

Invited Review

# The biology of kinetoplastid parasites: insights and challenges from genomics and post-genomics

Keith Gull\*

School of Biological Sciences, University of Manchester, 2.205 Stopford Building, Oxford Road, Manchester M13 9PT, UK

Received 2 October 2000; received in revised form 19 December 2000; accepted 19 December 2000

## Abstract

Kinetoplastid parasites exhibit a rich and diverse biology which mirrors many of the most interesting topics of current interest and study in the broader biological sciences. These evolutionarily ancient organisms possess intriguing mechanisms for control of gene expression, and exhibit complex patterns of cell morphogenesis orchestrated by an internal cytoskeleton. Their cell shapes change during a set of complex cell type differentiations in their life cycles. These differentiations are intimately linked to interactions with mammalian hosts or insect vectors, and often, these differentiations appear central to the successful transfer of the parasite between vector and host, and host and vector. The basics of this rich and complex cell and life cycle biology were described (with often rather forgotten clarity and prescience) in the early period of the last century. The last 30 years have seen major developments in our understanding of this biology. Ultrastructural differences in the various cells of the life cycle stages of *Trypanosoma brucei*, *Trypanosoma cruzi* and the various *Leishmania* species have been documented, and such studies have proven highly informative in defining important aspects of parasite adaptation. They have also proven to be a rich source of information for defining unusual aspects of parasite cell biology, novel organelles and cell architecture. This ultrastructural cell biology has been mirrored in a set of biochemical explanations defining unusual aspects of metabolism, surface molecules, and organelles. Finally, the application of molecular biology to these parasites revealed fascinating layers of complexity in the control of gene expression. These molecular studies have given us particular insights into polycistronic transcription, trans-splicing, RNA editing and gene rearrangements during antigenic variation. In contrast to other microbial systems, these cell biological, biochemical and molecular studies have not been greatly aided by insights gained from genetics – the diploid nature of the genome has discouraged the application of selectional genetics, mutant isolation and analysis. This is an important fact, since in general, it means that we have only recently started to analyse the phenotypes of mutants produced in the context of reverse genetics. In the following, I will argue that this lack of investment in the analysis of mutant phenotype is just one of the challenges that will need to be met if we are to gain the expected added value from the parasite genome projects. In this presentation, I will use some of the current areas of interest in the biology of *T. brucei*, *T. cruzi* and the *Leishmania* species to rehearse some of the insights and challenges that are likely to stem from the application of genomics and post-genomic studies to the kinetoplastid parasites. In some cases, I will exemplify points by illustrations from my laboratory's work, interests and hypotheses. The presentation slants therefore towards *T. brucei* biology, however, in each case the reader will, no doubt, see the generalities of application to other kinetoplastid parasites. © 2001 Australian Society for Parasitology Inc. Published by Elsevier Science Ltd. All rights reserved.

**Keywords:** Trypanosome; Genomics; Proteomics; *Leishmania*; Parasite; RNA interference

## 1. Biology of kinetoplastids

Table 1 contains a selected list of important aspects of the basic cell biology of kinetoplastids currently being addressed in many laboratories. Each of these represents an area that is informed by similar studies in many other cell systems. However, in addition, each has particular aspects that are key to understanding events central to life as a parasite. I will argue that each of these areas of current activity will be massively influenced in the coming years by

the very rapid accumulation of genome sequence information. I have recently argued (Gull, 2000) that the parasite genome projects will produce read-outs at a variety of levels. The first of these will be directly at the level of the parasite genome itself and the genome-related events that are central to orchestrating parasite functions.

Table 2 lists some of these first level read-outs that I expect to emerge from parasite genome projects. Each of these is likely to speak to issues of the basic cell biology of the parasites such as those rehearsed in Table 1. By enhancing these core studies of parasite biology, a second level of influence is likely to ensue. Here, the genome information will complement hypothesis driven research and provide

\* Tel.: +44-161-275-5108; fax: +44-161-275-5763.

E-mail address: k.gull@man.ac.uk (K. Gull).

Table 1  
Some of the interesting aspects of kinetoplastid basic cell biology

---

Nuclear structure in relation to VSG expression
Chromosome segregation
Flagellum morphogenesis/function
Protein targeting
Secretion/uptake
Plasma membrane domains
Membrane proteins
Cytoskeleton
Organelles such as the glycosome and mitochondrion
Metabolic specialisation/compartimentalisation
Cell differentiation
Cell cycle
Kinetoplast/nucleus co-ordination
Virulence factors
Host and vector interactions
Recognition and receptors
Invasion
Entry and establishment
Survival

---

novel insights to parasitism, virulence, parasite evolution and epidemiology (Table 2). The major hope is that this level of influence reads through to produce, via translational science, a direct benefit to areas such as diagnostics, drugs, vaccines and other opportunities for intervention in the disease process. The publication of the first complete sequence of a chromosome from a kinetoplastid parasite (Myler et al., 1999) did, in fact, produce information that speaks to many of the first level issues rehearsed in Table 2.

The intimate interrelationship of complete genome information to understanding the biology of the kinetoplastid parasites is exemplified by consideration of the complex nature of both the genome and the life cycle of *Trypanosoma brucei*.

The *T. brucei* genome contains three main chromosome types classified by their size ranges (Table 3). The megabase chromosomes (Fig. 1) have the intriguing organisation of internal regions of protein coding genes (organised in a manner that facilitates polycistronic transcription), coupled with telomeric expression sites for metacyclic and bloodstream versions of the variable surface glycoprotein genes (Ersfeld et al., 1999; El-Sayed et al., 2000). We are particularly interested in the minichromosomes (Ersfeld and Gull, 1997; Gull et al., 1998), and using high resolution pulsed field gels, we have recently shown that the minichromosome karyotype is clone specific, but that individual chromosomes show size variations consistent with telomeric growth and shortening. Full sequence information for the megabase chromosomes will soon provide fascinating insights into the genomic environment responsible for the different forms of gene expression (and silencing) from these chromosomes and will, no doubt, provide more 'first level' information on the various elements and domains listed in Table 2.

Consideration of even a simplified version of the events of the *T. brucei* life cycle (seen in Fig. 2) illustrates the

exquisite co-ordination of events occurring as the parasites proceed through their life cycle. They make a series of transitions between three major environments: mammalian host bloodstream, tsetse midgut and tsetse salivary gland. The life cycle is characterised by changes in cell shape, cell cycle, metabolism, surface coat, etc. At two of these transition points, bloodstream to tsetse midgut and tsetse salivary gland to mammalian bloodstream, there is a specific pattern of events. In each case, the transmitted parasite (stumpy form and metacyclic, respectively) exhibits particular differentiated attributes useful to its survival in the next environment. On reaching that environment, in each situation, the cell enters a proliferative cell cycle and colonises the midgut or bloodstream, respectively. In one of the other transitions, tsetse midgut to the salivary gland, the parasite differentiates to a form that uses its flagellum for attachment to surfaces – the epimastigote. Consideration of these life cycle events illustrates how each of the basic cell biological phenomena outlined in Table 1 are likely to be involved in particular steps. The genome projects will reveal the full set of players

Table 2  
Read-outs from the parasite genome projects<sup>a</sup>

---

First level <sup>b</sup>
Genome composition and evolution
Molecular karyotype
Chromosomal rearrangement
Genome plasticity
Synteny
Ploidy significance
Multi-gene families: maintenance and diversity
Repetitive sequences
Promoters
Transcription
Processing descriptors: cis- and trans-splice sites; Poly A sites
Centromeres
Telomeres
Organelle genome dependency
Differential gene expression
Control elements
Antigenic variation
Redundancy
Second level <sup>c</sup>
Insights to:
New basic biological phenomena
Parasitism
Parasite/vector biology
Virulence
Parasite evolution
Epidemiology
Complementation of hypothesis driven research
Added value and efficiency to all studies of parasites
Translational science
Diagnostics
Drugs
Vaccines
Intervention opportunities

---

<sup>a</sup> A series of areas that are likely to benefit from the genome projects.

<sup>b</sup> The first level describes areas related to the genome, genes and gene expression.

<sup>c</sup> The second level describes rather broader areas and themes.

Table 3

Properties of the three main chromosome classes of the *Trypanosoma brucei* genome

Megabase chromosomes	
11 pairs (at least)	
Diploid	
1–6 Mb	
Length polymorphisms	
Haploid VSG gene expression sites at telomeres	
Expressed genes and repetitive sequence blocks	
Interior VSG gene clusters	
Intermediate chromosomes	
200–900 kb	
Number and size variation	
Ploidy?	
Minichromosomes	
Linear	
30–150 kb	
177 bp repeat (~90% of sequence?)	
GC and AT rich repeats	
Silent telomeric VSG gene	
10–20% genome	
Ploidy?	

within the genome that can potentially contribute to these different aspects of the cell biology of *T. brucei*. A major, immediate benefit should be that one will be able to make judgements, of which candidate genes are likely to be the most important, rather than working, as at present, with those few known genes. I have suggested (Gull, 2000) that this will improve the quality of questioning and reasoning in the field. The quality and burden of proof demanded by colleagues and reviewers will also rise rapidly!

## 2. Genomics and post-genomics

Provision of the in silico transcriptome and proteome potential through the genome projects will provide many challenges for those studying kinetoplastids. Given the

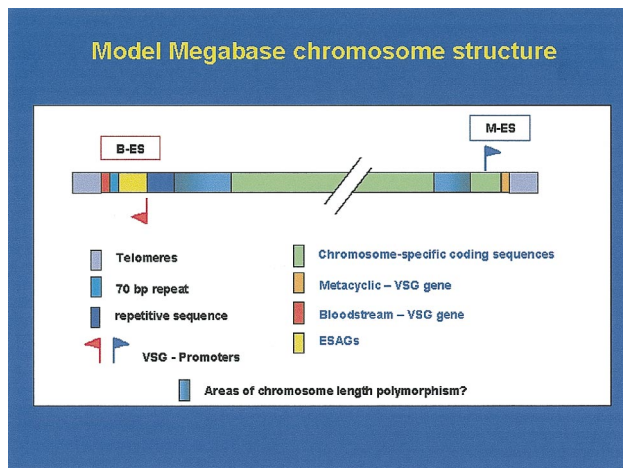


Fig. 1. A general diagram of the likely organisation of a megabase chromosome of *Trypanosoma brucei*.

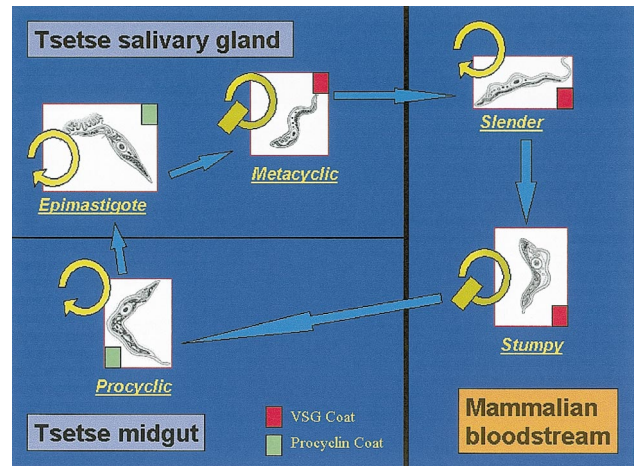


Fig. 2. Life cycle cartoon of the main stages of the *Trypanosoma brucei* life cycle, with thanks to Keith Vickerman for his classic drawings. Here, I have focussed the interpretation of the life cycle on the parasite's use of cell cycle control to orchestrate the major host/vector and vector/host transitions. The yellow lines denote a proliferative (arrow) or non-proliferative (bar) cell cycle. The cell surface molecule type is denoted by the coloured box.

likely size of the revealed gene set of an individual kinetoplastid, the individual scientist and the field in general will need clear strategies for experimental interrogation of this data set. Workers with other organisms have travelled this route and we should be able to learn much by being slightly late down the road! The obvious challenges are whether to interrogate gene function globally – organised array programmes, co-ordinated mass gene inactivation programmes, etc. – or whether to move more progressively and selectively? Desired speed and time of knowledge acquisition and cost/benefit analyses will inform these decisions. Gaining an early, yet clear overview of the global expression pattern of the genome within vector and host stages of the parasite will be a major target. It will be interesting to see whether array technologies, proteomic or other analyses will reveal global patterns of gene expression in these parasites, given their particular reliance on post-transcriptional and translational levels of gene regulation.

Subsequent or parallel interrogation of kinetoplastid parasite biology is aided by now having some excellent tools for reverse genetics available. Gene-knockouts, antisense, RNA interference (RNAi), conditional expression and many other approaches are now in place or are rapidly being developed. These technologies will not only be important for academic studies of parasite functions, but will be critical to the translational studies aiming at target identification and validation in drug discovery initiatives. I suspect that the most important challenge for post-genomic studies of protozoan parasites is that of defining phenotypes. We are probably better equipped to make mutants, knockout genes and analyse gene expression patterns than we are to analyse complex phenotypes. Here, I make a connection with my early comments about the lack of usefulness of selectional genet-

ics in these parasites. Due to this, we are not well equipped with the long history of phenotype analysis that characterises model organisms such as *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila* and *Caenorhabditis*. Moreover, much of the interesting biology of the kinetoplastid parasites occurs in rather experimentally intractable locations in the host or vector. We will need to invest heavily in defining new approaches to phenotype analysis if we wish to reveal the function of a reasonable proportion of the large number of genes expected to be represented in these genomes. Given the nature of genome organisation, transcriptional control and processing in the parasitic protozoa, we can expect these genomes to be ‘gene-dense’ (El-Sayed et al., 2000; Ersfeld et al., 1998; McDonagh et al., 2000; Myler et al., 1999). The 12 Mb genome of yeast revealed around 6000 genes on completion of the genome project. We may be looking at more genes in the kinetoplastid parasites. Bioinformatics and comparative genomics is likely to be a powerful approach to defining the potential functions or relationships of many genes in the repertoire. However, it is salutary to remember that in each of the many completed genome projects, around 40–60% of all genes identified have had to be labelled as ‘hypothetical protein’ on first annotation. In the yeast genome project, this figure of ‘orphan’ genes without identity or without even ‘in silico’ clues to biochemical or cellular function was 56% of the total: 3480 genes! The figures for *Escherichia coli* were 60% of the genes: 2583 genes! Given the cellular complexity of the kinetoplastid protozoa, it is clear that we must look forward to similar figures for the initial analysis of these genomes.

What functions might such ‘orphan genes’ and their products help orchestrate and influence in *Trypanosoma* and *Leishmania*. In Table 4, I have outlined some thoughts on what we might expect. These include critical functions in parasitism influencing both survival and virulence in host and vector, cytoarchitecture specialisations, nuclear and

Table 4

Insights to new functions defined by orphan genes in parasite genomes?<sup>a</sup>

Parasitism
Virulence
Survival in particular environments
Host interactions
Vector interactions
Nuclear and organelle genome plasticity, expression and interaction
Extent, global control and integration of metabolism
Cytoplasmic regionalisation and specialisation
Cell structures: a new molecular cytology
Signalling and communication: in space and time
The ‘slow growth or no growth’ condition
Insurance policies and control of alternative life cycle pathways

<sup>a</sup> Many genes will be identified in the genome project and annotated as hypothetical since they have no relation to known genes. This collection of ‘orphan genes’ is likely to include many genes which will provide unique insight to novel functions in parasite biology.

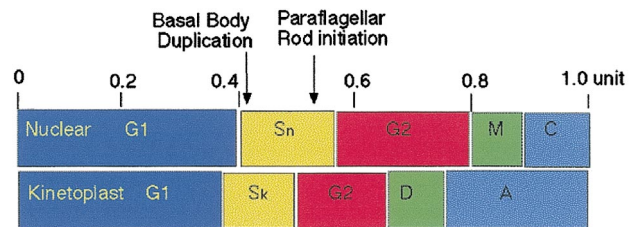


Fig. 3. Cell cycle periods and timings for *Trypanosoma brucei* procyclic cells.

kinetoplast properties and expression patterns. In the following sections, I will use some of our current interests and recent discoveries to illustrate some of the opportunities and challenges presented by *T. brucei* genomics and techniques that will be useful for post-genomics and phenotype analysis in this parasite.

### 3. The two-unit genome replication/segregation problem

The trypanosome cell possesses a two-unit genome — a nuclear genome — and, in contrast to many other eukaryotes, there is a single mitochondrion and the unit genome of this mitochondrion is localised in a specific structure, the kinetoplast. In trypanosomes, both of these genomes are replicated periodically in the cell cycle. Using immunofluorescence detection of bromodeoxyuridine incorporation into replicated DNA (Woodward and Gull, 1990) allowed us to determine the timings of such events in the *T. brucei* procyclic cell cycle (Fig. 3).

The two genomes can be easily visualised by fluorescence staining using DAPI. Observation of the cells within an exponentially growing population of *T. brucei* procyclic cells illustrates the observation, established early in the last century, that the kinetoplast DNA segregates before the replicated nuclear genome is separated at mitosis. This provides a very useful marker within the cell cycle since, in contrast to other eukaryotes, it is possible to use the configuration of the nucleus and kinetoplast to separate the cell cycle morphologically into four discrete periods, classifying cells as 1KIN, 2KIN, 2KmitoticN and 2K2N (Sherwin and Gull, 1989). The movement apart of the flagellum basal bodies mediates segregation of the mitochondrial genome during the cell cycle. A direct physical connection can be demonstrated between the basal bodies and the kinetoplast (Robinson and Gull, 1991). The cell cycle timing data suggest that DNA synthesis is triggered in the mitochondrion at essentially the same time as in the nucleus. This temporal pattern suggests a unique regulation of critical events within the trypanosome cell cycle. Genomic and comparative genomic analyses are likely to be of great value in identifying nuclear encoded gene products likely to be involved in mediating and regulating kinetoplast DNA synthesis, as well as those co-ordinating these events with nuclear S phase.

#### 4. New tubulins in trypanosomes

The main components of microtubules — alpha and beta tubulin — were identified around 30 years ago by biochemical means. Since then, two other members of the tubulin superfamily, gamma and delta, have been identified by genetic routes. First identified in *Aspergillus* (Oakley and Oakley, 1989) and *Chlamydomonas* (Dutcher and Trabuco, 1998), respectively, these two tubulins have subsequently been recognised in other organisms. The alpha and beta tubulins of *T. brucei* have been recognised for some time, and a little while ago, we were able to clone the gene encoding gamma tubulin (Gull, 1999; Scott et al., 1997). Last year, we initiated a search for the *T. brucei* gene encoding delta tubulin. We used PCR and genomic sequence comparison methods, similar to approaches that we had used successfully for gamma tubulin. Areas of the tubulins are highly conserved and we used PCR methods to search for the presence of the *T. brucei* homologue of delta tubulin.

To our surprise, after we cloned the *T. brucei* delta tubulin homologue, we also identified two new, divergent tubulin-like sequences (Vaughan et al., 2000). We found that both of the new sequences were also present within the *T. brucei* genome project databases at The Sanger Center and The Institute for Genome Research as partial sequences. Since these two new tubulins, epsilon and zeta, possess relatively low homology with other members of the family, we collaborated with Dr Terri Attwood in Manchester to utilise her technique of protein fingerprinting used to create the PRINTS pattern database. Fingerprinting is a multiple-motif iterative process that commences with sequence alignment and excision of conserved regions. Diagnostic performance is enhanced by iterative database scanning and the motifs 'mature' with each database pass, as more sequences are matched and assimilated into the process. The fingerprint analysis also clearly assigned both of the two novel *T. brucei* sequences as independent new members of the tubulin superfamily (Vaughan et al., 2000). This is an intriguing discovery. Tubulins represent one of the most studied gene and protein families in biology. The microtubule cytoskeleton of trypanosomes has received a fair degree of attention over the past 15 years without much hint of the extent of the family becoming apparent until our recent studies. The discovery has many implications. Clearly, it means that considerations of the biology of the microtubule cytoskeleton of eukaryotes need to become rather more sophisticated. Tubulin sequences have also been used extensively in molecular phylogeny and studies of early eukaryotic cell evolution. Again, with hindsight, such analyses have clearly been rather restricted in their view of the evolution of the tubulin family, and hence, its usefulness as a molecular marker.

#### 5. Comparative genomics of the new tubulins

During our identification of both epsilon and zeta tubulin,

we were aided by analyses of partial and complete sequences in other genome databases. We were interested to note that epsilon tubulin was present in the mammalian EST databases (and subsequently was published). We also note that epsilon, delta and zeta appeared to have a restricted occurrence within the completed genomes of organisms such as yeast, *Drosophila* and *Caenorhabditis elegans*. Table 5 indicates our analysis of the present occurrence of the members of the tubulin superfamily within selected eukaryotes. We have detected homologues in the genome databases of *Leishmania* and of *Plasmodium*.

No doubt, the occurrence of these new tubulins is reflective of a particular pattern of microtubule biology. What might be the reason behind the absence in yeast and certain other organisms of these and, maybe yet, other new tubulins? We find it intriguing that, at present, the possession of these new tubulin genes (epsilon, delta and zeta) correlates reasonably well with the expression of a triplet microtubule basal body and a 9 + 2 microtubule axoneme. The only difficulty here is *Drosophila*, which appears to have a triplet basal body at some stages of its development. It is clear that the use of gene inactivation techniques in *T. brucei* will be very instructive in providing insights into the function of this extended superfamily, and hence, the evolutionary divergence of the microtubule cytoskeleton within eukaryotic cells.

#### 6. Synteny

I believe that synteny is one of the most interesting and important aspects of the genome projects. Before the kinetoplastid genome projects got underway, a number of small-scale analyses of genes and gene families indicated that a reasonable amount of synteny might exist between the chromosomes of *T. brucei*, *Trypanosoma cruzi* and *Leishmania*. An example of this synteny and its usefulness in gene identification came in our analysis of the new members of the tubulin superfamily. The new zeta tubulin sequence that we cloned is on chromosome 1 of *T. brucei* and this chromosome was in the process of being sequenced at the Sanger Center. Using sequence reads and contigs from this source and from the TIGR *T. brucei* sequencing project, I was able to construct a map of the likely open reading frames from

Table 5  
The tubulin family: occurrence so far<sup>a</sup>

<i>Trypanosoma</i>
Members: alpha, beta, gamma, delta, epsilon, zeta
<i>Saccharomyces</i> , <i>Caenorhabditis</i> , <i>Drosophila</i>
Members: alpha, beta, gamma
Human, Mouse
Members: alpha, beta, gamma, delta, epsilon

<sup>a</sup> So far, six tubulins have been identified in *Trypanosoma brucei*. This table describes the occurrence and distribution of these tubulins within other complete and partially sequenced genomes.

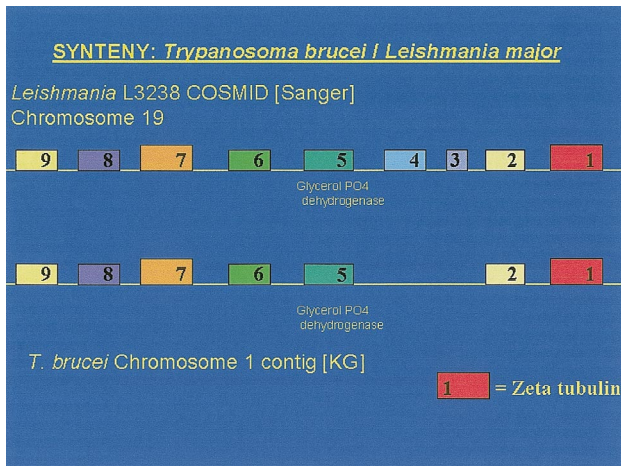


Fig. 4. Synteny between the *Trypanosoma brucei* and *Leishmania major* genomes around the zeta tubulin gene.

the zeta tubulin gene. This map is extremely similar to that of a small region of chromosome 19 of *Leishmania*, and a comparison of these maps is shown in Fig. 4. There is a high level of synteny between these regions of the genome of these two kinetoplastids, and larger scale comparisons that we have produced show synteny levels to be very significant across these two genomes. Synteny will be a significant tool aiding the task of identifying genes in kinetoplastids, and moreover, in understanding many of the expected first level read-outs described in Table 2 for these parasite genomes.

## 7. The flagellum

The procyclic *T. brucei* cell has a precisely defined pattern and polarity (Fig. 5). The anterior end of the cell points towards the direction of movement of the cell and is the narrower end. The flagellum which is attached to the cell body along its length is subtended by a basal body at the posterior end of the cell. In G<sub>1</sub>, there is a small probasal body close to the basal body. Our early studies, using whole-mount negatively stained *T. brucei* cytoskeletons, allowed us to define very precisely the structural changes occurring during the cell cycle (Sherwin and Gull, 1989). The application of specific monoclonals then allowed us to quantify the timings and order of replication and segregation of these organelles within the unit cell cycle (Gull, 1999; Sherwin and Gull, 1989; Woodward and Gull, 1990). The first morphological evidence of entry into the cell cycle is the duplication of the basal bodies at around 0.4 of the unit cell cycle. The new flagellum then starts to extend and this elongation proceeds through much of the subsequent portion of the cell cycle. The new flagellum is subtended by the basal body that remains at the posterior end of the cell. The old and new flagellum basal bodies then move apart within the cell, such that the mitotic spindle forms in the nucleus in a position where its long axis crosses the old flagellum.

The major events of this cell cycle are seen in the cartoon shown in Fig. 5 (Gull, 1998). The flagellum of kinetoplastid parasites is an interesting organelle. It possesses the intriguing paraflagellar rod (PFR) in addition to an axoneme, and we have concentrated much attention on this novel structure. The flagellum is often thought of as an organelle for motility of the cell, but I feel that it is likely to have many additional functions in kinetoplastid protozoa. Table 6 outlines some of these thoughts for *T. brucei* and I will exemplify some of them by reference to our work. Obviously, these general concepts and hypotheses on the wider biology of the flagellum apply to the other kinetoplastids, and indeed, to other parasitic protozoa.

## 8. Paraflagellar rod mutants and RNA interference gene-silencing

Our studies of the PFR, along with those from Jon

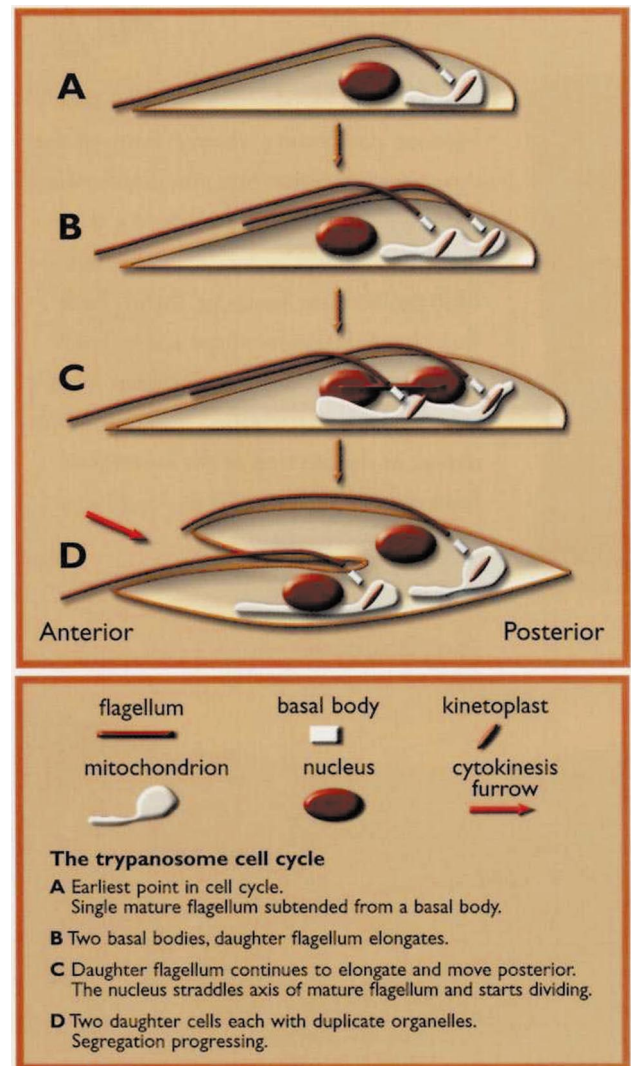


Fig. 5. A cartoon of the main structural events of *Trypanosoma brucei* procyclic cell cycle, from Gull (1998).

Table 6  
Pathogenicity functions that may be associated with the *Trypanosoma brucei* flagellum<sup>a</sup>

Motility of the cell	Pathogenicity Life cycle Tissue location movements Invasion
Motility within and on the cell	Surface transport phenomena  Distribution and movement of surface molecules Capping-like phenomena of antibodies Endocytosis and exocytosis The flagellum pocket, directed access and egress
Attachment	Salivary glands
Differentiation	Epimastigote: metacyclic differentiation
Recognition	Surface receptors: parasite/host, parasite/parasite and parasite/environment

<sup>a</sup> The flagellum is much more than an organelle for cell motility. This table defines some of the known and conjectured functions that I argue will apply to the *Trypanosoma brucei* flagellum.

LeBowitz's lab (Maga and LeBowitz, 1999; Maga et al., 1999; Santrich et al., 1997), on the same structure in *Leishmania*, have proven to be useful in defining a motility function for this structure and providing insight into the other processes that occur during morphogenesis of the flagellum. Our laboratory's initial aim of trying to use an antisense approach to understand PFR function led us to isolate mutants whose phenotype has been very instructive (Bastin et al., 1998). However, our studies of the molecular explanation of the phenotype provided us with an understanding of gene-silencing technologies that we have been able to extend and exploit for other studies of gene function in trypanosomes. This technology is likely to prove of great use in *T. brucei*, especially when coupled with inducible expression systems.

Aspects of the PFR biology and biochemistry have been reviewed recently and are illustrated in Fig. 6. The major components of the *T. brucei* PFR are two proteins, PFRA and PFRC. The PFRA protein is encoded by a set of four tandem genes, with two such clusters being present in the diploid genome. Some years ago, we set out to use antisense technologies to attempt to downregulate the expression of PFRA protein, and so, affect the structure and function of the PFR itself. PFRA antisense constructs expressed from 'ectopic' sites in the genome produced no phenotype in the resulting trypanosomes. However, in one such experiment, we identified a rare clone (Bastin et al., 1998). From one well in a multi-well plate from a transfection, we isolated the clone which exhibited a very strong phenotype. It was viable, but almost completely paralysed. Biochemical analysis revealed the expression of the PFRA RNA and protein was essentially ablated and electron microscopy and immunofluorescence microscopy revealed that the Sn11 mutant lacked most of the PFR structure (Fig. 7; Bastin et al., 1998).

Molecular analysis revealed that in this mutant clone, the

antisense construct had inserted in the genome in one of the two PFRA gene clusters. These experiments revealed that, given particular expression products, effective gene-silencing effects could be achieved in trypanosomes. We conjectured that the effect could be indicative of 'antisense interference with early processing' (Bastin et al., 1998). Recently, the description of a phenomenon called RNAi has extended conventional views about antisense mechanisms (Bosher and Labouesse, 2000). Microinjection of a mixture of in vitro synthesised sense and antisense RNA in *C. elegans* was found to be much more efficient at blocking the expression of genes than antisense (or sense) alone. This presence of double-stranded (ds) RNA leads to a specific, potent and rapid degradation of the corresponding mRNA. Mutants incapable of RNAi have recently been identified and indicate the existence of a complex pathway (see for instance: Bass, 2000; Chuang and Meyerowitz, 2000; Grishok et al., 2000; Ketting and Plasterk, 2000). RNAi initiated by exogenously synthesised dsRNA has been demonstrated in several organisms, including trypanosomes (Ngo et al., 1998) where ablation of alpha tubulin expression by such means led to a specific phenotype. Analysis of these experiments clarified other rather cryptic experimental results, and again, showed that RNAi was an effective means of gene-silencing in trypanosomes. Our consideration of the genotype of the paralysed snl-1 mutant suggested a number of possibilities whereby expression of the antisense constructs from this genomic environment would provide possibilities for local dsRNA formation. First, the insertion of a strong promoter in the opposite orientation to normal (polycistronic) transcription of the PFRA genes, would provide conditions for in situ, overlapping expression of sense and antisense RNA. Alternatively (or additionally), the lack of correct processing signals in a tandem antisense PFRA transcript might lead to loop formation in the single RNA molecule (Bastin et al., 1998). These and other phenomena would be capable of forming local dsRNA with the potential to influence, through an RNAi effect, the transcripts from downstream genes in the PFRA cluster and those from the homologous cluster in this diploid organism (Fig. 8).

We tested and verified this hypothesis by the generation of a new cell line (*snl-2*) expressing an RNA containing linked copies of sense and antisense PFRA from a tetracycline-inducible promoter (Bastin et al., 1999a, 2000). This construct was engineered in an inducible vector. The induction of expression of this PFRA dsRNA reproduced PFRA ablation, disappearance of the PFR itself and cell paralysis (Fig. 9). The wild-type population phenotype was recovered upon removal of the inducer (tetracycline). These results indicate the usefulness of a heritable and inducible RNAi system for interrogation of gene function in parasitic protozoa (Bastin et al., 2000; Shi et al., 2000). This approach should be extremely useful in the post-genomics analyses referred to earlier.

The analysis of the Sn11 and Sn1 2 mutant phenotype has

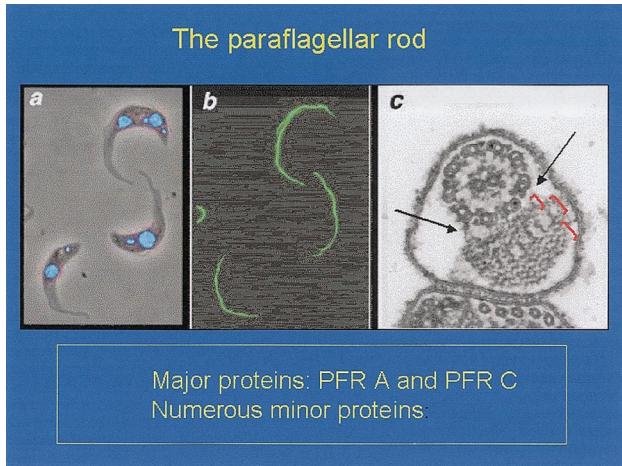


Fig. 6. Parts (a) and (b) show three *Trypanosoma brucei* procyclic cells at different stages of their cell cycle illustrating the growth of the paraflagellar rod: (a), phase contrast/DAPI; (b), anti-PFR immunofluorescence. Part (c) illustrates the various zones of the paraflagellar rod by transmission electron microscopy.

been very instructive in terms of the events involved in flagellum morphogenesis and function in kinetoplastid protozoa and eukaryotes in general (Bastin and Gull, 1999; Bastin et al., 1999b,c). In particular, these studies have highlighted the central role of protein targeting and intraflagellar transport systems (Kozminski et al., 1995; Rosenbaum et al., 1999). Analysis of the formation and removal of the blob of material at the end of the new flagellum in our trypanosome mutants appears to reflect the cell cycle-related action of an intraflagellar transport system (Bastin et al., 2000, 1999b, 1999c). This system, whereby there is directed retrograde and anterograde transport (Rosenbaum et al., 1999), is likely to have implications for the more general biology of kinetoplastid and other parasites. Construction of the main secretion and endocytotic area of the trypanosome — the flagellar pocket — cox-

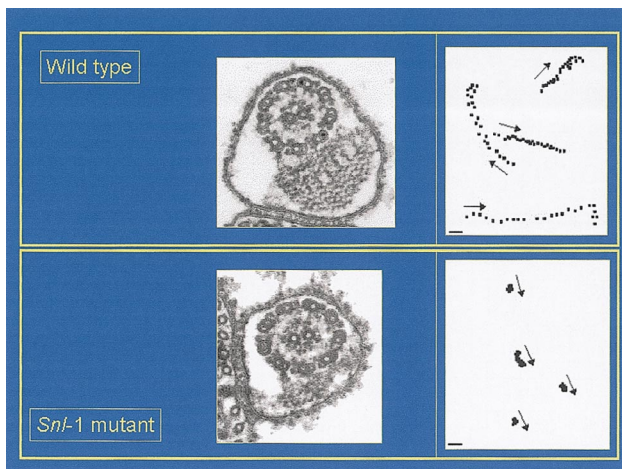


Fig. 7. A composite figure illustrating the ultrastructure of the flagellum and the movement of the cell across a microscope slide (or not!) for wild-type and a paralyzed mutant of *Trypanosoma brucei*.

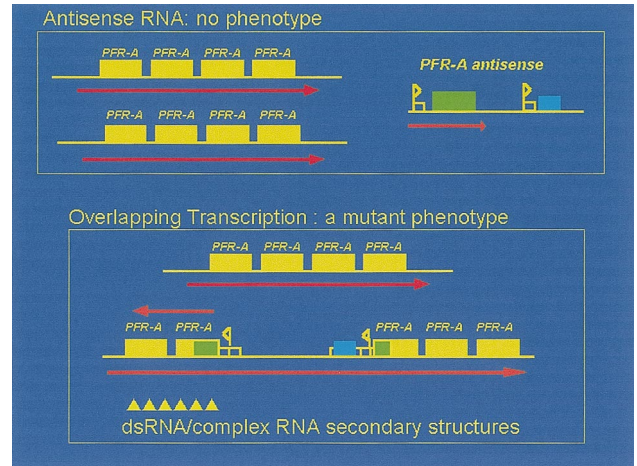


Fig. 8. Cartoon depicting the molecular phenotype of the *Snf1* mutant in the lower panel with some of the transcriptional consequences. The top panel denotes the situation in a large number of mutants which we have made where antisense is expressed, but from a site within the genome where it apparently does not produce the double stranded RNA required for the RNA interference effect.

ial with the flagellum basal area itself has major implications. It is likely that intraflagellar transport mechanisms might well also act to move surface and secreted molecules and to provide mechanisms for partition and distribution. The potential for bidirectional transport of cytoplasmic and membrane cargo by intraflagellar motor proteins is also likely to be mimicked in the cytoplasm. Our studies of microtubule polarity revealed a defined polar-

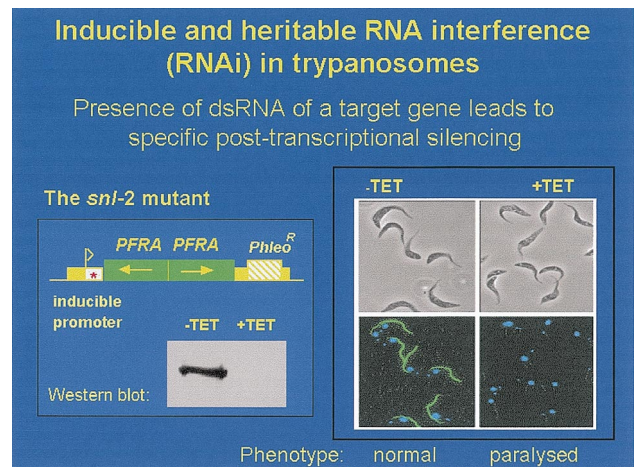


Fig. 9. A summary of the construct used to produce the inducible and heritable RNA interference phenomenon. The inducible construct is shown in the first panel, together with the results of the Western blot showing ablation of PFRA expression when the RNA interference effect is induced by the addition of tetracycline. The second panel shows the reversibility of this phenotype. In the absence of an RNA interference effect (no tetracycline, therefore no expression of double-stranded RNA for PFRA), the cells are normal and possess paraflagellar rods. In the presence of an RNA interference effect (plus tetracycline, therefore expression of double-stranded RNA for PFRA), the cells are paralyzed and do not possess paraflagellar rods as seen by immunofluorescence.



ity for the majority of the sub-pellicular microtubules and with a likely anti-parallel polarity for the four specialised microtubules nucleated close to the basal body/flagellar pocket (Robinson et al., 1995). The close association of the sub-pellicular microtubules and the four microtubules with specialised domains of the plasma membrane and endomembrane compartments suggests a role in directing membrane molecules. The particular polarities of these sets of sub-pellicular microtubules means that both plus-end and minus-end microtubules motors have the potential to act to influence the transport of surface proteins. Such interactions with the inner face of the plasma membrane and endomembrane compartments could be important in directing VSG (Variant Surface Glyco protein) molecules and receptor molecules both to and from the flagellar pocket domain. Also, the well known phenomena that antibody can be capped on the surface of trypanosomes and that the flagellum pocket is intimately involved in antibody clearance phenomena suggest a role for molecular motors in directing movements of surface molecules and antibody complexes (Table 6). The presence of at least three sets of microtubules with the potential to direct molecular motors towards the pocket area suggests that the architecture for such microtubule/motor/internal and external surface molecule movements does exist in trypanosomes. The use of paralysed mutants of the bloodstream forms of trypanosomes is likely to be of some interest in these studies and we are currently attempting to make these. We have also identified, cloned and sequenced a number of putative kinesin/dynein genes, and our analysis of the partial genome sequence databases using consensus sequences for microtubule motor proteins shows that there are still larger numbers of kinesins and dyneins encoded within the *T. brucei* genome. This fascinating area of microtubule/membrane biology will be another aspect of kinetoplastid biology to benefit from genomics and post-genomics.

## 9. The challenge of the orphan genes

Finally, I will return to the challenge of the orphan genes. Earlier, I rehearsed the likelihood that the kinetoplastid genomes would follow the pattern of revealing a vast proportion of genes with no known function or homologues. This is the missing biology! I have outlined some of my thoughts on what functions might be orchestrated by these genes and gene products in Table 4. I have pointed out elsewhere my prediction that included in this gene set, amongst protozoan parasites and other microbes, will be a large number of genes involved in what I have termed the 'slow growth' or 'no growth' state. I argue that for the last 200 years, most of our studies of microbes, both pathogenic and free-living, have relied on assays that demand growth of the organism. With notable exceptions, we have not asked about the no growth or very slow growth condition. No doubt, many parasites and microbes in their natural envi-

ronment spend a great deal of their time in this condition. Thinking of assays that involve or directly interrogate this condition is likely to be both difficult and rewarding. This, and the other read-outs which I have mentioned, mean that genomics and post-genomics of kinetoplastid parasites will certainly bring new insights into the biology of these organisms, but will also offer a whole new set of challenges.

## Acknowledgements

The author would like to thank past and present members of his laboratory for stimulating discussions and the Wellcome Trust and the BBSRC for funding.

## References

- Bass, B.L., 2000. Double-stranded RNA as a template for gene silencing. *Cell* 101, 235–8.
- Bastin, P., Gull, K., 1999. Assembly and function of complex flagellar structures illustrated by the paraflagellar rod of trypanosomes. *Protist* 150, 113–23.
- Bastin, P., Sherwin, T., Gull, K., 1998. Paraflagellar rod is vital for trypanosome motility. *Nature* 391, 548.
- Bastin, P., Ellis, K., Kohl, L., Gull, K., 1999a. Expression of an inverted repeat of a target gene leads to specific disappearance of the corresponding mRNA and encoded protein. *Mol. Biol. Cell* 10, 599.
- Bastin, P., MacRae, T.H., Francis, S.B., Matthews, K.R., Gull, K., 1999b. Flagellar morphogenesis: protein targeting and assembly in the paraflagellar rod of trypanosomes. *Mol. Cell. Biol.* 19, 8191–200.
- Bastin, P., Pullen, T.J., Sherwin, T., Gull, K., 1999c. Protein transport and flagellum assembly dynamics revealed by analysis of the paralysed trypanosome mutant *snl-1*. *J. Cell Sci.* 112, 3769–77.
- Bastin, P., Ellis, K., Kohl, L., Gull, K., 2000. Flagellum ontogeny in trypanosomes studied via an inherited and regulated RNA interference system. *J. Cell Sci.* 113, 3321–8.
- Bosher, J.M., Labouesse, M., 2000. RNA interference: genetic wand and genetic watchdog. *Nat. Cell Biol.* 2, E31–6.
- Chuang, C.F., Meyerowitz, E.M., 2000. Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 97, 4985–90.
- Dutcher, S.K., Trabuco, E.C., 1998. The UNI3 gene is required for assembly of basal bodies of *Chlamydomonas* and encodes delta-tubulin, a new member of the tubulin superfamily. *Mol. Biol. Cell* 9, 1293–308.
- El-Sayed, N.M., Hegde, P., Quackenbush, J., Melville, S.E., Donelson, J.E., 2000. The African trypanosome genome. *Int. J. Parasitol.* 30, 329–45.
- Ersfeld, K., Gull, K., 1997. Partitioning of large and minichromosomes in *Trypanosoma brucei*. *Science* 276, 611–4.
- Ersfeld, K., Asbeck, K., Gull, K., 1998. Direct visualisation of individual gene organisation in *Trypanosoma brucei* by high-resolution in situ hybridisation. *Chromosoma* 107, 237–40.
- Ersfeld, K., Melville, S.E., Gull, K., 1999. Nuclear and genome organization of *Trypanosoma brucei*. *Parasitol. Today* 15, 58–63.
- Grishok, A., Tabara, H., Mello, C.C., 2000. Genetic requirements for inheritance of RNAi in *C. elegans*. *Science* 287, 2494–7.
- Gull, K., 1998. Divide and Rule – The African trypanosome and sleeping sickness. *Wellcome Trust Rev.* 7, 26–9.
- Gull, K., 1999. The cytoskeleton of trypanosomatid parasites. *Annu. Rev. Microbiol.* 53, 629–55.
- Gull, K., 2000. Genomics and post-genomics in parasitology: genome babble or a real opportunity? *Biochem. Soc. Trans.* 28, 14–8.
- Gull, K., Alsford, S., Ersfeld, K., 1998. Segregation of minichromosomes in

- trypanosomes: implications for mitotic mechanisms. *Trends Microbiol.* 6, 319–23.
- Ketting, R.F., Plasterk, R.H.A., 2000. A genetic link between co-suppression and RNA interference in *C. elegans*. *Nature* 404, 296–8.
- Kozminski, K.G., Beech, P.L., Rosenbaum, J.L., 1995. The *Chlamydomonas* kinesin-like protein Fla10 is involved in motility associated with the flagellar membrane. *J. Cell Biol.* 131, 1517–27.
- Maga, J.A., LeBowitz, J.H., 1999. Unravelling the kinetoplastid paraflagellar rod. *Trends Cell Biol.* 9, 409–13.
- Maga, J.A., Sherwin, T., Francis, S., Gull, K., LeBowitz, J.H., 1999. Genetic dissection of the *Leishmania* paraflagellar rod, a unique flagellar cytoskeleton structure. *J. Cell Sci.* 112, 2753–63.
- McDonagh, P.D., Myler, P.J., Stuart, K., 2000. The unusual gene organization of *Leishmania* major chromosome 1 may reflect novel transcription processes. *Nucleic Acids Res.* 28, 2800–3.
- Myler, P.J., Audleman, L., DeVos, T., Hixson, G., Kiser, P., Lemley, C., Magness, C., Rickel, E., Sisk, E., Sunkin, S., Swartzell, S., Westlake, T., Bastien, P., Fu, G.L., Ivens, A., Stuart, K., 1999. *Leishmania major* Friedlin chromosome 1 has an unusual distribution of protein-coding genes. *Proc. Natl. Acad. Sci. USA* 96, 2902–6.
- Ngo, H., Tschudi, C., Gull, K., Ullu, E., 1998. Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc. Natl. Acad. Sci. USA* 95, 14687–92.
- Oakley, C.E., Oakley, B.R., 1989. Identification of gamma-tubulin, a new member of the tubulin superfamily encoded by *mipA* gene of *Aspergillus nidulans*. *Nature* 338, 662–4.
- Robinson, D.R., Gull, K., 1991. Basal body movements as a mechanism for mitochondrial genome segregation in the trypanosome cell cycle. *Nature* 352, 731–3.
- Robinson, D.R., Sherwin, T., Ploubidou, A., Byard, E.H., Gull, K., 1995. Microtubule polarity and dynamics in the control of organelle positioning, segregation, and cytokinesis in the trypanosome cell cycle. *J. Cell Biol.* 128, 1163–72.
- Rosenbaum, J.L., Cole, D.G., Diener, D.R., 1999. Intraflagellar transport: the eyes have it. *J. Cell Biol.* 144, 385–8.
- Santrich, C., Moore, L., Sherwin, T., Bastin, P., Brokaw, C., Gull, K., LeBowitz, J.H., 1997. A motility function for the paraflagellar rod of *Leishmania* parasites revealed by PFR-2 gene knockouts. *Mol. Biochem. Parasitol.* 90, 95–109.
- Scott, V., Sherwin, T., Gull, K., 1997.  $\gamma$ -Tubulin in trypanosomes: molecular characterisation and localisation to multiple and diverse microtubule organising centres. *J. Cell Sci.* 110, 157–68.
- Sherwin, T., Gull, K., 1989. The cell division cycle of *Trypanosoma brucei*: timing of event markers and cytoskeletal modulations. *Philos. Trans. R. Soc. Lond. Biol.* 323, 573–88.
- Shi, H.F., Djikeng, A., Mark, T., Wirtz, E., Tschudi, C., Ullu, E., 2000. Genetic interference in *Trypanosoma brucei* by heritable and inducible double-stranded RNA. *RNA* 6, 1069–76.
- Vaughan, S., Attwood, T., Navarro, M., Scott, V., McKean, P., Gull, K., 2000. New tubulins in protozoal parasites. *Curr Biol.* 10, 258–9.
- Woodward, R., Gull, K., 1990. Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*. *J. Cell Sci.* 95, 49–57.