ISRST ISBN: 978-1-60651-017-9

BCBGC-10

In-Silico Studies of Essential Metabolic Reactions of *Trypanosoma brucei*: Identifying Potential Drug Target*s*

Shalom Nwodo Chinedu
Department of Biological Sciences, College of Science & Technology, Covenant University, PMB 1023 Ota, Ogun State, Nigeria.
E-mail: sncresearch@gmail.com

Segun A. Fatumo
Department of Computer & Information Sciences, College of Science & Technology, Covenant University, PMB 1023 Ota, Ogun State, Nigeria.
E-mail: segunfatumo@yahoo.co.uk

Ezekiel F. Adebiyi
Department of Computer & Information Sciences, College of Science & Technology, Covenant University, PMB 1023 Ota, Ogun State, Nigeria.
E-mail: eadebiyi@sdsc.edu

Abstract

The aim of this work is to analyze the metabolic pathways of Trypanosoma brucei and identify essential reactions that may be considered as drug targets. Data from the metabolic reaction database of TrypanoCyc version10.0.0 was used to establish a connected graph. The analysis of the biochemical network is based on its topology. The metabolic network was implemented with Perl. Two reactions were defined as neighbours if a metabolite that is the product of one reaction and substrate of the other exists. This yielded a bipartite graph of alternating reaction and metabolic compound nodes. A graph based algorithms were used to analyze the structure of the biochemical networks to infer differences when exposed to changing nutrient and environmental conditions. Choke-points and load-points were used to estimate if reactions are essential for the organism. This produced a network of 809 metabolites and 798 reactions. With this strategy, we have identified 99 essential enzymatic reactions. These reactions can serve as drug targets to inhibit a normal metabolic flow in the parasite without harming the host. Further work is required to design drugs that can inhibit the targets.

Introduction

Human African trypanosomiasis (HAT), also known as sleeping sickness [1], is caused by species of *Trypanosoma brucei*. There are two forms of the disease: the acute form caused by *Trypanosoma brucei rhodesiense* which occurs mainly in East Africa and the chronic form caused by *Trypanosoma brucei gambiense* which arises mainly in West and Central Africa. Although their clinical infections differ in presentation and prognosis [2], the two protozoan parasites have identical morphologic appearances and are transmitted by tsetse fly. WHO estimates that up to 500,000 individuals are currently suffering from human trypanosomiasis in Africa, most of who are infected by *Trypanosoma brucei gambiense* [3].

There have been three severe epidemics of HAT over the last century: one between 1896 and 1906, mostly in Uganda and the Congo Basin, one in 1920 in several African countries and the one that began in 1970 and is still in progress. The 1920 epidemic was arrested using mobile teams which systematically screened millions of 'at-risk' individuals. The disease practically disappeared between 1960 and 1965, but reappeared in several foci in the 1970s due to relaxed screening and lack of effective surveillance [4]. The resurgence has been attributed to civil wars, population movements, economic decline, reduced health financing, and lack of human resources in these areas [5-7]. Today, HAT occurs in 36 sub-Saharan countries within the distribution of the tsetse fly. There are about 250 foci of disease transmission and over 60 million people are at risk of contracting the disease [8]. There is no vaccine for HAT yet, and the

few drugs available are becoming ineffective due to parasite resistance. This underscores the need for new effective chemotherapeutic agents against the parasites.

The delineation of the genome of *T. brucei* has raised the hope for new drug targets and vaccines [9]. Though several proteins have been identified as potential targets for drug treatment, *T. brucei* has over 800 genes that code for proteins which the parasite mixes and matches to evade immune system detection; this makes the production of vaccines for the disease difficult [10]. Identification of novel drug targets is required to develop new classes of drugs in order to overcome drug resistance and replace less efficacious treatments. Analysis of metabolic pathways provides a useful conceptual framework for the identification of potential drug targets and also for improving our understanding of microbial responses to nutritional, chemical and other environmental stresses. A number of metabolic databases are available as tools for such analyses. In this work, data from the metabolic reaction database of *TrypanoCyc version10.0.0* [11] was used to identify reactions which can serve as drug targets by inhibiting a normal metabolic flow in the parasite without harming the host.

Material and Methods

Metabolic Network Construction

Data from the metabolic reactions database of *TrypanoCyc version10.0.0.* [11] was used for the analysis. A connected graph was established by defining neighbours of reactions: two reactions are neighbours if a metabolite that is the product of one reaction and the substrate for the other exists. This yields a bipartite graph of alternating reaction and metabolic compound nodes. A description of metabolic network is given in equation 1-4 and figure 1a and 1b. Metabolites that were highly connected and therefore pathway unspecific, such as water, oxygen and ATP, were discarded. The metabolic network was implemented with Perl.

Topology Analysis to Identify Essential Reactions

We analyzed the biochemical network of *Trypanosoma brucei* based on its topology. A graph based algorithms of analysing the structure of biochemical networks was employed to infer differences when exposed to changing nutrients and environmental conditions and Reaction without deviation (RWD) analysis was used to identify potential drug targets [12]. The strategy was used to test the reactions coming out of the gene expression analysis on serving as crucial choke and load points for the organism while being harmless to the human. These reactions can serve as drug targets to inhibit a normal metabolic flow in the parasite without harming the host. Furthermore, concepts of choke-points and load-points were used to estimate if reactions are essential for the organism [12, 13]. Choke-points basically and uniquely consume or produce a certain metabolite which may make them indispensable. It could be shown that inactivating choke-points lead to an organism's failure. With a choke-point analysis [13], enzymes catalyzing choke-point reactions were identified; each enzyme was assumed to have only one active site, unless annotated as multifunctional. If an enzyme catalyzed at least one choke-point reaction, it was classified as a potential drug target. The two concepts, RWD and choke-points, were implemented.

Drug Targets in African Trypanosomes

Several possible targets which could be exploited for the development of new drugs against African trypanosomiasis have been identified. Some drug targets against HAT are available at the TDR Targets Database [14]. The website aims to capture, collate and make public available expert knowledge on potential drug targets against parasitic diseases. Fifty of such targets are listed in Table 1.

Given a set of reactions in a pathway:

- *R1:* $1\ G6P + 1\ NADP \xrightarrow{ZWF} 1\ 6PGL + 1\ NADPH$ ……..*equation 1*
- *R2:* $1\ 6PGL + 1\ H20 \xrightarrow{PGL} 1\ 6PG$ …….. *equation 2*
- *R3:* $1\ 6PGL + 1\ H2O \xrightarrow{GND} 1\ R5P + 1\ NADPH$ …….. *equation 3*
- *R4:* $1\ R5P \xrightarrow{RPE} 1\ X5P$ ……..*equation 4*

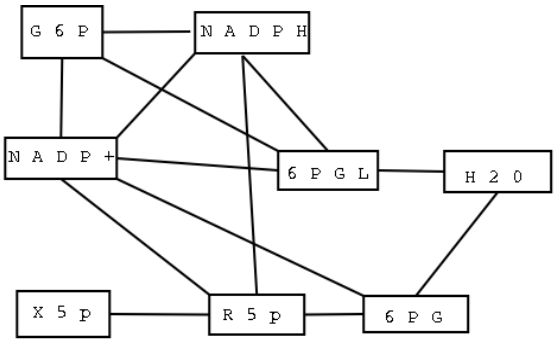

Figure 1a: Substrate graph

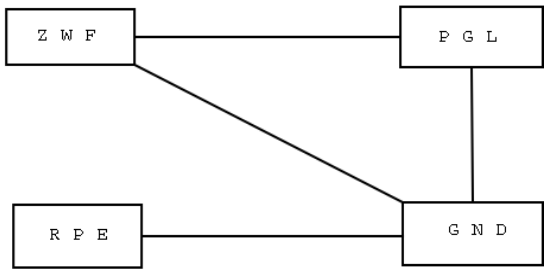

Figure 1b: Reaction graph

Table 1: Drug Targets against African Trypanosomes

S/N	EC Number	Enzyme	Source Database / ID:	Target Name
1.	EC 2.7.1.67	Phosphatidylinositol 4-kinase	Tb927.4.1140	Phosphatidyl-inositol
2.	EC 6.3.2.2	Gamma-glutamylcysteine synthetase	Tb10.389.1360	Glutamate-cysteine
3.	EC 4.1.1.50	S-adenosylmethionine decarboxylase proenzyme	GeneDB / Tb927.6.4460	(AdoMetDC)
4.	EC 3.6.3.14	ATP synthase complex	Tb927.3.1380; Tb927.7.7420.	ATP synthase
5.	EC 2.7.1.37	cell division related protein kinase 2 (CRK3:CYC6)	Tb10.70.2210.	CRK3
6.	EC 2.5.1.18	GPI:protein transamidase complex	Tb10.61.3060	GPI8
7.	EC 6.3.4.2	Trypanosome CTP Synthetase	Tb927.1.1240	Trypanosome CTP Synthetase
8.	EC 2.7.11.1	Casein kinase 1.2	Tb927.5.800; Tb927.5.790.	CK1.2
9.	N/A	Peroxin 14 (PEX14)	Tb10.100.0130	PEX14
10.	N/A	Peroxin 5 (PEX5)	Tb927.5.1100.	PEX5
11.	EC 2.7.1.40	Pyruvate kinase (PYK)	Tb10.61.2680	PYK
12.	EC 5.4.2.1	Phosphoglycerate mutase (PGAM)	Tb10.6k15.2620	PGAM
13.	EC 2.7.1.11	Phosphofructokinase (PFK)	Tb927.3.3270	PFK
14.	EC 4.2.1.11	Enolase (ENO)	Tb10.70.4740	ENO
15.	EC 4.1.2.13	Fructose-1,6-bisphosphate aldolase (ALD)	Tb10.70.1370	ALD
16.	EC 2.7.2.3	Phosphoglycerate kinase (PGK)	Tb927.1.700.	PGK-C or gPGK
17.	EC 1.1.1.49	Glucose-6-phosphate dehydrogenase (G6PDH)	Tb10.70.5200	G6PDH
18.	N/A	Vps34	Tb927.8.6210	TbVPS34
19.	EC 1.1.1.8	NAD-dependent glycerol-3-phosphatedehydrogenase (NAD-GPDH)	Tb927.8.3530	NAD-GPDH
20.	EC 5.3.1.9	Glucose-6-phosphate isomerase (phosphoglucose isomerase - PGI)	Tb927.1.3830	PGI

21.	EC 5.3.2.1	Triosephosphate isomerase (TIM)	Tb11.02.3210	TIM
22.	EC 1.2.1.12	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	Tb927.6.4280Tb927.6.4300	gGAP
23.	N/A	5'-3' exonuclease (XRNA)	Tb927.7.4900	XRNA
24.	N/A	Pumilio domain protein (PUF9)	Tb927.1.2600	PUF9
25.	N/A	Small RNA binding proteins (UBP1 and 2)	Tb11.03.0620Tb11.03.0580	UBP1 and 2
26.	N/A	Nucleolar protein (DRBD1)	Tb927.3.5280	DRBD1
27.	N/A	Cyclin F box protein (CFB2)	Tb927.1.4650	CFB2
28.	N/A	MRNA degradation CAF1 deadenylase	Tb927.6.600	CAF1
29.	N/A	Alternative oxidase	Tb10.6k15.3640	N/A
30.	EC 2.7.1.37	Glycogen synthase kinase (GSK)	Tb10.61.3140	GSK
31.	N/A	Aurora kinase 1	Tb11.01.0330	TbAUK1
32.	EC 4.1.1.17	Ornithine decarboxylase (ODC)	Tb11.01.5300	ODC
33.	EC 2.7.1.-	Polo-like kinase (PLK)	Tb927.7.6310	PLK
34.	N/A	Histone deacetylases (DAC1 and DAC3)	Tb10.70.6220Tb927.2.2190	DAC1 & DAC3
35.	EC 1.11.1.9	Non-selenium glutathione peroxidase (TbTPNI)	Tb09.160.4250 Tb09.160.4280	TbTPNI
36.	EC 1.15.1.1	Superoxide dismutases (SODs)	Tb11.01.7550Tb11.01.6660	TbSODB2
37.	EC 6.3.1.9 / EC 3.5.1.78	trypanothione synthetase / amidase	Tb927.2.4370	trypanothione synthetase / amidase
38.	EC 1.8.1.12	Trypanothione reductase	Tb927.2.4370	trypanothione reductase
39.	N/A	Membrane-bound histidine acid phosphatase1	Tb11.01.4701	TbMBAP1
40.	N/A	Premature differentiation to the stumpy form.	Tb927.6.4220	TbMAPK5
41.	EC 3.1.4.17	Cyclic nucleotide specific phosphodiesterases (PDEs)	Tb09.160.3630 Tb09.160.3590	Phosphodiesterase
42.	EC 2.7.1.1	Hexokinase (HK)	Tb10.70.5820	TbHK1
43.	N/A	Metacaspases	Tb927.6.940	MCA2
44.	EC 5.1.3.2	UDP-Glc 4'-Epimerase		UDP-Glc 4'-Epimerase
45.	N/A	N-myristoyltransferase	Tb10.61.2550	N-myristoyl transferase,
46.	EC 2.7.7.23	UDP-GlcNAc diphosphorylase	Tb11.02.0120	UDP-GlcNAc diphosphorylase
47.	N/A	Kinetoplatstid RNA editing ligase 1	Tb09.160.2970	Kinetoplatstid RNA editing ligase
48.	N/A	Oligosaccharyl transferase (OST)	N/A	OST
49.	EC 2.5.1.58	Protein farnesyltransferase (PFT)	Tb927.3.4490Tb927.7.460	Protein farnesyltransferase
50.	EC 3.5.1.89	GlcNAc-PI de-N-acetylase	Tb11.01.3900	GlcNAc-PI de-N-acetylase

Targets lacking necessary information such as EC number (in red colour) were not considered in the analysis.

Results and Discussion

A connected graph was established by defining neighbours of reactions yielded a bipartite graph of alternating reaction and metabolic compound nodes. This produced a network of 809 metabolites and 798 reactions. Each reaction was considered reversible since there was no contrary information available. Note that these reactions can serve as drug targets to inhibit a normal metabolic flow in the parasite without harming the host. With this strategy, we have identified 99 essential enzymatic reactions which can serve as drug targets (Table 2).

In medical diagnostic test evaluation, more common metrics for evaluation are *sensitivity* and *specificity*. *Sensitivity* is the accuracy among positive instances and *specificity* among negative. S*ensitivity* and *specificity* overcome negative sides of accuracy (error type resolution and condition prevalence). Accuracy reflects the overall correctness of the classifier.

Essential	6	93	99
Non-essential	28	671	699
Total	34	764	798

Accuracy = 6+671/798 = 84.8%

Sensitivity = 6/34 = 17.6%

Specificity = 671/764 = 87.8%

Precision = 6/6+93 = 6/99 = 6.06%

$$Recall = \frac{TP}{TP + FN} = Sensitivity = True\ Positive\ Rate$$

$$Precision = \frac{TP}{TP + FP}$$

A graph based algorithms of analysing the structure of biochemical networks to infer differences when exposed to changing nutrients and environmental conditions was employed for *Plasmodium falciparum* to test the reactions coming out of the gene expression analysis on serving as crucial choke and load points for the organism while being harmless to the human [12]. These reactions can serve as drug targets to inhibit a normal metabolic flow in the parasite without harming the host. Furthermore, concepts of choke-points and load-points were used to estimate if reactions are essential for the organisms [13]. Choke-points basically uniquely consume or produce a certain metabolite which may make them indispensable. It has been shown that inactivating choke-points lead to an organism's failure. For example, in *P. falciparum* d-aminolevulinate dehydratase (ALAD) has been considered as such a choke-point [13]. With a choke-point analysis, 216 enzymes catalysing choke-point reactions were identified *P. falciparum* [13]. Within the 216 identified potential targets, they identified three targets of clinically proven drugs and 24 proposed drug targets with biological evidence (such as *in vitro* growth inhibition of the parasite with target inhibition). Hence, with the combination of RWD and chokepoint analyses, we can confidently conclude that we have identified 99 essential enzymatic reactions which can serve as drug targets.

Table 2: Reaction without Deviation (RWD) and Chokepoint Analyses

Method	***Essential reactions***
Reaction without Deviation (RWD)	130
Chokepoint	211
Intercession (RWD and Chokepoint)	99

References

[1] Kennedy, PG. Human African trypanosomiasis of the CNS: current issues and challenges. J Clin Invest; 2004; 113:496.

[2] Barrett, MP, Burchmore, RJ, Stich, A. The trypanosomiasis. Lancet; 2003; 362:1469.

[3] WHO Media centre. *Fact sheet N°259: African trypanosomiasis or sleeping sickness*. 2006; http://www.who.int/mediacentre/factsheets/fs259/en/

[4] Smith, DH, Pepin, J, Stich, AH. Human African trypanosomiasis: an emerging public health crisis. Br Med Bull; 1988; 54:341.

[5] Mhlanga, JD. Sleeping sickness: perspectives in African trypanosomiasis. Sci Prog; 1996; 79 (3): 183.

[6] Stich, A, Abel, PM, Krishna, S. Human African trypanosomiasis. BMJ; 2002; 325:203.

[7] Barrett, MP. The rise and fall of sleeping sickness. Lancet; 2006; 367:1377.

[8] WHO Scientific Working Group Report on African trypanosomiasis (sleeping sickness) June 2001. www.who.int/dir.WHO. Geneva c2003.

[9] Bloom S. Parasite genome similarities offer hope for new drugs and vaccines. J Clin Invest.;2005; 115 (9):2300-1.

[10] Berriman, M, Ghedin, E, Hertz-Fowler, C. The genome of the African trypanosome Trypanosoma brucei. Science; 2005; 309:416 -22.

[11] Chukualim B., Peters N., Fowler C. H. and Berriman M. TrypanoCyc – a metabolic pathway database for *Trypanosoma brucei*. *BMC Bioinformatics 2008;* 9 (Suppl 10): 5 doi:10.1186/1471-2105.

[12] Fatumo S, Plaimas K, Mallm JP, Schramm G, Adebiyi E, Oswald M, Eils R, König R. Estimating novel potential drug targets of Plasmodium falciparum by analysing the metabolic network of knock-out strains in silico. Infect Genet Evol.;2009; 9(3):351-8.

[13] Yeh RK, Chen J, Williams JL, Baluch M, Hundley TR, Rosenbaum RE, Kalala S, Traganos F, Benardini F, del Soldato P, Kashfi K, Rigas B. NO-donating nonsteroidal antiinflammatory drugs (NSAIDs) inhibit colon cancer cell growth more potently than traditional NSAIDs: a general pharmacological property? Biochem Pharmacol. 2009; 67(12):2197-205.

[14] The TDR Targets Database: Identification and ranking of targets against neglected tropical diseases. 2010; http://tdrtargets.org/survey