Systems Analysis of Metabolism in the Pathogenic Trypanosomatid *Leishmania major*

Arvind K. Chavali, Jeffrey D. Whittemore, James A. Eddy, Kyle T. Williams
and Jason A. Papin

Department of Biomedical Engineering
University of Virginia
Charlottesville, VA 22908
United States of America

Supplementary material I

TABLE OF CONTENTS

Figure S1

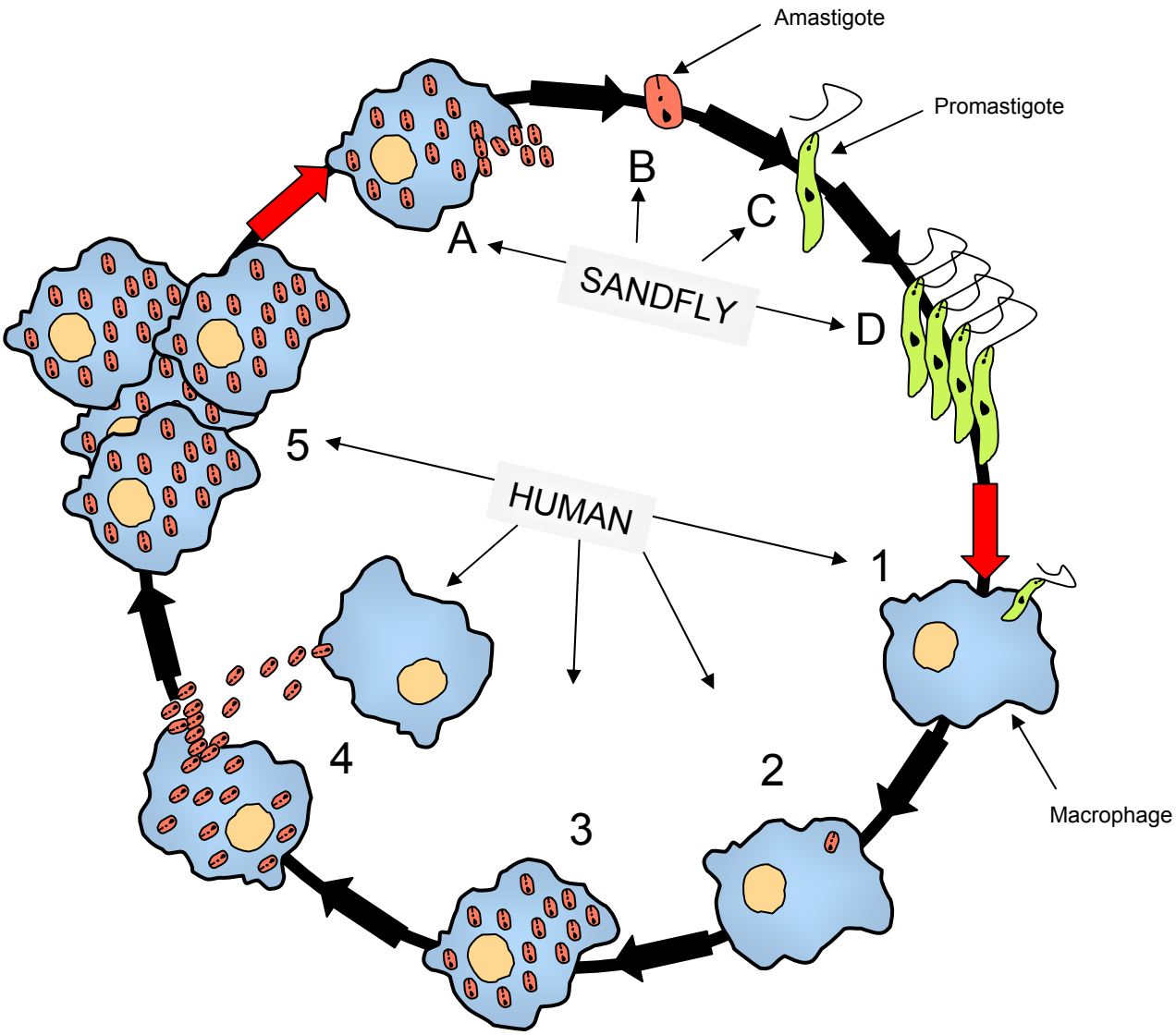

Figure S1: Life-cycle of Leishmania parasites
The lifecycle of *Leishmania* spp. is presented. Specifically, when a sandfly bites to take a blood meal, promastigotes are transferred to the human host (1). Subsequently, macrophages phagocytose promastigotes, which undergo a morphological change into amastigotes (2). As amastigotes, *Leishmania* replicate inside macrophages (3) and with sufficient numbers can lyse host macrophages and spread infection (4, 5). When another sandfly takes a blood meal from an infected human host, infected macrophages containing amastigotes are taken up by the sandfly (A). Amastigotes (B) transform into promastigotes (C, D) and the life-cycle continues.

Figure S2

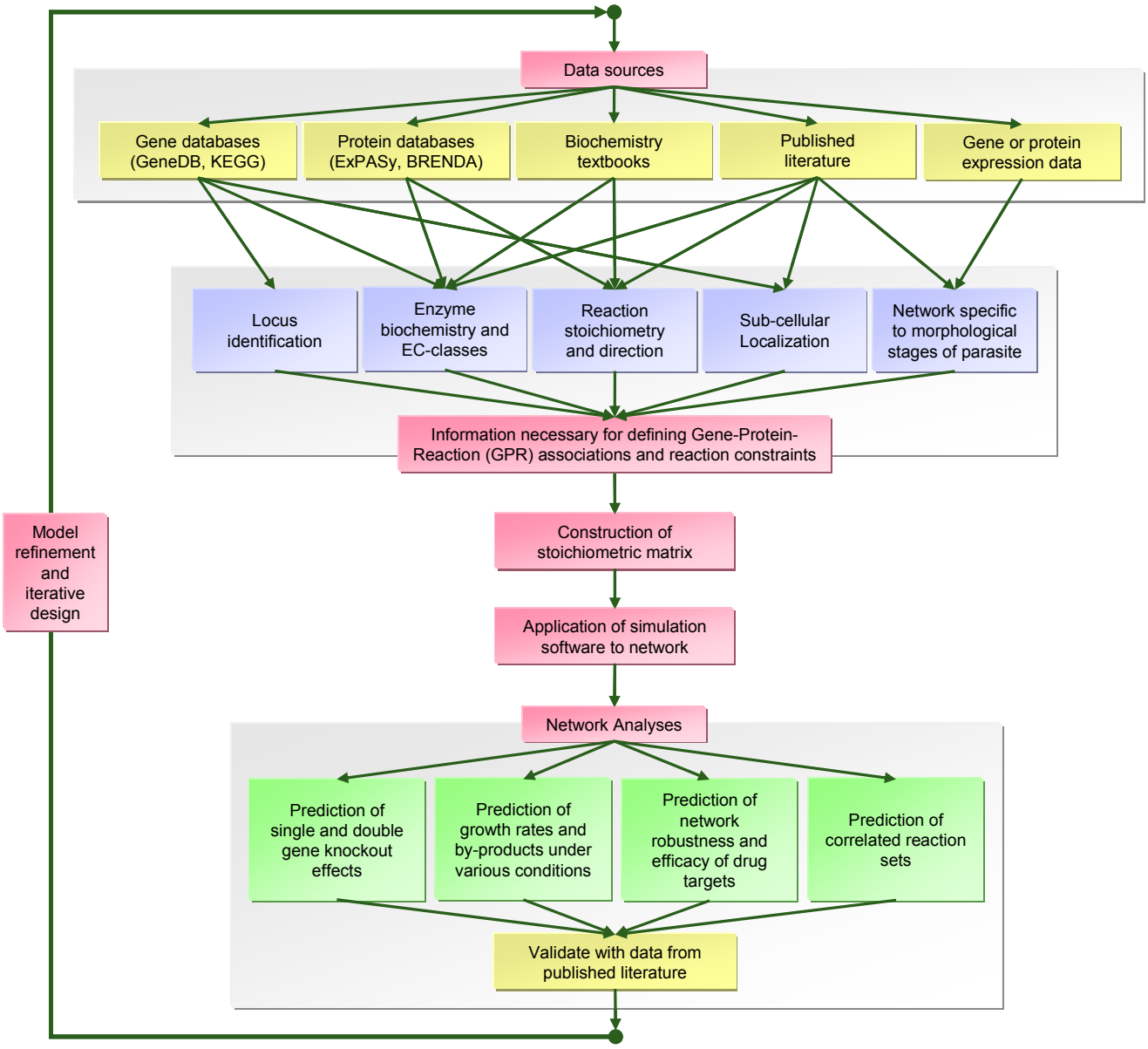

Figure S2: Process of reconstructing a metabolic network
The metabolic reconstruction process required assimilating information from gene/protein databases, biochemistry textbooks, published literature and expression data. This compilation of data is used to construct gene-protein-reaction (GPR) relationships. Subsequently, a stoichiometric matrix (S-matrix) consisting of the stoichiometric coefficients of metabolites participating in all the reactions within the network is assembled. Simulation software is applied to calculate the flux distribution. Several types of network analyses can be performed such as gene essentiality studies, growth rate predictions and robustness analysis. The model is iteratively refined such that predictions match data from experimental studies.

BIOMASS CALCULATIONS FOR *LEISHMANIA MAJOR*

To generate a biomass demand reaction for the metabolic network of *Leishmania major*, the cellular composition was estimated as in Table 1.

Component	% dry weight
Protein	45
DNA	1.6
RNA	11
Carbohydrate	27
Lipid	15
Polyamine pools	0.4
Total	**100**

Table 1: L. major cellular composition (estimated)

Specifically, fractions of protein, RNA, and carbohydrate were estimated from the protozoan *Tetrahymena* (Gates *et al*, 1982; Hellung-Larsen & Andersen, 1989), the most related-species for which such data was available. Carbohydrates were split into two categories: mannan and other carbohydrates. Mannan comprises approximately 90% of cellular carbohydrate in *L. mexicana* (Ralton *et al*, 2003). Lipid content was estimated from related *Leishmania* promastigotes (*L. enrietti* and *L. donovani*) (Beach *et al*, 1979; Glew *et al*, 1988). Percent dry weight of polyamine pools, specifically putrescine and spermidine, was estimated from the metabolic network reconstructions of *Helicobacter pylori* and *Methanosarcina barkeri* (Feist *et al*, 2006; Thiele *et al*, 2005). The remaining composition was assumed to be DNA. The contribution of free amino acids was assumed to be insignificant.

PROTEIN CONTENT

To determine the relative abundance of amino acids (AA) in protein (see Table 2), the AA sequence associated with every ORF in the *L. major* genome was downloaded from GeneDB (http://www.genedb.org/genedb/leish/), the numbers of each AA in these ORFs were counted, and the percentage prevalence of each AA was calculated. As an example, consider the calculation of the cellular content of arginine (in mmol per gram of dry cell weight):

$$\text{Arginine (mmol/gDW)} = \frac{0.0995 \text{ gms of Arginine}}{1 \text{ g of Protein}} * \frac{0.45 \text{ gms of Protein}}{1 \text{ gDW}} * \frac{1 \text{ mole Arginine}}{175.11 \text{ gms of Arginine}} * \frac{1000 \text{ mmol}}{1 \text{ mole}} = 0.2557$$

Amino Acid	% Prevalence	MW (g/mol)	% (by weight)	mmol/gDW
Alanine (A)	12.04	89.05	8.47	0.4281
Arginine (R)	7.19	175.11	9.95	0.2557
Asparagine (N)	2.62	132.05	2.73	0.0932
Aspartic acid (D)	4.86	132.04	5.07	0.1728
Cysteine (C)	1.89	121.02	1.81	0.0672

Glutamate (E)	6.00	146.05	6.92	0.2134
Glutamine (Q)	4.10	146.07	4.73	0.1458
Glycine (G)	6.46	75.03	3.83	0.2297
Histidine (H)	2.70	155.07	3.31	0.0960
Isoleucine (I)	2.98	131.09	3.09	0.1060
Leucine (L)	9.20	131.09	9.53	0.3271
Lysine (K)	3.33	147.11	3.87	0.1184
Methionine (M)	2.26	149.05	2.66	0.0804
Phenylalanine (F)	2.95	165.08	3.85	0.1049
Proline (P)	5.79	115.06	5.26	0.2059
Serine (S)	8.98	105.04	7.45	0.3193
Threonine (T)	6.00	119.06	5.64	0.2134
Tryptophan (W)	1.08	204.09	1.74	0.0384
Tyrosine (Y)	2.40	181.07	3.43	0.0853
Valine (V)	7.17	117.08	6.63	0.2550

Table 2: Amino acid contribution to biomass

DNA CONTENT

The percent prevalence of the four nucleotides in DNA, shown in Table 3, was calculated assuming a G+C content of 59.7% (Ivens *et al*, 2005). The equation to calculate the cellular content (in mmol/gDW) for each deoxynucleotide was the same as for AA contribution.

DNA	% Prevalence	MW (g/mol)	% (by weight)	mmol/gDW
dAMP	20.15	329.07	20.41	0.0099
dCMP	29.85	305.06	28.03	0.0147
dGMP	29.85	345.06	31.71	0.0147
dTMP	20.15	320.06	19.85	0.0099

Table 3: Deoxynucleotide contribution to biomass

RNA CONTENT

The RNA monomer abundance, shown in Table 4, was determined in the same way as AA abundance described above. However, in this instance, the DNA sequence associated with each ORF in the *L. major* genome was downloaded from GeneDB. The equation to calculate the cellular content (in mmol/gDW) for each ribonucleotide was the same as for AA and DNA.

RNA	% Prevalence	MW (g/mol)	% (by weight)	mmol/gDW
AMP	19.32	345.06	19.70	0.0628
CMP	30.94	321.05	29.35	0.1006
GMP	31.49	361.06	33.59	0.1023
UMP	18.25	322.04	17.36	0.0593

Table 4: Ribonucleotide contribution to biomass

CARBOHYDRATE CONTENT

The molecular weight (MW) of mannan was approximated as 70kDa (see Table 5) (Roy *et al*, 1998). Since mannan accounts for the majority of carbohydrate content in *Leishmania* (Ralton *et al*, 2003), the contribution of other carbohydrates to biomass was assumed to be insignificant.

Carbohydrate	% dry weight	MW (g/mol)	mmol/gDW
Mannan	24	70000	0.0034

Table 5: Mannan contribution to biomass

FATTY ACID CONTENT

The fatty acid composition of *L. major* promastigotes (see Table 6) was obtained from (Vessal *et al*, 1974). Those fatty acids with less that 2% overall contribution were omitted, and the percentages were normalized. The MW's represent the weights of atoms in the alphatic chain excluding the carboxylic acid end group. Atomic masses (in atomic mass units, amu, or Daltons, Da) of 12.011 for carbon and 1.00794 for hydrogen were used (http://www.nist.gov/). The carboxylic acid end group was included in the formula for the compound to which the fatty acid chain was being attached. For example, for fatty acid 14:0, the MW was determined as follows:

$$\mathrm{MW}\ 14:0 = 13*12.011 + 27*1.00794 = 183.36\ \mathrm{Da}$$

Fatty acid	% Prevalence (Normalized)	MW (g/mol)
14:0	3.10	183.36
16:0	8.68	211.41
18:0	9.92	239.46
18:1	33.37	237.45
18:2	28.82	235.43
18:3	16.11	233.42
Weighted MW (g/mol)		**232.48**

Table 6: Fatty acid composition of L. major promastigotes (Vessal et al, 1974)

LIPID CONTENT

The estimated lipid composition for *Leishmania* is shown in Table 7.

Lipid	% (by weight)
Neutral	34.5
Polar	65.5

Table 7: Estimated lipid composition for Leishmania spp. (Beach et al, 1979; Glew et al, 1988)

The MW's of neutral lipids (Table 8), namely monoacylglycerol, diacylglycerol and triacylglycerol, were determined using the weighted fatty acid MW (see Table 6). For example, diacylglyc-

erol can be written as $C_5H_6O_5$ with 2 fatty acid chains, and therefore, using atomic masses of 12.011 for carbon, 1.00794 for hydrogen, and 15.9994 for oxygen (http://www.nist.gov/), its MW was computed as follows:

MW of diacylglycerol = 146.1 + 2*232.48 = 611.06 Da

Neutral Lipid	% (by weight)	MW (g/mol)	mmol/gDW
Sterols (ergosterol)	43.3	396.34	0.0565
Triacylglycerol (triglyceride)	43.0	870.54	0.0256
Sterol esters (zymosterol)	9.1	384.34	0.0123
Diacylglycerol	1.5	611.06	0.0013
Monoacylglycerol	1.0	351.58	0.0015

Table 8: Neutral lipid composition of Leishmania spp. (Beach et al, 1979; Glew et al, 1988)

The MW's of phospholipids were similarly determined and are presented in Table 9. For some phospholipids, additional atomic masses of 14.00674 for nitrogen and 30.973762 for phosphorous were used (http://www.nist.gov/). The neutral and polar lipid compositions were obtained for *L. tarentolae* (Beach *et al*, 1979; Glew *et al*, 1988).

Polar Lipid	% (by weight)	MW (g/mol)	mmol/gDW
Phosphatidylethanolamine	20	734.11	0.0268
Phosphatidylcholine	49	776.19	0.0620
Phosphatidylinositol	8	852.18	0.0092
Diphosphatidylglycerol (cardiolipin)	3	1436.13	0.0021

Table 9: Polar lipid composition of Leishmania spp. (Beach et al, 1979; Glew et al, 1988)

POLYAMINE CONTENT

As shown in Table 10, the amounts of putrescine and spermidine were estimated as 0.3 and 0.1 percent dry weight respectively from the metabolic reconstructions of *Helicobacter pylori* and *Methanosarcina barkeri* (Feist *et al*, 2006; Thiele *et al*, 2005).

Polyamine	% dry weight	MW (g/mol)	mmol/gDW
Putrescine	0.3	90.1	0.0333
Spermidine	0.1	148.16	0.0067

Table 10: Polyamine contribution to biomass (estimated)

ATP MAINTENANCE

From (ter Kuile & Opperdoes, 1992), the growth associated ATP maintenance was estimated as 32.26 mmol ATP per 1 gram of dry cell weight. Therefore, the coefficients for metabolites in the ATP maintenance reaction are as shown in Table 11.

Metabolite	mmol/gDW
ATP	32.26
H_2O	32.26
ADP (prod)	32.26
H (prod)	32.26
Pi (prod)	32.26

Table 11: Coefficients for metabolites in the ATP maintenance reaction

FINAL BIOMASS EQUATION

(0.4281) ala-L + (0.2557) arg-L + (0.0932) asn-L + (0.1728) asp-L + (0.0672) cys-L + (0.1458) gln-L + (0.2134) glu-L + (0.2297) gly + (0.0960) his-L + (0.1060) ile-L + (0.3271) leu-L + (0.1184) lys-L + (0.0804) met-L + (0.1049) phe-L + (0.2059) pro-L + (0.3193) ser-L + (0.2134) thr-L + (0.0384) trp-L + (0.0853) tyr-L + (0.2550) val-L + (0.0099) damp + (0.0147) dcmp + (0.0147) dgmp + (0.0099) dtmp + (0.0628) amp + (0.1006) cmp + (0.1023) gmp + (0.0593) ump + (0.0034) mannan + (0.0565) ergst + (0.0256) triglyc_LM + (0.0123) zymst + (0.0013) 12dgr_LM + (0.0015) mag_LM + (0.0268) pe_LM + (0.0620) pc_LM + (0.0092) ptd1ino_LM + (0.0021) clpn_LM + (0.0333) ptrc + (0.0067) spmd + (32.26) atp + (32.26) h2o --> (32.26) adp + (32.26) h + (32.26) pi

Note: All MW's listed are for the charged form of a compound. Unless estimated, the MW's were obtained from KEGG database (http://www.genome.jp/kgg/). The method specified by Forster et al. was used in the biomass estimation for *L. major* (Forster *et al*, 2003).

REFERENCES

Beach DH, Holz GG, Jr., Anekwe GE (1979) Lipids of Leishmania promastigotes. *J Parasitol* **65:** 201-216.

Feist AM, Scholten JC, Palsson BO, Brockman FJ, Ideker T (2006) Modeling methanogenesis with a genome-scale metabolic reconstruction of Methanosarcina barkeri. *Mol Syst Biol* **2:** 0004.

Forster J, Famili I, Fu P, Palsson BO, Nielsen J (2003) Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. *Genome Res* **13:** 244-253.

Gates MA, Rogerson A, Berger J (1982) Dry to Wet Weight Biomass Conversion Constant for Tetrahymena-Elliotti (Ciliophora, Protozoa). *Oecologia* **55:** 145-148.

Glew RH, Saha AK, Das S, Remaley AT (1988) Biochemistry of the Leishmania species. *Microbiol Rev* **52:** 412-432.

Hellung-Larsen P, Andersen AP (1989) Cell volume and dry weight of cultured Tetrahymena. *J Cell Sci* **92 (Pt 2):** 319-324.

Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E, Aert R, Anupama A, Apostolou Z, Attipoe P, Bason N, Bauser C, Beck A, Beverley SM, Bianchettin G, Borzym K, Bothe G*, et al.* (2005) The genome of the kinetoplastid parasite, Leishmania major. *Science* **309:** 436-442.

Ralton JE, Naderer T, Piraino HL, Bashtannyk TA, Callaghan JM, McConville MJ (2003) Evidence that intracellular beta1-2 mannan is a virulence factor in Leishmania parasites. *J Biol Chem* **278:** 40757-40763.

Roy SK, Yoko-o T, Ikenaga H, Jigami Y (1998) Functional evidence for UDP-galactose transporter in Saccharomyces cerevisiae through the in vivo galactosylation and in vitro transport assay. *J Biol Chem* **273:** 2583-2590.

ter Kuile BH, Opperdoes FR (1992) Comparative physiology of two protozoan parasites, Leishmania donovani and Trypanosoma brucei, grown in chemostats. *J Bacteriol* **174:** 2929-2934.

Thiele I, Vo TD, Price ND, Palsson BO (2005) Expanded metabolic reconstruction of Helicobacter pylori (iIT341 GSM/GPR): an in silico genome-scale characterization of single- and double-deletion mutants. *J Bacteriol* **187:** 5818-5830.

Vessal M, Rezai HR, Pakzad P (1974) Leishmania species: fatty acid composition of promastigotes. *Exp Parasitol* **36:** 455-461.