The silicon trypanosome

BARBARA M. BAKKER*, R. LUISE KRAUTH-SIEGEL, CHRISTINE CLAYTON, KEITH MATTTHEWS, MARK GIROLAMI, HANS V. WESTERHOFF, PAUL A. M. MICHELS, RAINER BREITLING and MICHAEL P. BARRETT

1 Department of Pediatrics, Center for Liver, Digestive and Metabolic Diseases, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9713 EZ Groningen, The Netherlands
2 Biochemie-Zentrum der Universität Heidelberg, 69120 Heidelberg, Germany
3 Zentrum für Molekulare Biologie der Universität Heidelberg, ZMBH-DKFZ Alliance, D69120 Heidelberg, Germany
4 School of Biological Sciences, University of Edinburgh, Ashworth Laboratories, Edinburgh EH9 3JT, United Kingdom
5 University of Glasgow, Department of Computing Science & Department of Statistics, Glasgow, G12 8QQ, United Kingdom
6 Department of Molecular Cell Physiology, VU University Amsterdam, 1081 HV Amsterdam, The Netherlands; and Manchester Centre for Integrative Systems Biology, Manchester Interdisciplinary BioCentre, The University of Manchester, Manchester M1 7ND, United Kingdom
7 Research Unit for Tropical Diseases, de Duve Institute and Laboratory of Biochemistry, Université catholique de Louvain, Brussels, Belgium and Faculty of Biomolecular and Life Sciences, University of Glasgow, Glasgow, G12 8QQ, UK
8 Groningen Bioinformatics Centre, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9751 NN Haren, The Netherlands
9 Faculty of Biomolecular and Life Sciences and Wellcome Centre of Molecular Parasitology, University of Glasgow, Glasgow Biomedical Research Centre, Glasgow G12 8TA, United Kingdom

(Received 8 December 2009; revised 17 February 2010; accepted 18 February 2010; first published online 6 May 2010)

SUMMARY

African trypanosomes have emerged as promising unicellular model organisms for the next generation of systems biology. They offer unique advantages, due to their relative simplicity, the availability of all standard genomics techniques and a long history of quantitative research. Reproducible cultivation methods exist for morphologically and physiologically distinct life-cycle stages. The genome has been sequenced, and microarrays, RNA-interference and high-accuracy metabolomics are available. Furthermore, the availability of extensive kinetic data on all glycolytic enzymes has led to the early development of a complete, experiment-based dynamic model of an important biochemical pathway. Here we describe the achievements of trypanosome systems biology so far and outline the necessary steps towards the ambitious aim of creating a ‘Silicon Trypanosome’, a comprehensive, experiment-based, multi-scale mathematical model of trypanosome physiology. We expect that, in the long run, the quantitative modelling enabled by the Silicon Trypanosome will play a key role in selecting the most suitable targets for developing new anti-parasite drugs.

Key words: Trypanosoma brucei, metabolism, gene expression, differentiation, Silicon Cell.

THE AMBITION OF SYSTEMS BIOLOGY

Systems biology seeks to understand how functional properties of living systems, such as biological rhythms, cellular differentiation or the adaptation to drugs may be particularly relevant. Current systems biology is to a large extent (but not exclusively) focused on single-cell systems. These are more amenable to global molecular analysis than multicellular organisms. This is partly because high-throughput post-genomic technologies (transcriptomics, proteomics and metabolomics) make it relatively easy to measure many components of a homogeneous population of cells simultaneously. Furthermore, dynamic measurements of the response of a cell population to a shared stimulus allow insight into the functional connectivity between components (Richard et al. 1996; Hynne et al. 2001; Nikerel et al. 2006, 2009; Schmitz et al. 2009).

Mathematical methods that enable quantitative descriptions of the dynamic interplay between the molecules in living cells are being developed and, for the first time, it is possible to envisage a comprehensive molecular description of the functional circuitry of cellular systems. The Silicon Cell project (Snoep et al. 2006; Westerhoff et al. 2009) involves an international consortium of researchers aiming at a mathematical description of life at the cellular level on the basis of complete and quantitative genomic, transcriptomic, proteomic, metabolomic and phenotypic information. So far, the most ambitious whole-cell modelling efforts have targeted the model
organisms Escherichia coli and Saccharomyces cerevisiae. The advanced state of understanding and the enormous amount of data relating to these organisms have made them obvious candidates for such a comprehensive description. Yet, the variety of organisms in the ‘JWS Online’ model repository (www.jj.bio.vu.nl) demonstrates that the silicon cell initiative is not limited to these organisms.

The Unique Advantages of Trypanosome Systems Biology

The African trypanosome, Trypanosoma brucei, the causative agent of human African trypanosomiasis and Nagana cattle disease (Barrett et al. 2003), has emerged as a front runner in systems biology analysis. The relative simplicity of the energy metabolism of its bloodstream form and the early availability of a comprehensive and uniform set of kinetic data of the enzymes involved, were crucial factors for the successful construction of a detailed computer model of trypanosome glycolysis (Bakker et al. 1997). The obvious questions for this model were initially in the realm of drug-target selection; the first studies analysed in depth how sensitive the pathway overall would be to varying extents of inhibition of each enzyme (Bakker et al. 1999a). Another important factor stimulating further development of trypanosome systems biology was the possibility of reproducible in vitro cultivation, first of the procyclic insect stage, but later also of the long slender bloodstream form (Hirumi and Hirumi, 1989; Haanstra, 2009). Transitions between distinct life-cycle stages can be studied in a tractable and synchronous differentiation system (Fenn and Matthews, 2007). Moreover, the complete genome of T. brucei has been sequenced and annotated, and a metabolic pathway database has been developed (Berriman et al. 2005; Chukualim et al. 2008).

An extension of the scope of trypanosome systems biology to include gene expression is facilitated by the absence of transcriptional regulation in trypanosomes. This implies that, unlike in most other organisms, the gene-expression cascade is regulated only post-transcriptionally. The genes of African trypanosomes – as well as those of the other kinetoplastids – are arranged in polycistronic transcription units which can be hundreds of kilobases long (Berriman et al. 2005; Siegel et al. 2009). All evidence so far indicates that RNA polymerase II transcribes constitutively, without intervention of regulatory factors (Lee et al. 2009; Palenchar and Bellofatto, 2006). Individual mRNAs are excised by a transsplicing complex which places identically capped 3’ ends at the 5’ end of every mRNA (Liang et al. 2003); this splicing is co-ordinated with polyadenylation of the RNA located immediately upstream. Indeed, regulation of mRNA biogenesis may well be restricted to the processing steps (Lustig et al. 2007; Stern et al. 2009), while steady-state levels are further influenced by the rate of mRNA degradation (Clayton and Shapira, 2007). In fact, the majority of evidence concerning regulation of gene expression has implicated mRNA decay as the dominant factor (Clayton and Shapira, 2007; Haanstra et al. 2008b), and this is the only step for which mechanistic details of the regulation are available.

After the complete sequencing of the trypanosome genome (Berriman et al. 2005), mRNA microarray analyses of the differentiation from the bloodstream to the procyclic form have demonstrated that the expression of whole sets of mRNAs is coordinately regulated (Queiroz et al. 2009). When gene-expression is studied during synchronous differentiation, accurate time profiles of extremely homogenous cell populations can be obtained (Kabani et al. 2009). Most results so far suggest that such regulation is mediated by RNA-binding proteins that bind to specific sequences in the 3’-untranslated regions of mRNAs (Archer et al. 2009; Clayton and Shapira, 2007; Estévez, 2008). The rate of mRNA translation and protein turnover are other factors influencing the steady-state protein levels. Here too, key regulatory proteins have been identified (Paterou et al. 2006; Walrad et al. 2009).

Milestones of Trypanosome Systems Biology

A quantitative mathematical model of energy metabolism in the long slender form of the trypanosome (i.e. the form that replicates in the mammalian bloodstream) has been developed (Bakker et al. 1997) and iteratively updated after experimental testing (Bakker et al. 1999a, b; Albert et al. 2005; Haanstra et al. 2008a). This model yields quantitative predictions of the flux through glycolysis, the concomitant ATP production flux, and the concentrations of glycolytic metabolites, at steady state as well as following a perturbation. Input data for the model are kinetic equations and parameters of enzymes and their concentrations. Through this model, the effects of drugs on the glycolytic pathway can be assessed quantitatively starting from their effects on the individual enzymes.

Free-energy metabolism in the bloodstream form of T. brucei has been a logical starting point for the ‘bottom-up’ construction of a ‘Silicon Trypanosome’. Bloodstream forms of the parasite depend exclusively on substrate-level phosphorylation for ATP production through glycolysis, which proceeds as far as pyruvate (Flynn and Bowman, 1973) (Fig. 1). Pyruvate is the end product and is secreted from the cell. Many of the glycolytic enzymes differ in terms of allosteric regulation from their mammalian counterparts, and this probably relates to the fact that in T. brucei the first seven enzymes of the pathway reside within membrane-bounded, peroxisome-like
organelles called glycosomes (Opperdoes and Borst, 1977; Parsons, 2004; Michels et al. 2006; Haanstra et al. 2008), which isolate most of the glycolytic pathway from the rest of the metabolic network involved in consumption of ATP and NAD(H). Even in growing and dividing trypanosomes, virtually all glucose is converted to pyruvate, as the amount required for biosynthesis is quantitatively negligible (Haanstra, 2009). This finding supports the initial choice to model the glycolytic pathway without any branches other than the one to glycerol. Glycerol production is crucial under anaerobic conditions (Fairlamb et al. 1977).

Since the publication of the first version of the glycolysis model (Bakker et al. 1997), there have been two major updates (Helfert et al. 2001; Albert et al. 2005). Both of these involved updates and extensions of the enzyme kinetic dataset, e.g. the explicit inclusion of individual enzymes that were previously grouped into a net multi-step conversion. In the second update (Albert et al. 2005) the enzyme expression levels (V_max) were adapted to reflect the concentrations observed in trypanosomes obtained from controlled state-of-the-art in vitro cultivation. Key missing pieces of information remain the mechanism and kinetics of the transport of glycolytic metabolites across the glycosomal membrane. The identification of semi-selective pores in peroxisomal membranes suggests that the smaller metabolites equilibrate across the glycosomal membrane, while bulkier molecules like ATP or NADH require specific transporters (Grunau et al. 2009; Rokka et al. 2009). This idea justifies, with hindsight, the choice to model the transport of a number of small intermediates as rapid-equilibrium processes.

A number of basic and applied biological questions have been addressed using the glycolysis model. For example, it was predicted and then experimentally confirmed (Bakker et al. 1999a, b) that the uptake of glucose across the plasma membrane was a major flux controlling step and therefore an interesting drug target. Enzymes that have been suggested to control glycolysis in mammalian cells, like hexokinase, phosphofructokinase and pyruvate kinase (Schuster and Holzhütter, 1995), exerted little control in trypanosomes, according to the model (Bakker et al. 1999a; Albert et al. 2005). Experiments, in which the expression of these enzymes was knocked down, confirmed this prediction qualitatively. However, the enormous overcapacity of some enzymes, which was predicted by the model, was shown to be exaggerated (Albert et al. 2005). This suggests that there are in vitro regulation mechanisms affecting these enzymes in a currently unknown fashion. Protein phosphorylation may contribute, since a glycosomal phosphatase has been identified in developmental signalling (Szoor and Matthews, unpublished data). The inhibition of anaerobic glycolysis by glycerol was also reproduced by the model, first qualitatively and then quantitatively (Bakker et al. 1997; Albert et al. 2005).

An interesting biological feature that was revealed by the model was the relationship between compartmentation of glycolysis in glycosomes and the virtual absence of allosteric regulation of the glycolytic enzymes. Glycolysis models predict that glycolytic intermediates accumulate readily due to the investment of ATP at the beginning of the pathway (Teusink et al. 1998; Bakker et al. 2000). This risky ‘turbo’ effect can be avoided either by allosteric feedback
regulation of hexokinase or by compartmentation of the pathway in glycosomes (Fig. 2). Compartmentation prevents the accumulation of intermediates, because the net ATP production occurs outside the glycosome and this excess of ATP cannot activate the first enzymes of glycolysis. This model prediction was recently confirmed experimentally (Haanstra et al. 2008a), providing a clear example of model-driven experimental design and hypothesis-driven systems biology. According to model predictions the glycolytic intermediates glucose 6-phosphate, fructose 6-phosphate and fructose-1,6-bisphosphate should accumulate on addition of glucose if the glycolytic enzymes are not properly located in the glycosome. Indeed, accumulation of glucose 6-phosphate could be measured in a PEX14-RNAi mutant in which protein import into the glycosomes is disturbed. A similar phenotype was observed on glycerol addition, which led to accumulation of glycerol 3-phosphate, both in the model and in the PEX14-RNAi cells. Also in accordance with model predictions, a down-regulation of the expression of the genes encoding hexokinase and glycerol kinase rescues the PEX14-RNAi cells on glucose and glycerol, respectively (Kessler and Parsons, 2005; Haanstra et al. 2008a).

More recently, a model of the gene-expression cascade, based on quantitative knowledge of transcription, RNA precursor degradation, trans-splicing and mRNA degradation for phosphoglycerate kinase (PGK) has been generated (Haanstra et al. 2008b). The model allowed a quantitative analysis of the control and regulation of the expression of the PGK isoenzymes. It was shown that regulation of mRNA degradation explains 80–90% of the regulation of mature mRNA levels, while precursor degradation and trans-splicing make only minor contributions.

In spite of the success of the model, it covers to date only a small part of trypanosome metabolism. This relates, for instance, to the fact that even the compartmentalised glycolysis does branch into other pathways, for example towards the biosynthesis of glycoconjugates and the pentose phosphate pathway. Although the fluxes into these branches may be small, they are vital for trypanosomes. Sufficient kinetic data have become available to enable extension of the model to include the pentose phosphate pathway which provides NADPH for reductive biosyntheses and also reducing equivalents to sustain cellular redox balance. Since redox balance is intimately related to the biosynthesis of trypanothione (from polyamine and glutathione precursors), a natural next step in a bottom-up systems biology approach to trypanosome metabolism would be the inclusion of the trypanothione–pentose phosphate pathway and related areas of redox metabolism (Fig. 3).

**GROWTH STAGES OF BUILDING A SILICON TRYpanosome**

Our current level of knowledge of trypanosome redox metabolism, as well as its biological importance (Krauth-Siegel and Comini, 2008), render it a natural choice for a next model extension (Fig. 3). The inclusion of redox metabolism is particularly
The silicon trypanosome

interesting as trypanosome redox metabolism is sufficiently different from its human counterpart to offer perspectives for drug discovery. The unusual polyamine–glutathione conjugate trypanothione or bis(glutathionyl)spermidine (Fairlamb and Cerami, 1992) takes on the majority of roles served by glutathione in most other cell types. In addition, work in the last few years revealed that the enzymes involved in the synthesis and reduction of trypanothione are essential for the parasite (Krauth-Siegel and Comini, 2008).

The trypanocidal drug eflornithine exerts its trypanocidal activity as an irreversible inhibitor of the enzyme ornithine decarboxylase (Bacchi et al. 1980), an enzyme involved in trypanothione biosynthesis (enzyme 10 in Fig. 3). A significant amount of information is available on kinetic parameters of that pathway, too. Preliminary attempts to model trypanothione metabolism have been made (Xu Gu, University of Glasgow PhD thesis, unpublished). Information available on the abundance of key metabolites measured in bloodstream form T. brucei grown in vitro (Fairlamb et al. 1987) and in vivo (Xiao et al. 2009), before and after exposure to eflornithine, was used to determine whether predicted behaviour under those perturbed conditions emulated the measured behaviour. The scarcity of kinetic data describing the whole pathway, however, has presented many challenges to constructing a model that captures observed behaviour. The acquisition of new kinetic data and the implementation of new mathematical tools to fill gaps in the data (Nikerel et al. 2006; Smallbone et al. 2007; Resendis-Antonio, 2009) should improve this.

An extension of the glycolysis model to include the pentose phosphate pathway (Hanau et al. 1996; Barrett, 1997; Duffieux et al. 2000) and trypanothione metabolism should be a suitable next step in the modular approach that we envisage towards a complete Silicon Trypanosome. Initial efforts in this direction (not published) have indicated the importance of the compartmentation of the pentose phosphate pathway. Although most of the enzymes of the pathway have a peroxisome targeting

---

**Fig. 3.** The glycolytic and trypanothione pathways are linked through the oxidative pentose phosphate pathway. Metabolites are presented in abbreviated form within rectangles. Enzymes and transporters are circles. Kinetic data is available for those shaded grey. Met = methionine; Arg = arginine; Gly = glycine; Glu = glutamate; Cys = cysteine; ATP = adenosine triphosphate; SAM = S-adenosylmethionine; dcSAM = decarboxylated S-adenosylmethionine; MTA = methylthioadenosine; Orn = ornithine; Put = putrescine; Spd = spermidine; γ-GC = γ-glutamylcysteine; GSH = glutathione; GSpd = glutathionylspermidine; T(SH)2 = reduced trypanothione; TSH = oxidised trypanothione; NADPH = reduced nicotinamide adenine dinucleotide phosphate; Glc = glucose; G6P = glucose 6-phosphate; Ru5P = ribose 5-phosphate; Oxy = oxidised cellular metabolites; Red = reduced cellular metabolites zeroth = methionine transport; 2. = arginine transport; 3. = glycine transport; 4. = glutamate transport; 5. = cysteine transport; 6. = methionine adenosyltransferase; 7. = arginase (N.B., a robust arginase gene orthologue has not been annotated in the T. brucei genome project, raising the possibility that arginine does not serve as a source of ornithine in these cells); 8. = S-adenosylmethionine decarboxylase; 9. = prozyme; 10. = ornithine decarboxylase; 11. = spermidine synthase; 12. = γ-glutamylcysteine synthetase; 13. = glutathione synthetase; 14. = γ-glutamylspermidine synthetase; 15. = trypanothione synthetase/amidase (in T. brucei 14. & 15. are catalysed by a single protein); 16. = trypanothione reductase; 17. = oxidative pentose phosphate pathway (glucose 6-phosphate dehydrogenase, 6-phosphogluconolactonase & 6-phosphogluconate dehydrogenase); 18. = glucose transporter; 19. = hexokinase (this enzyme links the redox pathway to glycolysis); 20. = pyruvate transporter; 21. = methionine cycle enzymes; 22. = The pathway of electrons from reduced trypanothione for final acceptance on oxidised cellular metabolites or macromolecules is complex, involving transfers via other redox active intermediates including tryparedoxin (thioredoxin-like proteins) and peroxidoxin.
sequence (PTS1), a significant fraction of their activity is often found in the cytosol (Michels et al. 2006; Heise and Oppendoes, 1999; Duffieux et al. 2000). A correct localisation of the enzymes as well as good estimates of the transport of intermediates across the glycosomal membrane will be key to a good model of the pentose phosphate pathway.

CHALLENGES OF TRYPANOSOME SYSTEMS BIOLOGY

The first initiatives in systems biology of trypanosomes as well as of other organisms dealt with enzymatic sub-systems, such as glycolysis. The models have depended largely on kinetic data for isolated enzymes. However, the abundance of these enzymes can, in principle, be controlled by the rates of transcription, RNA processing, translation, protein modification and turnover. These processes themselves may be regulated through complex signalling networks in response to both internal and external conditions (Westerhoff et al. 1990).

To include gene expression in a Silicon Trypanosome requires a dramatic increase in model complexity—as well as the acquisition of new types of data on a large scale. Fortunately, the absence of transcriptional control of most individual open reading frames makes trypanosome gene expression simpler than that of yeast or even E. coli, rendering it uniquely amenable to mathematical modelling.

It may well be possible to describe much of trypanosome mRNA metabolism using the following parameters: the rate constant of processing of the precursor RNA, i.e. of trans-splicing; the rate constant of degradation of the precursor (which competes with its trans-splicing); the rate constant of polyadenylation; and the rate constant of mRNA degradation. The rates of degradation of the precursor and the mature mRNA can be measured by inhibiting splicing and transcription. To measure the rate of mRNA processing two approaches are possible. First, one can inhibit transcription alone, and assay precursor decay; this approach is, however, compromised by practical constraints since splicing is very rapid. Second, the splicing rate can be calculated based on the steady-state abundance of the precursor mRNA, and the half-life and abundance of the products. This methodology has already been applied to the mRNA encoding PGK and it was demonstrated that splicing occurred within less than one minute; mRNA decay was the primary determinant of mRNA abundance (Haanstra et al. 2008b).

Previous microarray studies with yeast have yielded estimates of the half-lives and polysomal loading of many mRNAs (e.g. Grigull et al. 2004). Deep sequencing technology—being more sensitive—should allow measurement of the abundances of all mRNAs and precursors on a genome-wide scale and to the accuracy required for the modelling; from these data, it should be possible to derive the steady-state abundances and half-lives of all RNAs, revealing regulated or inefficient processing. This—combined with global polysome profiling—will provide quantitative data which allow quantifying the regulation of the processing, degradation and translation of each mRNA (Daran-Lapujade et al. 2007). The next challenge would then be to integrate such measurements with metabolic modelling in order to provide a complete model of pathways, from DNA to metabolic end-products.

ANTICIPATED OUTCOMES FROM A SILICON TRYPANOSOME

So far, the systems biology approach to trypanosomes has contributed to a fundamental understanding of cellular regulation (Bakker et al. 1999a; Haanstra et al. 2008a), as well as to improvements in the drug-target selection process (Bakker et al. 1999a, b; Albert et al. 2005; Hornberg et al. 2007). Since the initial systems biology analysis only addressed processes associated with less that 1% of the organism’s genome, we would expect many more new insights to lie ahead.

Drugs currently used against human African trypanosomiasis are unsatisfactory for a number of reasons, including their extreme adverse effects in the patient and the emergence of resistant parasites. New drugs are urgently needed and there is hope that a better understanding of the control points of the metabolic network can guide the selection of optimal drug targets. This has already been achieved for enzymes of the glycolytic pathway (Hornberg et al. 2007). This information can be used alongside enhanced chemoinformatics (Frearson et al. 2007) in order to determine those components of the trypanosome that are most amenable to drug targeting.

As a consortium we have embarked on the construction of a Silicon Trypanosome. In this review we have discussed the current status and future directions of trypanosome systems biology that form the context of this endeavour. Our ambition is to achieve a comprehensive, quantitative description of the flow of information from gene, through transcript and protein, to metabolism and back. This will allow prediction of how the parasite responds to changes in its environment, with respect to nutrients, temperature and/or chemical inhibitors. It will also assist the deciphering of complex phenotypes generated by genetic perturbations in the laboratory or in the field. Thus, model predictions will improve our biological understanding of the differentiation and adaptation of the parasite as well as stimulate the discovery of inhibitors that attack processes which control trypanosome growth. The latter should contribute to the development of new optimised drugs for trypanocidal chemotherapy. Pioneering efforts

Barbara M. Bakker and others
have focused on energy metabolism and recently started to include adaptations of the parasite via gene expression (Haanstra, 2009).

The construction of a complete Silicon Trypanosome, which integrates metabolism, gene expression and signal transduction, is an ambitious project. Clearly the route towards this objective will be long, and many challenges will emerge as the datasets required to build such a model are collected and analysed. However, the emergence of methods to allow collection of massive datasets, at every level, suggests that we may, in time, be able to generate a reasonably complete mathematical description of trypanosome cellular biology. Even if completion is not feasible, the evolving description will always represent the best conceivable dynamic representation of our knowledge of trypanosome biology. As a result, drug development programmes will have at their disposal a predictive model of the trypanosome to help identify those parts of metabolism most amenable to targeting by novel drugs and to controlling vital functions of the parasite. The project will be strengthened by parallel world-wide systems biology projects of human metabolism, in which some of us will be involved. After all, killing trypanosomes is easy. The difficulty is to kill the trypanosome without harming its host (Bakker et al., 2002). A careful comparison of the behaviour of our Silicon Trypanosome to quantitative knowledge of the control of human metabolism, will allow the identification of selective targets.

ACKNOWLEDGEMENTS

The work of BMB was funded by NWO Vernieuwingsimpuls and by a Rosalind Franklin Fellowship. HVW thanks BBSRC and EPSRC for support through the MCISB grant (http://www.systembiology.net/). MPB is grateful to the BBSRC for their support of the BBSRC-ANR “Systrep” consortium. The Silicon Trypanosome consortium is supported by a grant from SysMO2 (www.sysmo.net).

REFERENCES


