, J.R.; Borges-Pereira, J. Chagas disease. What is known and what should be improved: A systemic review. *Rev. Soc. Bras. Med. Trop.* **2012**, *45*, 286–296. [CrossRef]
56. Moraes, C.B.; Giardini, M.A.; Kim, H.; Franco, C.H.; Araujo-Junior, A.M.; Schenkman, S.; Chatelain, E.; Freitas-Junior, L.H. Nitroheterocyclic compounds are more efficacious than CYP51 inhibitors against *Trypanosoma cruzi*: Implications for Chagas disease drug discovery and development. *Sci. Rep.* **2014**, *4*, 4703. [CrossRef] [PubMed]
57. Pinto, A.V.; de Castro, S.L. The trypanocidal activity of naphthoquinones: A review. *Mol. Basel Switz.* **2009**, *14*, 4570–4590. [CrossRef]
58. da Silva Santos-Júnior, P.F.; Schmitt, M.; de Araújo-Júnior, J.X.; da Silva-Júnior, E.F. Sterol 14 $\alpha$ -Demethylase from Trypanosomatidae Parasites as a Promising Target for Designing New Antiparasitic Agents. *Curr. Top. Med. Chem.* **2021**, *21*, 1900–1921. [CrossRef]
59. Cotuá, J.; LLinás, H.; Cotes, S. Virtual Screening Based on QSAR and Molecular Docking of Possible Inhibitors Targeting Chagas CYP51. *J. Chem.* **2021**, *2021*, e6640624. [CrossRef]
60. AlphaFold Protein Structure Database. Available online: <https://alphafold.ebi.ac.uk/> (accessed on 16 September 2022).

61. León-Del-Río, A. Biotin in metabolism, gene expression, and human disease. *J. Inherit. Metab. Dis.* **2019**, *42*, 647–654. [\[CrossRef\]](#)
62. Pendini, N.R.; Bailey, L.M.; Booker, G.W.; Wilce, M.C.J.; Wallace, J.C.; Polyak, S.W. Biotin protein ligase from *Candida albicans*: Expression, purification and development of a novel assay. *Arch. Biochem. Biophys.* **2008**, *479*, 163–169. [\[CrossRef\]](#)
63. Häberle, J.; Görg, B.; Rutsch, F.; Schmidt, E.; Toutain, A.; Benoist, J.-F.; Gelot, A.; Suc, A.-L.; Höhne, W.; Schliess, F.; et al. Congenital glutamine deficiency with glutamine synthetase mutations. *N. Engl. J. Med.* **2005**, *353*, 1926–1933. [\[CrossRef\]](#)
64. Morris, C.R.; Hamilton-Reeves, J.; Martindale, R.G.; Sarav, M.; Ochoa Gautier, J.B. Acquired Amino Acid Deficiencies: A Focus on Arginine and Glutamine. *Nutr. Clin. Pract. Off. Publ. Am. Soc. Parenter. Enter. Nutr.* **2017**, *32*, 30S–47S. [\[CrossRef\]](#)
65. Viotti, R.; Vigliano, C.; Lococo, B.; Alvarez, M.G.; Petti, M.; Bertocchi, G.; Armenti, A. Side effects of benznidazole as treatment in chronic Chagas disease: Fears and realities. *Expert Rev. Anti Infect. Ther.* **2009**, *7*, 157–163. [\[CrossRef\]](#)
66. Desai, C. Meyler's side effects of drugs: The international encyclopedia of adverse drug reactions and interactions. *Indian J. Pharmacol.* **2016**, *48*, 224.
67. Álvarez, M.G.; Vigliano, C.; Lococo, B.; Bertocchi, G.; Viotti, R. Prevention of congenital Chagas disease by Benznidazole treatment in reproductive-age women. *An observational study. Acta Trop.* **2017**, *174*, 149–152. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Komoto, J.; Yamada, T.; Watanabe, K.; Takusagawa, F. Crystal structure of human prostaglandin F synthase (AKR1C3). *Biochemistry* **2004**, *43*, 2188–2198. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Qin, K.N.; New, M.I.; Cheng, K.C. Molecular cloning of multiple cDNAs encoding human enzymes structurally related to 3 alpha-hydroxysteroid dehydrogenase. *J. Steroid Biochem. Mol. Biol.* **1993**, *46*, 673–679. [\[CrossRef\]](#)
70. Bresson, E.; Boucher-Kovalik, S.; Chapdelaine, P.; Madore, E.; Harvey, N.; Laberge, P.Y.; Leboeuf, M.; Fortier, M.A. The human aldose reductase AKR1B1 qualifies as the primary prostaglandin F synthase in the endometrium. *J. Clin. Endocrinol. Metab.* **2011**, *96*, 210–219. [\[CrossRef\]](#)
71. Kubata, B.K.; Kabututu, Z.; Nozaki, T.; Munday, C.J.; Fukuzumi, S.; Ohkubo, K.; Lazarus, M.; Maruyama, T.; Martin, S.K.; Duszynko, M.; et al. A key role for old yellow enzyme in the metabolism of drugs by *Trypanosoma cruzi*. *J. Exp. Med.* **2002**, *196*, 1241–1251. [\[CrossRef\]](#)
72. Andrade, H.M.; Murta, S.M.F.; Chapeaurouge, A.; Perales, J.; Nirdé, P.; Romanha, A.J. Proteomic analysis of *Trypanosoma cruzi* resistance to Benznidazole. *J. Proteome Res.* **2008**, *7*, 2357–2367. [\[CrossRef\]](#)
73. Murta, S.M.F.; Krieger, M.A.; Montenegro, L.R.; Campos, F.F.M.; Probst, C.M.; Avila, A.R.; Muto, N.H.; de Oliveira, R.C.; Nunes, L.R.; Nirdé, P.; et al. Deletion of copies of the gene encoding old yellow enzyme (TcOYE), a NAD(P)H flavin oxidoreductase, associates with in vitro-induced benznidazole resistance in *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* **2006**, *146*, 151–162. [\[CrossRef\]](#)
74. Arnal, A.; Villanueva-Lizama, L.; Teh-Poot, C.; Herrera, C.; Dumonteil, E. Extent of polymorphism and selection pressure on the *Trypanosoma cruzi* vaccine candidate antigen Tc24. *Evol. Appl.* **2020**, *13*, 2663–2672. [\[CrossRef\]](#)
75. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50*, D439–D444. [\[CrossRef\]](#)
76. Irwin, J.J.; Shoichet, B.K. ZINC—A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* **2005**, *45*, 177–182. [\[CrossRef\]](#)
77. Watts, D.J.; Strogatz, S.H. Collective dynamics of “small-world” networks. *Nature* **1998**, *393*, 440–442. [\[CrossRef\]](#) [\[PubMed\]](#)
78. de Silva, E.; Stumpf, M.P.H. Complex networks and simple models in biology. *J. R. Soc. Interface* **2005**, *2*, 419–430. [\[CrossRef\]](#) [\[PubMed\]](#)
79. Ravasz, E.; Somera, A.L.; Mongru, D.A.; Oltvai, Z.N.; Barabási, A.-L. Hierarchical Organization of Modularity in Metabolic Networks. *Science* **2002**, *297*, 1551–1555. [\[CrossRef\]](#) [\[PubMed\]](#)
80. Barrat, A.; Barthélemy, M.; Pastor-Satorras, R.; Vespignani, A. The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3747–3752. [\[CrossRef\]](#)
81. Assenov, Y.; Ramírez, F.; Schelhorn, S.-E.; Lengauer, T.; Albrecht, M. Computing topological parameters of biological networks. *Bioinform. Oxf. Engl.* **2008**, *24*, 282–284. [\[CrossRef\]](#) [\[PubMed\]](#)
82. Gao, Z.; Li, H.; Zhang, H.; Liu, X.; Kang, L.; Luo, X.; Zhu, W.; Chen, K.; Wang, X.; Jiang, H. PDTD: A web-accessible protein database for drug target identification. *BMC Bioinform.* **2008**, *9*, 104. [\[CrossRef\]](#) [\[PubMed\]](#)
83. Hui-fang, L.; Qing, S.; Jian, Z.; Wei, F. Evaluation of various inverse docking schemes in multiple targets identification. *J. Mol. Graph. Model.* **2010**, *29*, 326–330. [\[CrossRef\]](#)
84. Lee, M.; Kim, D. Large-scale reverse docking profiles and their applications. *BMC Bioinform.* **2012**, *13*, S6. [\[CrossRef\]](#)
85. Hopkins, A.L. Network pharmacology: The next paradigm in drug discovery. *Nat. Chem. Biol.* **2008**, *4*, 682–690. [\[CrossRef\]](#)
86. Rawal, K.; Singh, P.; Sinha, R.; Ridhima; Preeti, P.; Nath, S.K.; Kumari, P.; Sahai, S.; Garg, N.; Sharma, T.; et al. COV-DOCK server: A web server for COVID-19 ligand-target docking. *OSF Preprints* **2021**. [\[CrossRef\]](#)
87. Voss, M.; Lorenz, N.I.; Luger, A.-L.; Steinbach, J.P.; Rieger, J.; Ronellenfitch, M.W. Rescue of 2-Deoxyglucose Side Effects by Ketogenic Diet. *Int. J. Mol. Sci.* **2018**, *19*, 2462. [\[CrossRef\]](#) [\[PubMed\]](#)
88. Gandara, D.R.; Perez, E.A.; Weibe, V.; De Gregorio, M.W. Cisplatin chemoprotection and rescue: Pharmacologic modulation of toxicity. *Semin Oncol* **1991**, *18*, 49–55.

89. Díaz-Chiguer, D.L.; Márquez-Navarro, A.; Noguera-Torres, B.; de la Luz León-Ávila, G.; Pérez-Villanueva, J.; Hernández-Campos, A.; Castillo, R.; Ambrosio, J.R.; Nieto-Meneses, R.; Yépez-Mulia, L.; et al. In vitro and in vivo trypanocidal activity of some benzimidazole derivatives against two strains of *Trypanosoma cruzi*. *Acta Trop.* **2012**, *122*, 108–112. [[CrossRef](#)] [[PubMed](#)]
90. Woolley, D.W. Some Biological Effects Produced By Benzimidazole And Their Reversal By Purines. *J. Biol. Chem.* **1944**, *152*, 225–232. [[CrossRef](#)]
91. de VC Sinatti, V.; Baptista, L.P.R.; Alves-Ferreira, M.; Dardenne, L.; da Silva, J.H.M.; Guimarães, A.C. In silico identification of inhibitors of ribose 5-phosphate isomerase from *Trypanosoma cruzi* using ligand and structure based approaches. *J. Mol. Graph. Model.* **2017**, *77*, 168–180. [[CrossRef](#)] [[PubMed](#)]
92. Loureiro, I.; Faria, J.; Clayton, C.; Macedo-Ribeiro, S.; Santarém, N.; Roy, N.; Cordeiro-da-Siva, A.; Tavares, J. Ribose 5-Phosphate Isomerase B Knockdown Compromises *Trypanosoma brucei* Bloodstream Form Infectivity. *PLoS Negl. Trop. Dis.* **2015**, *9*, e3430. [[CrossRef](#)] [[PubMed](#)]
93. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [[CrossRef](#)] [[PubMed](#)]
94. Faria, J.; Loureiro, I.; Santarém, N.; Cecílio, P.; Macedo-Ribeiro, S.; Tavares, J.; Cordeiro-da-Silva, A. Disclosing the essentiality of ribose-5-phosphate isomerase B in Trypanosomatids. *Sci. Rep.* **2016**, *6*, 26937. [[CrossRef](#)]
95. Bethony, J.M.; Cole, R.N.; Guo, X.; Kamhawi, S.; Lightowers, M.W.; Loukas, A.; Petri, W.; Reed, S.; Valenzuela, J.G.; Hotez, P.J. Vaccines to combat the neglected tropical diseases. *Immunol. Rev.* **2011**, *239*, 237–270. [[CrossRef](#)]
96. Querec, T.D.; Akondy, R.S.; Lee, E.K.; Cao, W.; Nakaya, H.I.; Teuwen, D.; Pirani, A.; Gernert, K.; Deng, J.; Marzolf, B.; et al. Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. *Nat. Immunol.* **2009**, *10*, 116–125. [[CrossRef](#)] [[PubMed](#)]
97. Nakaya, H.I.; Wrammert, J.; Lee, E.K.; Racioppi, L.; Marie-Kunze, S.; Haining, W.N.; Means, A.R.; Kasturi, S.P.; Khan, N.; Li, G.-M.; et al. Systems biology of vaccination for seasonal influenza in humans. *Nat. Immunol.* **2011**, *12*, 786–795. [[CrossRef](#)] [[PubMed](#)]
98. Li, S.; Roupheal, N.; Duraisingham, S.; Romero-Steiner, S.; Presnell, S.; Davis, C.; Schmidt, D.S.; Johnson, S.E.; Milton, A.; Rajam, G.; et al. Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. *Nat. Immunol.* **2014**, *15*, 195–204. [[CrossRef](#)] [[PubMed](#)]
99. Xu, X.; Olson, C.L.; Engman, D.M.; Ames, J.B. NMR structure of the calflagin Tb24 flagellar calcium binding protein of *Trypanosoma brucei*: NMR Structure of Calflagin Tb24. *Protein Sci.* **2012**, *21*, 1942–1947. [[CrossRef](#)] [[PubMed](#)]
100. Versteeg, L.; Adhikari, R.; Poveda, C.; Villar-Mondragon, M.J.; Jones, K.M.; Hotez, P.J.; Bottazzi, M.E.; Tijhaar, E.; Pollet, J. Location and expression kinetics of Tc24 in different life stages of *Trypanosoma cruzi*. *PLoS Negl. Trop. Dis.* **2021**, *15*, e0009689. [[CrossRef](#)]
101. Díaz-Viraqué, F.; Chiribao, M.L.; Trochine, A.; González-Herrera, F.; Castillo, C.; Liempi, A.; Kemmerling, U.; Maya, J.D.; Robello, C. Old Yellow Enzyme from *Trypanosoma cruzi* Exhibits In Vivo Prostaglandin F2 $\alpha$  Synthase Activity and Has a Key Role in Parasite Infection and Drug Susceptibility. *Front. Immunol.* **2018**, *9*, 456. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.