More than one hundred years after it was first discovered by Carlos Ribeiro Justiniano Chagas, the disease bearing his name remains one of the most neglected topical diseases in the world. Trypanosoma cruzi, the causative agent of Chagas disease, is a protozoan parasite which has become unable to synthesize purines and pteridines instead salvaging them from its human host. It is mainly transmitted by the feces of the Kissing bug, but it is also spread by organ transplant, blood transfusion, and from mother to child by breastfeeding. Trypanosoma cruzi has a sophisticated life cycle with four stages each with unique metabolism. Once in the body the metacyclic trypomastigote penetrates a cell and transforms into an amastigote. The amastigote divides by binary fission in cells of infected tissue then transform into a trypomastigote which burst out from cells into the blood stream, the trypomastigote can infect other cells (usually cells of the heart or gastrointestinal tract) and transform back into an intracellular amastigote in a new infection cycle, leading to chronic infection. Benznidazole and nifurtimox have been the only drugs used to treat Chagas disease for 40 years. They work well to treat the infection in the acute stages but their potency diminishes the longer a person has been infected, becoming ineffective at chronic stages of infection. The 2 month duration of treatment, adverse effect rate of 40%, and increase in drug resistance are top reasons why a safer, polypharmacological drug is needed [1]. Currently an estimated 10 million people, mostly in Latin America, are infected with Trypanosoma Cruzi, and it will kill about 50,000 people this year. Although, Chagas was once only seen in South and Central America with the rise in globalization it has spread to other countries and might become a worldwide problem [2]. New drugs with specificity to multiple targets and potency at chronic stages of infection are desperately needed.

The aim of my research is to guide drug design towards a safe polypharmacological drug

for Chagas disease by using enzyme Sterol 14-alpha demethylase (14DM), a validated therapeutic target, as a template to find other therapeutic target proteins. Since there are already multiple compounds which inhibit 14DM, strategic modifications of their structures can lead to an increase in their polypharmacology [3]. I plan to use SMAP, a ligand-binding site comparison tool, and Nimrod, an execution tool used to run parameter sweeps across multiple compute resources, in order to guide drug design towards a polypharmacological safe drug against Chagas disease.

Pharmaceutical companies have long ignored infectious tropical disease, mainly because of monetary reasons. There are multiple verified drug targets in T. *cruzi* such as Cruzain, Sterol C14 alpha-demethylase, lanosterol synthase, 5A- Glyceraldehyde-3-phosphate dehydrogenase which are possible jumping points for polypharmacological drug design [4]. Yet, drug discovery is currently based on the "one gene, one drug, one disease" philosophy, leading to a compound which is created with one specific target, usually a protein, in mind [5]. The idea behind polypharmacology is if instead of designing a drug to bind to the active site of a single therapeutic target it were designed with multiple targets with similar binding sites in mind the drug could decrease the probability of drug resistance and increase in potency. Drug resistance in multi-target drugs is less likely because in order for T. *cruzi* to become resistance to a multi-target drug, it would require mutations in multiple protein targets. While multi-target drugs decrease the susceptibility of resistance they need to be specific enough not to interact with human proteins which could lead to adverse effects. Finding a single drug which can bind to the active site of multiple therapeutic targets would also increase its potency by reducing the amount of vital compounds created by T. *cruzi*. Difficulty arises in finding therapeutically important

proteins with binding sites similar enough to react to the same compound. Looking for proteins with a similar primary structure to the validated therapeutic target is not helpful because the similar primary structures do not reveal much information about the 3D architecture of the binding site. Also, the time it would take to compare the 3D binding sites of all the proteins in the T. *cruzi* and human proteomes to a template binding site one by one would take a large amount of time.

I used a program called SMAP to identify proteins from the T. *cruzi* and human proteome which had similar drug bind site 3D structure to that of Sterol 14- alpha demethylase (14DM), my template protein. In order to run SMAP you need to have the PDB IDs for the proteins you are interested in, meaning they have to have 3D structures available. I choose 14 DM as the template because it was a validated drug targets in T. *cruzi*, had a 3D structure, and was the focus of drug development. SMAP is designed for web accessible 3D ligand-binding site comparison and similarity searching on a structural proteome scale [6]. I used SMAP mainly for its ability to bilaterally compare the drug binding sites of proteins on a 3D platform independent of primary sequence and its ability to produce a raw score, p-value, template and query coverage, Tanimoto coefficient and RMSD. The p-value is an estimate of the statistical significance of the raw score, the template/query coverage is the ratio of overlapped pocked volume in the template and query pockets, the Tanimoto coefficient is a way to calculate the volume similarity between the two pockets [7]. Take together a p-value below 1.0E-04 and a Tanimoto coefficient above 0.5 indicates the chance of a biological meaningful similarity [8]. I accumulated a list of T. *cruzi* proteome PDB IDs from the PDB website, collected about 141 with a search for “Trypanosoma Cruzi” and from a databank on tropicaldisease.org, which contained 3415 T. cruzi model

proteins. After removing repeated PDB IDs I had a list of 3573 model proteins for T. *cruzi*. The total T. *cruzi* proteome contains 2784 proteins [9]. I account for having more model proteins that the amount in the actual proteome by taking into consideration homologous proteins with different PDB IDs. I received a list of 5103 PDB IDs from the human proteome from Dr. Lei Xie. Overall I ran two experiments on Nimrod: one where SMAP compared the T. *cruzi* proteome to the template enzyme 14DM and another experiment where SMAP compared the human proteome to 14DM. Using Nimrod to run the SMAP jobs on the grid greatly reduced the amount of time it would have taken to run the experiments. Nimrod was able to execute SMAP on multiple computer resources, automatically scheduling each comparison, cutting the time it would take to compare a whole proteome tremendously. After the results were produced for the T. *cruzi* proteome experiment the proteins with the most biologically significant SMAP scores were compared to 14DM using the PDB jCE algorithm comparison tool and any with an identity greater than 15% were ignored.

The experiment which compared the template protein Sterol 14-alpha demethylase with the T. *cruzi* proteome yielded five biologically significant results which can be used in polypharmacological drug design. The result with the most biologically significant p-value and Tanimoto coefficient was quinonoid dihydropteridine reductase (qDHPR). qDHPR is a key enzyme required for regeneration and maintenance of tetrahydrobiopterin (BH4) pools. BH4 is an essential cofactor for enzymes of metabolic importance in many eukaryotes. Trypanosomatids are incapable of creating pteridines on their own and must obtain them by salvage pathways from their insect or mammal host. While BH4's exact roles is not completely understood they are known to be of importance for virulence and differentiation in Leishmania [10]. Interestingly the

human homolog of qDHPR had a p-value of 4.41E-05 and Tamnimoto coefficient of 1 signifying greater similarity between the 14DM and human qDHPR binding pockets. Yet humans can also create BH4 through the conversion of 6-pyruvoyl-BH4 to BH4 with sepiapterin reductase [11]. Therefore a compound inhibiting both human and T. *cruzi* qDHPR would cause the most harm to T. *cruzi*.

The similarity between the drug binding sites of 14DM and a T. *cruzi* model of fructose-1,6-Bisphosphate (cytosolic) is another biologically significant result with a larger p-value but smaller tanimoto coefficient. Fructose-1,6-Bisphosphate is a key rate controlling enzyme in gluconeogenesis, the metabolic process for the synthesis of glucose. Gluconeogenesis is an important process in the amastigote stage of T. *cruzi* in particular because once in the amastigote stage T. *cruzi* is dependent on gluconeogenesis for the synthesis of glycoproteins and glycophospholipids [12]. While there is no in vivo evidence for the vital necessity of fructose-1,6-bisphosphate in T. *cruzi,* it is likely that without it the parasite will be at a disadvantage. While the human homolog of fructose-1,6-bisphosphate is important for glucose metabolism it did not appear in the top hits of the 14DM run against the human proteome signifying compounds targeting 14DM would not affect human fructose-1,6-bisphosphate.

Purine synthesis, necessary for the creation of nucleotides,is another synthetic area where T. *cruzi* lacks the enzymatic machinery for de novo synthesis and is dependent on salvage mechanism. T. *cruzi* is highly dependent on the enzyme Inosine-5-monophophate dehydrogenase (IMPDH) for synthesis of purine nucleotides since IMPDH is the enzyme that catalyzes the first committed step in guanosine synthesis, and is required for the irreversible oxidation of insosine monophsthate [13]. Again, the SMAP score of the human IMPDH homolog (p-value of 9.61E-05

and a Tanimoto coefficient of 0,85) signified it had greater pocket site similarity to 14DM than the T. *cruzi* version. While IMPDH remains important in guanosine synthesis in humans there are other pathways for its synthesis such as the conversion of guanine to guanosine by pruine nuceloside phosphorylase. The secondary human purine synthesis pathway provides an alternative route should human IMPDH be inhibited.

While not as biologically relevant as the preceding proteins T. cruzi pteridine reductase 2 (TcPTR2) is of great importance in the accumulation of pteridine from its host [14]. TcPTR2 can reduce only dihydrobiopterin and dihydrofolate and not oxidized pteridines and with no homologous human protein with similar binding sites it a great candidate for consideration in the drug design of a polypharmacological compound.

All trypanosomes possess a unique thiol metabolism in which the mammalian glutathione reductase system is replaced by trypanothione reductase. [15] While trypanothione reductase might have been the least biologically significant protein hit according to the SMAP score, including it as a target for a polypharmacological compound would be beneficial because it is essential for T. cruzi survival and is absent in the human proteome [16].

Of the human proteins with the most similar drug binding site to that of 14DM the top 25 of them were from the cytochrome p450 gene family. Cytochrome p450 enzymes are mainly involved in metabolism of drugs, hormones, and lipids. 11 of the 25 cytochrome enzymes were responsible for drug metabolism, 3 of them were related to vitamin D metabolism, 2 related to nicotine metabolism, 2 metabolized cholesterol, 1 metabolized retinoic acid, another one metabolized androgens, and one metabolized prostacylin. The remaining cytochrome enzymes were theoretical models. Another interesting result came from subunits alpha, beat, delta, and

gama from a model human acetylcholine receptor with SMAP results which were all biologically significant. These results could be used to predict the causes of side effects related to drugs that bind to 14DM. The results can also help guide drug design to prevent 14DM from these possible biologically significant targets.

From my perspective the future of the project would focus on designing drug that would be able to take advantage of the multiple protein targets with similar binding pockets identified by SMAP. I would suggest that the active sites of the top proteins are compared to find what they have in common and use those residues as anchors for drug design. I would also suggest that any compound which is found to bind to multiple drug targets is also used in docking experiments against the top proteins found in the human proteome experiment such as the acetylcholine receptors. I think the docking experiments with the human proteome would help in finding which drug candidates would be likely to cause side effects and could lead to changes in drug design early in the process to avoid adverse side effects.

Works Cited

[1]World Health Organization Chagas Disease (American trypanosomiasis) http://www.who.int/mediacentre/factsheets/fs340/en/index.html
[2]Ibid
[3]Lepesheva, I. Galina , Hargrove, Y. Tatiana , Anderson, Spencer, et al. (June 8, 2010). Structural Insight Into Inhibition of Sterol 14 alpha-demethylase in the Human Pathogen Trypanosoma Cruzi. Journal of Biochemistry. Doi: 10.1074/jbc.M110.133215
[4]Duschak G. Vilma, Couto S. Alicia. An Insight on Targets and Patented Drugs for Chemotherapy of Chagas Disease. Recent Patents on Anti-Infective Drug Discovery, 2007; 2(1): 19-43.
[5]Durrant D. Jacob, Amaro E. Rommie, Xie Lei. Et al. A Multidimensional Strategy to Detect Polypharmacological Targets in the Absence of Structural and sequence Homology. PLoS Computational Biology, 2010; 3-17
[6]Ren Jingyuan, Xie Lei, et al. SMAP-SW: a parallel web service for structural proteome-wide ligand-binding site comparison. Nucleic Acids Research, 2010; 38(w441-w444):1-5
[7]Ibid
[8]Ibid
[9]Attawod, J.A, Weatherly D.B, Minning T.A. The Trypanosoma Cruzi Proteome. Science, 2005; 309: 473-475
[10]Lye, Lon-Fey, Cunningham L., Mark, Beverley M. Stephen. Characterization of quinonoid-Dihydropteridine Reductase (QDPR) from Lower Eukaryote Leishmania major. Journal of Biological Chemistry, 2002; 227 (41): 38245-38253
[11]Choi YK, Kong JS, Park YS. Functional role of sepiapterin reductase in the biosynthesis of tetrahydropteridines in Dictyostelium discoideum Ax2. Biochim Bioshys Acta., 2006; 1760(6): 877-882
[12] Attawod, J.A, Weatherly D.B, Minning T.A. The Trypanosoma Cruzi Proteome. Science, 2005; 309: 473-475
[13]Marr J. Joseph, Docampo Roberto. Chemotherapy for Chagas' Disease: A Perspective of Current Therapy and Considerations for Future Research. Reviews of Infectious Diseases, 1986; 8 (6): 897-899
[14]Schormann Norbert, Pal Biswajit, Senkovich Olga, et al. Crystal Structure of Trypanosoma cruzi pteridines reductase 2 in complex with a substrate and an inhibitor. Journal of Structural Biology, 2005; 152 (1): 64-75

[15]Krauth-Siegel R. Louise, Inhoff Oliver. Parasite-specific trypanothione reductase as a drug target molecule. Parasitol Res, 2003: 90:557-885
[16]Ibid